# THE CENTRALISATION OF INVENTORY

# AND THE MODELLING OF DEMAND

John Edward Boylan

A thesis submitted in accordance with the regulations concerning admission

to the degree of Doctor of Philosophy.

Warwick Business School

University of Warwick

March 1997

# CONTENTS

# DETAILED CONTENTS

## INTRODUCTION

# LIST OF TABLES

## LIST OF FIGURES

# GLOSSARY

| | |
|---|---|
| Amalgamation | An umbrella term for 'centralisation' and 'consolidation' (defined below). |
| Centralisation | Amalgamation of inventories from more than one depots to one depot. |
| Complete Line Fill | Proportion of total lines which are filled completely (notation - $L_c$) |
| Complete Order Fill | Proportion of total orders which are filled completely (notation - $O_c$) |
| Consolidation | Amalgamation of inventories from more than one depot to fewer depots but not to one depot. |
| Cycle stock | Those stocks held to achieve economies of scale in ordering and production (notation - CS). |
| Demand | Demand occurs when a customer orders a product, regardless of whether the demand is completely or partially satisfied or not satisfied at all. The total number of units ordered is defined as 'demand'. |
| Demand Incidence | The total number of order occasions is defined as 'demand incidence'. |
| Economic Order Quantity (EOQ) | Amount ordered to satisfy a cost minimisation function which takes into account ordering and holding costs. |
| Inter-correlations | Correlations between demands on depots from different sub-consolidations. |
| Intra-correlations | Correlations between demands on depots from the same sub-consolidation. |
| Joint demand model | A model of demand which reflects variations in demand for an SKU attributable to both individual SKU effects and joint effects (such as market fluctuations) which affect many SKUs. |
| Operationalisation | Translation of abstract concepts into measures enabling observations to be made. |
| Order line | An order for a particular SKU. |
| Order size | The number of units demand of a particular SKU in a single order. |

| | |
|---|---|
| Partial Line Fill | Proportion of total lines which are filled (at least) partially (notation - $L_p$). |
| Partial Order Fill | Proportion of total orders which are filled (at least) partially (notation -$O_p$). |
| PES | Pillar Engineering Supplies Ltd. |
| Portfolio Effect | Denotes the percentage saving in safety stock arising from centralisation or consolidation. |
| Portfolio Cost Effect | Reduction in cost (including holding cost, transportation cost, investment cost and procurement cost) arising from inventory amalgamation. |
| Portfolio Quantity Effect | Reduction in stock quantity arising from depot amalgamation. |
| Return on investment (ROI) | Ratio of profit before tax to total equity, including inventory. |
| ROQ | Order quantity which maximises ROI. |
| Safety stock | Those stocks held as a buffer against unforeseen fluctuations of demand (notation - SS). |
| Service level | A measure of the availability of stock to customers. |
| SKU | Stock keeping unit. |
| Square root law | The total stock in a system is proportional to the square root of the number of locations at which a product is stocked. |
| Super consolidation | An alternative name for 'centralisation'. |
| Sub consolidation | An alternative name for 'consolidation'. |
| Total stock | The total of cycle stock and safety stock (notation - TS). |
| Unified Order Quantity | An order quantity which optimises the total costs across the whole of the distribution chain. |
| Unit fill rate | Proportion of units ordered filled from stock (notation - U). |
| Value fill rate | Proportion of value of units ordered filled from stock (notation - V). |
| Variance law | A relationship between the variance and mean levels of demand, which holds across a group of items. |

ACKNOWLEDGEMENTS

There are many people who have helped and supported me throughout the period of my part-time PhD study but without the following the road would have been much harder.

I would especially like to thank Roy Johnston, my supervisor, for our meetings which I found to be unfailingly challenging, stimulating and encouraging. His advice has been much valued.

I am indebted to Alan Thompson and Stuart Coggins, of Pillar Engineering Supplies, for helping me to obtain data on demand histories of a large sample of stock keeping units. I am also grateful to my employers at Buckinghamshire College for financial support and reduction in teaching load, particularly during the sabbatical semester, which has been invaluable in allowing me to complete the thesis.

Finally, I would like to thank Jan, my wife, for all her love and encouragement throughout the time of my PhD study and for proof-reading the thesis.

## DECLARATION

Thirteen of the fifteen chapters in this thesis have not been published before. Two chapters, in slightly different form, have been published as journal articles. The details are given below.

* Chapter 7 is based on :

  BOYLAN, JE and JOHNSTON, FR (1994) Relationships between service level measures for inventory systems. *J Opl Res Soc*, **45**, 838-844.

* Chapter 12 is based on :

  BOYLAN, JE and JOHNSTON, FR (1996) Variance laws for inventory management. *Int J Production Economics*, **45**, 343-352.

Both articles are the product of original research by the first-named author. The second-named author was included in recognition of his constructive comments and criticism, acting in his capacity as PhD supervisor.

John Edward Boylan

For Jan Jan, with love.

# SUMMARY

The motivation for this research arose from two projects in which the author advised on inventory centralisation. Since it was found that the literature was of limited value, inventory centralisation was identified as a suitable topic for research.

Operationalisation is not adequately addressed in the literature on centralisation models. 'Operationalisation' denotes the translation of abstract concepts such as 'inventory service' into measures enabling observations to be recorded. In the literature, 'inventory service' is often equated to 'probability of stock-out' but many other measures are used in practice. This thesis presents a network of relationships between six commonly used measures. This is useful when decentralised depots do not share a common measure, or when the measure changes after centralisation.

This thesis argues that, under all circumstances which arise in practice, it will be possible to achieve inventory availability benefits from centralisation with no added investment in stocks. A counter-example to the universal application of this rule have been presented in the literature but it is shown that such counter-examples are artificial.

Since savings from centralisation may be offset by increases in transport costs, the reduction in stock-holdings needs to be estimated. The models presented in the literature assume that the estimation of demand variance and the correlation of demand between depots is not problematic. In practice, reliable estimates may be difficult to obtain, particularly for slow-moving items. Consequently, it is difficult to decide which items should be centralised and which, if any, should not. This thesis proposes a 'quadratic variance law' approach', linking the variance of demand to its mean. This approach is underpinned by a model which allows correlation effects to be taken into account. The 'variance law' approach is a contribution towards the operationalisation of centralisation models, since reliable estimates of mean demand are easier to obtain than estimates of variance and correlation.

The 'quadratic variance law' is examined empirically using a sample of 230 stock-keeping units from an engineering supplies company. The approach is shown to be well-supported by the data. All the model assumptions are supported except one. The postulated independence between mean order-size and mean demand is rejected since a weak correlation was found. However, the simpler quadratic law is found to be more robust than a more complex law which would have taken this correlation into account.

# INTRODUCTION

In this introductory section of the thesis, four issues will be addressed :

1. The importance of inventory centralisation to organisations.
2. The principal shortcomings of published models and criteria for 'good' centralisation models.
3. The structure of the thesis.
4. The methods used in the research undertaken.

This material will provide a background to the research and the main arguments to be developed. A summary of the research findings is given in chapter 15.

## 0.1 Business context

Logistics is a matter of strategic importance for organisations since it can contribute to competitive advantage by cost reduction and service improvement. In this section, the role of inventory centralisation in achieving competitive advantage is discussed.

### 0.1.1 Benefits of inventory centralisation

The first reason why inventory centralisation is of importance is that substantial cost savings may be achieved. Such savings are predicted by theory, to be reviewed in part I of the thesis. More importantly, substantial savings have been observed in practice. For example, Abrahamsson (1993) describes the case of Atlas Copco Tools AB, manufacturers and distributors of hand-held pneumatic tools. From a complex system of central, national and regional warehouses, the company centralised their stock-holdings to one location in southern Sweden (later relocated to Belgium). The centralisation reduced the value of inventory holding from 29% to 20% of sales. Substantial reductions in warehousing costs were also achieved. Abrahamsson claims that, despite the introduction of some air freight, transportation costs remained constant.

Although the above discussion is limited to one company, Evans (1993) found that 25% of respondents to a KPMG survey cited 'centralisation of distribution' as a 'response to market pressures'. The sample for the KPMG survey consisted of 153 directors of leading European companies.

The second reason why inventory centralisation is important is that a more consistent level of inventory service may be provided to all customers. If stocks are de-centralised, then different availability levels may be observed at the depots. It is even possible that the same availability may be achieved according to a formal measure and yet a worse service may be offered to some customers. For example, in a study of European distribution of Austin Rover parts, Eastburn and Nockold (1986) found that, *"the P&A stock control system ... is wilfully (and cleverly) abused in Bologna where they stock-pile the very cheap parts in preference to the very expensive parts in order to achieve the target 92% availability with as little investment as possible. Hence, they would be able to supply the nuts and bolts to attach a new alternator, but could be out of stock of the alternator itself."* If, on the other hand, stocks are centralised, then the company service policy can be implemented more straightforwardly and all customers may be guaranteed the same level of service.

### 0.1.2 Implementation of centralisation

The first step in examining centralisation options is to identify customer lead-time requirements. If customers require delivery within 48 hours of order submission, then a detailed analysis of lead-times will enable the minimum number of stock-holding points to be identified.

Lead-times may be split into two components : physical lead-time and administrative lead-time. Within the EU, the relaxation of border controls and the opening of the channel tunnel have contributed towards shorter physical lead-times. More significantly, advances in information systems have enabled substantial reductions in administrative lead-times to be achieved. For example, Taylor (1991) describes how Unipart Group Ltd achieved significant improvements in its delivery service to parts wholesalers. The new service was based on daily receipt of stock-orders and provided a reduction in turn-round time from receipt of order to delivery from five working days to just one working day. Implementation of software allowing wholesalers to interface directly with the Unipart mainframe computer enabled direct transmission of orders. The new system reduced preparation time by wholesalers and eliminated the need to input orders from hard copy at Unipart, thereby cutting the order cycle time. Reductions in order-cycle times such as these have given many companies greater scope for centralisation.

In conclusion, centralisation of inventories has been achieved by many organisations, is a feasible approach for many others, and may yield significant cost reductions and service benefits.

## 0.2 Modelling of inventory centralisation

### 0.2.1 The need for inventory centralisation models

An inventory centralisation model enables a manager to obtain estimates of the reduction in stock-holding which would result from centralisation. This is useful since it provides a forecast of the capital which will be freed for other activities. In the Atlas Copco case, for example, it was intended to use the money saved by inventory reductions to boost the marketing operations of the national sales companies.

Inventory centralisation models may also be used to compare different options for inventory amalgamation. Abrahamsson (1993) argues that such models are unnecessary and a 'time based distribution' approach should be used: *"... the lead-time is the only variable to consider to calculate the number of warehouses."* This assumes that the total cost curve always rises as the number of warehouses increases. This assumption is valid if transport costs do not rise as a result of centralisation, as in the case of Atlas Copco.

At Atlas Copco, centralisation resulted in a large number of small deliveries to a large number of customers. In these circumstances, centralised transport costs should be greater than in the decentralised scenario. One possible explanation for the constancy of transport costs is that, since inventory availability was very poor at the decentralised depots (only 70%), significant costs of expediting were saved by centralisation, offsetting the increases in delivery costs. If higher availabilities were attained at the decentralised depots, then transport costs may have been considerably higher after centralisation.

This discussion highlights the problems in generalising the findings from a case-study. Since a number of variables change after centralisation, it is not possible to claim that total logistics costs always reduce as a result of inventory centralisation. A deeper understanding is afforded by modelling the effect of centralisation on inventory, warehousing and transport costs in each individual case. Then, together with an assessment of the impact on inventory availability, a more informed decision may be made.

### 0.2.2 Weaknesses of published centralisation models

A detailed review and critique of published models is given in part I of the thesis. Three issues, in particular, have been under-researched in the literature :

1. *The measurement of service levels.*

   It is generally assumed in the inventory centralisation literature that 'probability of stock-out during lead-time' is the measure used. However, many organisations use other measures such as line-fill or order-fill. For centralisation models to be useful in these circumstances, or in situations where different measures are employed at different decentralised depots, a set of relationships between service levels is required.

2. *The estimation of demand variance.*

   All of the published models refer to demand variance, but few discuss how this variable should be estimated. This may be a problem for slow-moving items or for items which experience a step-change in demand. Since slow-moving stock is at particular risk of obsolescence, centralisation is an option which should be considered for this category of inventory. More accurate estimation of demand variance would allow a better assessment to be made of which SKUs should be centralised and which should not.

3. *The estimation of demand covariance.*

   Again, few published models discuss the estimation of demand covariance between depots. The same estimation difficulties occur in the circumstances highlighted for demand variance. If the effect of demand covariance could be estimated more accurately, this would enable the assessment of stock centralisation to be improved when there are some correlations between demands at decentralised depots.

### 0.2.3 Criteria for 'good' centralisation models

A 'good' inventory centralisation model satisfies a number of criteria. Firstly, the model assumptions should not be too restrictive. Early publications assumed equal mean demands at all decentralised depots and no correlation between demands. Both of these assumptions are unlikely to hold in practice.

The second criterion is that the model assumptions should be clearly stated. For example, some users of the 'square root law', to be discussed in chapter 2, may be unaware of the assumptions behind the 'law'. Not all publications on this 'law' give a comprehensive list of model assumptions.

The third criterion is that the model is built upon sound theory and does not rely on empirical evidence alone. (Naturally, any theoretical model should be tested against empirical evidence). In particular, the structure of demand must be understood in order to assess the effect of inventory centralisation on demand variability. Without such a structure, 'what if' analyses cannot be performed.

The final criterion is that the method of estimation or measurement of all variables included in the model formulation is specified. Without such a specification, a model is of theoretical interest only.

## 0.3 Thesis structure

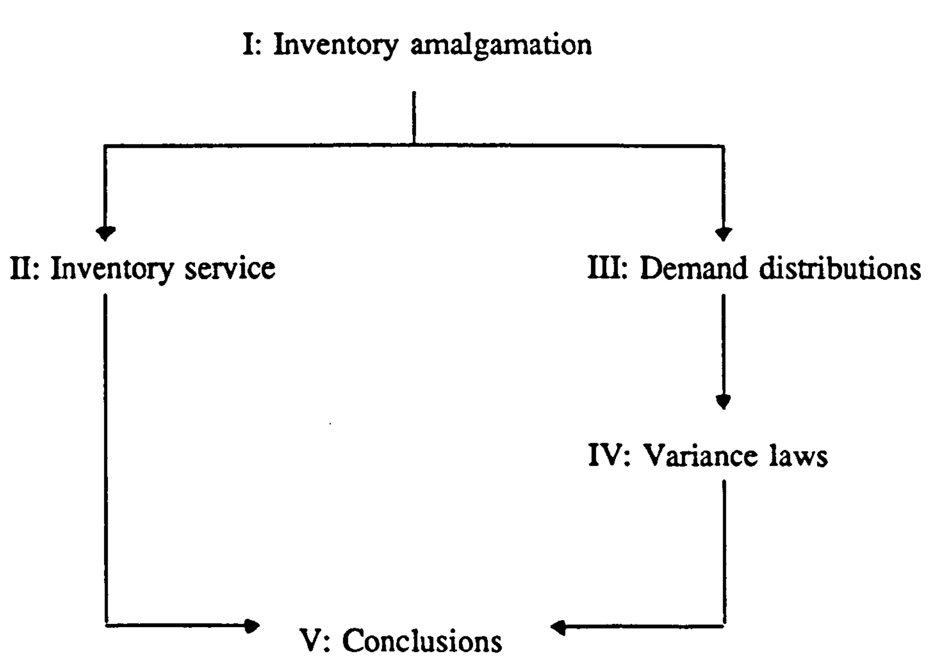
### 0.3.1 Overall structure

The thesis is divided into five parts, namely :

| | |
|---|---|
| Part I | Inventory amalgamation models |
| Part II | Inventory service models |
| Part III | Demand distribution families |
| Part IV | Demand variance laws |
| Part V | Conclusions |

The logical interdependence of the thesis components is summarised in figure 0.1 :

Figure 0.1
Logical interdependence of the five parts of the thesis



### 0.3.2 Inventory amalgamation models

This part of the thesis contains a detailed literature review and critique of previously published models. The models are assessed with particular reference to the clarity of their stated assumptions. Where necessary, assumptions are corrected or expanded, with the aid of a classification system for inventory amalgamation models. In some instances, the models are corrected for mathematical errors.

It is found that, although early models had quite restrictive assumptions, later generalisations have overcome many of these limitations. However, the three issues of service measures, variance estimation and covariance estimation, mentioned in sub-section 0.2.2, are identified as requiring further research.

### 0.3.3 Inventory service models

Inventory service models are considered from two perspectives in part II of the thesis. Firstly, the effect of centralisation on service is examined. The conditions under which service benefits occur are delineated. It is found that the conditions under which disbenefits will result are unlikely to occur in practice. These findings are conjectural since complete mathematical proofs have not yet been obtained.

The second perspective from which service models are viewed is the measurement of service itself. Many different measures are used in practice and it may happen that measures vary at decentralised depots or will change after centralisation. To overcome these difficulties, a 'network of relationships' between six commonly used measures is presented.

### 0.3.4 Demand distribution families

The third part of the thesis, on demand distribution families, is a precursor to the fourth part of the thesis, on demand variance laws, as shown in figure 0.1.

It is argued that many proposed distributions lack any underpinning theory to justify their use as demand distributions; the gamma is an example of such a distribution. Some proposed distributions are examined from a theoretical standpoint and possible underlying mechanisms are suggested as appropriate. It is shown that the compound Poisson family of distributions possesses attractive theoretical properties for demand modelling.

Estimation issues are considered and a number of demand incidence and demand distributions are tested against empirical demand data on a sample of 230 SKUs from Pillar Engineering Supplies Ltd. It is found that Poisson demand incidence is well supported by the data, as is negative binomial demand.

### 0.3.5 Demand variance laws

The fourth part of the thesis follows on from part III, by establishing demand variance laws based upon compound Poisson demand. The variance of demand is shown to be a quadratic function of the mean demand under certain assumptions.

Using the Pillar Engineering Supplies data, each of the assumptions for a quadratic variance law are tested. Three assumptions are found to hold very strongly. The fourth assumption, that mean order size is independent of mean demand, is not found to hold as there is a weak correlation between mean order-size and the square root of mean demand. However, the quadratic variance law is found to be more appropriate for the data than a more complex expression which includes an allowance for the correlation between mean order-size and mean demand.

### 0.3.6 Conclusions

In the final part of the thesis, the conclusions are summarised, limitations of the research are identified and avenues for further research are outlined.

One of the main conclusions of the research is that a quadratic variance law may be used to estimate demand variance. The variance law allows the effect of demand covariance to be taken into account automatically.

The variance law predicts that as the mean demand rises, the proportional stock-savings from centralisation decrease. The predictions allow a comparison with marginal increases in transport costs to be made, thus identifying those SKUs which should be centralised and those which should not.

## 0.4 Methodology

### 0.4.1 A deductivist approach

The thesis is characterised by a deductivist approach. Theories of inventory amalgamation, inventory service, demand distributions and demand variance 'laws' have been deduced from various assumptions. The assumptions have been examined in two different ways. Firstly, they have been checked for their general plausibility: do they make sense in an inventory context ? Secondly, they have been tested empirically: are they supported by data from a real organisation ?

An alternative approach is to analyse inventory data before and after centralisation and to infer some general principles from the observations. Abrahamsson's (1993) 'time based distribution' principle, discussed in sub-section 0.2.1, is a good example of the dangers of the inductive approach. After depot centralisation, a number of variables may change in addition to the number of inventory locations. In Abrahamsson's case, the level of inventory availability increased markedly. This may have affected the transport costs and casts doubt on the general proposition that total logistics costs always decrease as the number of inventory locations decrease.

In contrast to the inductive approach, this thesis is 'scientific' in the sense of Popper (1959). According to Popper, there are a number of stages in scientific research. The first stage is the formulation of theory or a hypothesis. The second stage, called 'operationalisation', is the translation of abstract concepts into indicators or measures, enabling observations to be made. The third stage is the empirical testing of the theory; if unfalsified, the theory stands as not yet unproven. Much of the literature on inventory amalgamation has not progressed beyond the first stage of Popper's method. The purpose of this thesis is to 'operationalise' centralisation models and to test the underpinning theory.

### 0.4.2 Service level measures

A set of relationships between inventory service measures is presented in chapter 7. Since the relationships are algebraic links of *definitions*, they are tautological. However, they have been formulated in such a way that only those variables most likely to be available in practice have been included. Thus, the relationships were operationalised. Empirical testing is not relevant in this case since the algebraic formulae are tautological.

### 0.4.3 Demand variance laws and demand distributions

Centralisation models in the literature have paid little attention to operationalisation. Variance estimates are required but may be difficult to obtain for slow-moving items, newly launched products or those which have experienced a step-change in demand. Variance laws, linking the variance of demand to its mean, provide a neat operationalisation, since accurate estimates of the mean are easier to obtain. The theory of such variance laws is underpinned by demand distribution assumptions. This necessitates empirical testing at two levels. The demand distribution models must be tested empirically and so must the variance laws themselves. Both sets of tests have been conducted in the thesis.

### 0.4.4 Amalgamation models

The amalgamation models presented in chapter 14 are based on the theories of compound Poisson demand and a quadratic variance - mean demand relationship. The quadratic variance law provides an operationalisation of the amalgamation models. Although both the compound Poisson distribution and the quadratic variance law have been tested empirically, the amalgamation models have not been subject to empirical testing. This represents the next stage of research after completion of the PhD thesis.

# PART I

# INVENTORY AMALGAMATION MODELS

## Summary of Part I

In this part of the thesis, a literature review and critique of previous work on inventory amalgamation is provided. The term 'amalgamation' covers both 'centralisation' to one depot and 'consolidation' to more than one depot.

In the first chapter, a classification system for inventory amalgamation models is developed. The system is used as a framework throughout the literature review. The classification provides a structured approach to the identification of the main model assumptions. This is an important issue in amalgamation modelling since such clarity has not always been achieved in this field of research.

Chapters 2 and 3 contain a detailed literature review and critique of safety stock amalgamation models. Work by Zinn, Levy and Bowersox (1989) on this problem is reviewed and, using the classification system, shown to be mis-categorised as a cost-driven model when it is actually a service-driven model. Some corrections to previously published papers are given, notably the preference rule between different types of consolidation (Mahmoud (1992)) and the centralisation model for variable lead-times (Tallon (1993)).

Throughout chapters 2 and 3, it is found that issues of estimation and measurement have been neglected in the inventory amalgamation literature. The question of service level measures has been the subject of some discussion (Ronen (1982) and Zinn, Levy and Bowersox (1989)). It is argued that this matter requires further analysis and is the subject of chapter 7 in the second part of the thesis.

Estimation of demand variance and covariance are highlighted as requiring further work. The models reviewed assume that these estimates will be readily available but there are many situations in which estimation may be difficult. This issue is examined theoretically and empirically in part IV of the thesis.

In chapters 4 and 5, cycle stock centralisation models are examined. Estimation issues are not discussed since estimation of mean demand is less problematic than estimation of demand variance and covariance. The robustness of the cycle stock square root law to unequal mean demands is examined. Also, in chapter 5, the robustness of the Economic Order Quantity (EOQ) approach is discussed, with reference to its application in centralisation modelling. It is shown that some of the limitations of the EOQ model at fixed locations do not apply to centralisation models, assuming parameters are unchanged by centralisation. However, the use of the alternative criterion of return on investment produces quite different results. Lower and upper bounds for the centralised to decentralised cycle stock ratio are presented for this case.

# CHAPTER 1

## *Classification of Inventory Models*

## 1.1 Introduction

### 1.1.1 Summary

In this chapter, classifications of inventory models will be reviewed. Classification systems have been designed for models of inventories at fixed locations but it will be argued that inventory amalgamation models require their own classification. The lack of such a classification is remedied by a new system, based on previous work on fixed location inventories but specifically adapted for amalgamation models. In the following three chapters, the new classification system will be applied to two types of inventory amalgamation: 'centralisation', where the inventories are amalgamated in one location, and 'consolidation', where the inventories are amalgamated in more than one location.

### 1.1.2 Reasons for classification

Before proceeding to the detailed review and proposal, it is worth reflecting on why a classification system is needed for inventory amalgamation models. There are two main reasons. First, such a system would help the modeller appreciate the place of each model within the overall body of research. The case for such a classification system for fixed-location inventory models is overwhelming, since many hundreds of models have been published. The case for amalgamation models may seem weaker as there are fewer published models. However, it will be shown that some researchers have incorrectly placed their own work within the discipline and it will be argued that greater clarity of assumptions is needed. One way of achieving this is by adopting a formal classification system which encapsulates essential model assumptions.

The second reason for investigating a classification system is to help identify those aspects of amalgamation modelling which have been under-researched and which may benefit from further investigation. Examples which will be highlighted in this chapter include 'Gross Models', 'Multiple Criteria Decision Models' and 'Multi-Echelon Models'. Although these issues are outside the scope of this thesis, they have emerged from the development of the classification system as areas for further research.

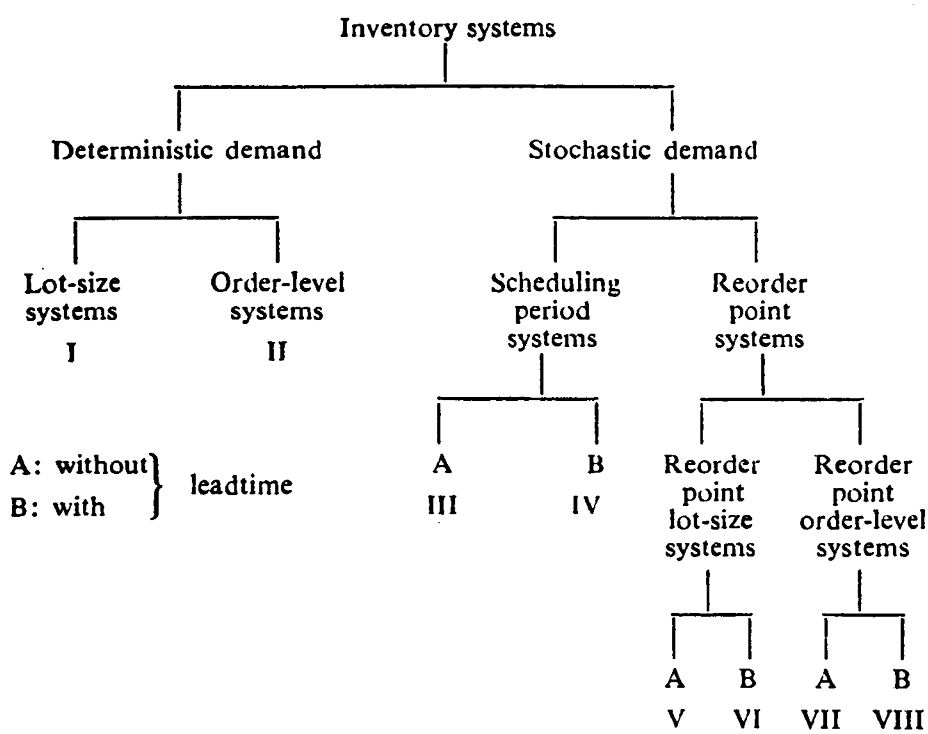
## 1.2 Review and critique of classification systems

Clearly, there is no definitive way in which inventory models should be classified and a number of authors have produced systematic classification systems of varying degrees of complexity. Three publications in this area are reviewed in this section : Naddor (1966), Aggarwal (1974) and Chikán (1990).

### 1.2.1 Naddor's classification

Naddor's (1966) classification of inventory models is shown in Figure 1.1 :

Figure 1.1
Naddor's classification system

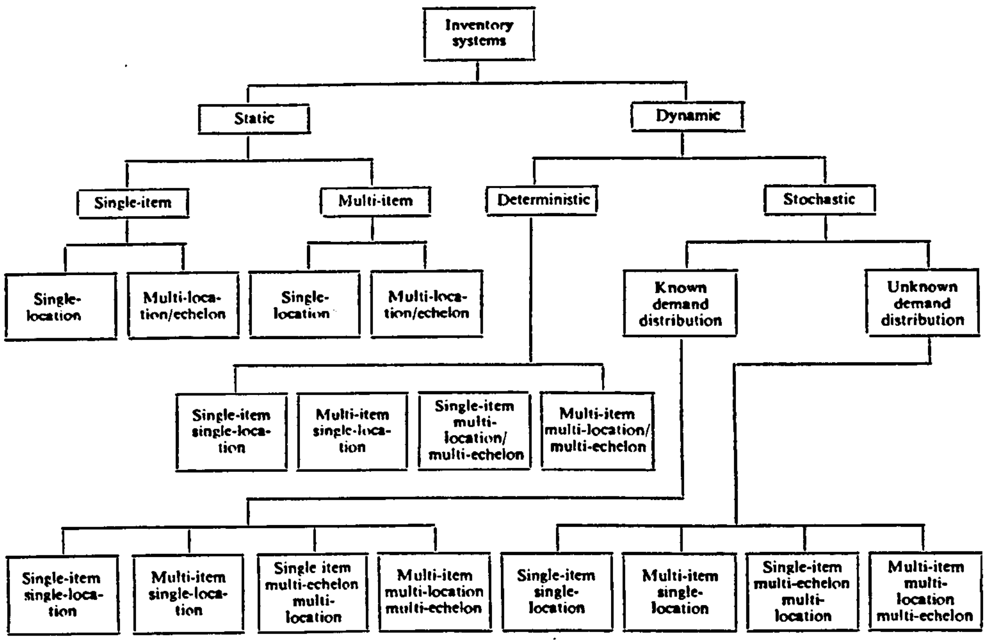



Naddor's classification system is quite rudimentary. It ignores three issues which are important in fixed-location and amalgamation models alike, namely : the system objective (cost-driven or service-driven), the nature of lead-time (deterministic or stochastic) and the multi-depot situation.

### 1.2.2 Aggarwal's classification

Aggarwal (1974) produced an alternative classification system, shown in Figure 1.2 below, based on similar principles to the work of Naddor :

Figure 1.2
Aggarwal's classification system




Aggarwal's classification includes explicit references to the depot-system : single depot-location and multi depot-location cases are shown and multi-echelon systems are considered. However, the other criticisms of Naddor's system also apply to Aggarwal's classification. The system objective (cost-driven or service-driven) and the nature of lead-time (deterministic or stochastic) are not considered.

### 1.2.3 Chikán's classification

A more comprehensive classification of fixed-location inventory models was given by Chikán (1990). Chikán uses ten 'main codes' (MC1 to MC10) to categorise inventory models :

*MC1 Number of items stored in the system*

| | | |
|---|---|---|
| Code 1 | : | one item stored |
| Code 8 | : | several items (ie more than one) stored |

*MC2 Number of stores in the system*

| | | |
|---|---|---|
| Code 1 | : | one store in the system |
| Code 8 | : | several stores in the system |

*MC3 Character of the inflow (input) process*

| | | |
|---|---|---|
| Code 0 | : | input is deterministic |
| Code 1 | : | input is stochastic |

Input is said to be deterministic if *all* elements of the input process are known with certainty. According to Chikán's definition, if any uncertainty exists, then the inflow is considered to be stochastic.

*MC4 Character of the outflow (output) process*

| | | |
|---|---|---|
| Code 0 | : | output is deterministic |
| Code 1 | : | output is stochastic |

The same considerations apply as to inputs. The deterministic or stochastic character of demand determines the character of the outflow.

*MC5 Way of treating time in the system ('dynamics' of the system)*

| | | |
|---|---|---|
| Code 0 | : | the model is static |
| Code 1 | : | the model is dynamic |

Chikán defines a model as 'dynamic' if the objective function over a whole planning period depends in its general form, at any given time, on the values of the same objective function taken at some other time. Otherwise, the model is said to be 'static'.

*MC6 The objective of the system*

| | | |
|---|---|---|
| Code 0 | : | optimisation model |
| Code 1 | : | reliability model |
| Code 2 | : | descriptive model |

The optimisation model category covers such criteria as cost minimisation, discounted cost minimisation and income maximisation. 'Reliability models' are driven by demand satisfaction objectives. Finally, 'descriptive' models describe the reaction of an inventory system to impulses from the 'external world'.

*MC7 Operation mechanism of the system (ordering rule)*

| | | |
|---|---|---|
| Code 0 | : | (t, S) |
| Code 1 | : | ($t_p$, S) |
| Code 2 | : | (s,q) or (s,$q_p$) |
| Code 3 | : | ($s_p$, q) |
| Code 4 | : | (t,q), ($t_p$, q) or (t, $q_p$) |
| Code 5 | : | (s, S) or (s, $S_p$) |
| Code 6 | : | ($s_p$, S) |
| Code 8 | : | (t, $S_p$) |

| | | |
|---|---|---|
| where t | - | fixed order-interval |
| s | - | re-order point |
| q | - | fixed order quantity |
| S | - | order-up-to point |
| p | - | fixed parameter |

A 'fixed parameter' is not reviewed at the same time that the stock-ordering decision is reviewed.

*MC8 Mode of review*

| | | |
|---|---|---|
| Code 0 | : | no inventory reviewing |
| Code 1 | : | periodic review |
| Code 2 | : | continuous review |

*MC9 Treatment of shortage*

| | | |
|---|---|---|
| Code 0 | : | shortage not allowed |
| Code 1 | : | shortage allowed, demand back-ordered |
| Code 2 | : | shortage allowed, lost sales |

Models with 'shortage not allowed' are relevant to situations with deterministic demand or stochastic demand constrained to a maximum value.

*MC10 Lead-time*

| | | |
|---|---|---|
| Code 0 | : | zero lead-time |
| Code 1 | : | constant lead-time |
| Code 2 | : | variable lead-time (but known in advance) |
| Code 3 | : | stochastic lead-time |
| Code 4 | : | fulfilment in more than one lot, or continuously |

A distinction is made between two cases where lead-times vary over time. Code 2 is used for situations when the lead-time is known and code 3 for when the lead-time is not known in advance.

In addition to the ten 'main codes' summarised above, Chikán devised a number of 'secondary codes', summarised in Appendix 1.1.

As an illustration of the main code system, suppose that there is a single store, single item, static optimisation model for a periodic (s,q) system. The model is based on deterministic inputs, stochastic demand, stochastic lead time and it is assumed that there will be lost sales if orders are unfilled. This model would be coded :

1 1 0 1 0 0 2 1 2 3

A criticism of this representation is that it is difficult to recognise a model represented by an array of ten numbers.

The most significant omission in Chikán's system is the classification of multi-echelon models. A secondary factor of 'connection between the stores' is included in the classification (see Appendix 1.1). Three codes are identified: no connection between the stores (code 0), linear connection between the stores (code 1) and parallel stores connected to a central store (code 2). However, this categorisation cannot capture the richness of multi-echelon inventory systems since the number of levels in the system is restricted to a maximum of two. Consequently, the arborescent structure of many multi-echelon systems cannot be represented within the classification.

Another omission in the classification is that no account is taken of the presence, or absence, of correlation between demands at different stores. At item-level, if lateral supply between depots is permitted, then this issue is relevant. In the context of inventory centralisation, demand correlation assumptions are always relevant.

In this section, the factors used in three classification systems have been examined. Since Chikán's system is the most comprehensive, it will be used as the basis for a classification of centralisation models. As some factors are omitted by Chikán, these will be incorporated as necessary.

## 1.3 Proposed classification of amalgamation models

In this section, Chikán's classification system will be used as a basis for developing a classification of amalgamation models. Factors which require coding for fixed-location models but not for amalgamation models will be removed. Factors not coded by Chikán but needed for amalgamation models will be highlighted and included in the new system. A 'user-friendly' notation for classification will be developed throughout this section.

### 1.3.1 Principle of parsimony

In deciding which factors to include in the classification system, a principle of parsimony is adopted and the number of factors used for the classification is kept to a minimum. By taking this approach :

* the essential aspects of a model are covered
* non-essential aspects may be given secondary codes, as in Chikán's system, without confusing the main classification.

### 1.3.2 Number of items

Chikán's single-item / multi-item flag requires re-interpretation for amalgamation models. In such models, the 'multi-item' case may refer to *aggregation* across a number of items, or to the treatment of aggregate inventory *as a whole*, without reference to individual SKUs. A number of examples of the former approach are given in chapters 2 and 4. An example of the latter approach is given by Johnston et al (1988a); the authors define 'stock cover' as follows :

$$\text{Stock Cover (C)} = \frac{\text{Total average stock}}{\text{Average monthly demand}}$$

A 'response equation' was found (Johnston et al (1988a), (1988b)), of the form :

$$C = a_0 + a_1D + a_2D^2 + a_3R^2 + a_4L^2 + a_5E^2$$
$$+ a_6RL + a_7RD + a_8RE + a_9LE + a_{10}LD + a_{11}ED$$

where : $D = 1/F$ and $F$ is average monthly demand per item
$R$ is review time
$L$ is lead-time
$E = \exp((S-100)/10)$ and $S$ is service level, expressed as a percentage.

Such an approach will be termed 'GROSS', since the relationships are derived from inventory as a whole. The alternative approach to multi-item modelling, where results are derived from relationships at item level, will be termed 'AGGREGATE'. The distinction between the two categories is intended to avoid confusion.

The options are amended accordingly :

| | |
|---|---|
| SKU | model of stock for an individual SKU |
| AGGREGATE | model of stock derived for a single SKU and aggregated over a set of SKUs. |
| GROSS | model of stock derived from inventory as a whole. |

'GROSS' models, which will not be covered further in this thesis, are identified as an area for further research.

### 1.3.3 Number of stores : a redundant factor

The second factor included in Chikán's classification is never required for amalgamation models. By definition, such models must cater for a multi-depot situation and the second main code (number of stores in the system) is always set to 'several stores'.

### 1.3.4 Deterministic and stochastic inflow

Two factors determine whether inflow to an inventory system is deterministic or stochastic: the quantity of inflow and the time of inflow. It has generally been assumed in the literature on inventory amalgamation that the quantity of inflow is deterministic: it is known in advance that the quantity received will be the quantity ordered. The robustness of this assumption is discussed in sub-section 5.3.10, as part of the review in chapter 5 of the Economic Order Quantity (EOQ) approach. Based on findings by Silver (1976), it will be concluded that amalgamation models are robust to deviations from the assumption. Hence, the 'quantity of inflow' will be assumed to be deterministic for the purpose of model classification.

The character of the time of inflow (deterministic or stochastic) is determined by the manner of delivery (received in a single lot, in several lots or continuously) and the character of the lead-time (deterministic or stochastic).

Chikán (1990) included the manner of delivery at the 'secondary' level of coding. Since it has received little attention in amalgamation problems, it would seem appropriate to maintain this secondary status and not include it in the new main classification system.

The character of lead-time (deterministic or stochastic) is an important factor in amalgamation models. Its coding will be discussed in sub-section 1.3.11.

### 1.3.5 Deterministic and stochastic outflow

The cycle stock centralisation models to be reviewed in chapter 4 are based on the Economic Order Quantity (EOQ) formulation. Since the standard EOQ model assumes deterministic demand, this assumption will be maintained for cycle stock models.

In the safety stock centralisation and consolidation models to be reviewed in chapters 2 and 3, it is always assumed that the demand is stochastic and, hence, that the inventory system's outflows are stochastic. A significant benefit of inventory amalgamation is the 'portfolio effect', which occurs when there are savings in safety stocks as a result of the reduction in the coefficient of variation of demand by pooling demand streams. This benefit does not apply to deterministic demand. Hence, demand will be assumed to be stochastic for safety stock models.

Under the above assumptions, it will be clear whether demand is assumed deterministic or stochastic from the context of cycle or safety stock modelling. Therefore, following the principle of parsimony, the distinction between stochastic and deterministic demand is relegated to secondary status and not included in the main classification.

### 1.3.6 The assumption of a 'static' model

Dynamic models have been studied extensively in the context of fixed-location inventory systems. It is known that the (s,S) strategy is optimal for dynamic models with stationary independent demands, full backlogging and fixed lead-times. However, Porteus (1985) showed that the optimal solution to the (s,S) model may be approximated by the less computationally intensive method of determining the order quantity and the re-order point separately, as for the static model.

In previous work on inventory amalgamation, a static model has generally been assumed. Since its solution provides a good approximation to the dynamic model, the distinction between static and dynamic models will not be included in the main classification of inventory amalgamation models.

### 1.3.7 The objective of the system

Chikán's classification system contains three codes for this factor : optimization model, reliability model and descriptive model. This is a classification of *model* objectives, not *system* objectives. If a model 'drives' a system (as is often the case in computerised inventory systems), then the distinction is rather fine. However, it is possible to design a descriptive model of a cost-driven system. For example, inventory levels may be determined by minimising a holding-cost plus shortage-cost function, and a model may be designed to show the effect on total cost of changing market conditions. In such cases, the distinction is of some practical significance.

In the new classification, any confusion will be avoided by concentrating on *system* objectives. In the inventory amalgamation literature, as in the fixed-location inventory literature, almost all discussion centres on two types of system objectives :

COST - minimisation of a cost function
SERV - satisfaction of service level criteria.

Chikán's classification assumes a single objective. Work by Bookbinder and Chen (1992) has demonstrated the importance of multiple-objective models in this field. So a further code is added :

MCDM - multiple criteria decision model based, for example, on both cost and service criteria.

The MCDM category has been included for comprehensiveness of the classification, but this line of research is not developed in the thesis. However, the area is identified as one for further research.

### 1.3.8 Ordering rules

Following Naddor (1966) and Chikán (1990), the ordering rule may be set to :

(t,q) - fixed order-interval (t), fixed order quantity (q)
(t,S) - fixed order-interval (t), order up to S

(s,q) - re-order point (s), fixed order quantity (q)
(s,S) - re-order point (s), order up to S.

As in Chikán's system, a 'p suffix' may be used to denote a fixed parameter. No modification is needed to this comprehensive set of codes. Sometimes, however, if a parameter is not specified, this should be reflected by the notation. For example, if it is known that a model is re-order point but q or S is not specified, then it will be written as (s,.).

### 1.3.9 Mode of review

Chikán listed this factor separately from the ordering rule. This seems appropriate since, for example, it is possible to run an (s,S) ordering rule using continuous or periodic review (see Silver and Peterson (1985) who designate the periodic variant as an (R,s,S) system). However, Chikán found that the 'mode of review' was not a key factor in discriminating between fixed-location inventory models. Likewise, it is not a key discriminating factor for amalgamation models.

### 1.3.10 Treatment of shortages

If a model is service-driven, the formulas for stock-holdings vary according to the way shortages are treated (see Ronen (1982)). However, amalgamation models may be adapted to both cases using the relationships between service level measures presented in chapter 7. The centralisation literature on cost-based models has been restricted to 'newsboy' problems, with lost sales at the end of each period. Therefore, the factor 'treatment of shortages' is relegated to secondary status and not included in the main classification.

### 1.3.11 Character of lead-time

Although this factor was not identified in Chikán's cluster analysis as an 'important' factor for item-level models, recent work in centralisation and consolidation theory, to be reviewed in sections 2.5 and 3.2, makes its inclusion necessary in a classification of amalgamation models.

Three codes are included for this factor :

| | |
|---|---|
| LT(zero) | - zero lead-time |
| LT(const) | - constant lead time (variance = 0) |
| LT(var) | - variable lead time (variance > 0). |

### 1.3.12 Demand correlation

At the end of section 1.2, it was noted that the correlation of demand between depots (stores) had been omitted from Chikán's classification. This factor is essential for categorising amalgamation models. Some models include correlation coefficients explicitly, allowing for any value between -1 and +1. Other models, of a different structure, do not include any reference to correlation and implicitly assume all correlation coefficients to be zero.

A suitable categorisation for the correlation of demand between depots in centralisation models is given below :

| | |
|---|---|
| CORR(0) | demand correlation is assumed to be zero between each pair of depots |
| CORR([-1,1]) | demand correlation is not assumed to be zero and may take any value between -1 and +1, inclusive, at each pair of depots. |

### 1.3.13 Single-echelon and multi-echelon models

In amalgamation problems, two questions may emerge :

1. What is the result of changing the echelon structure ?
   (eg having an extra layer, with wholesalers receiving stock from 'super-wholesalers', rather than depots.)
2. What is the result of changing the number of depots within an echelon ?

Previous amalgamation models have addressed the second question, to estimate the effect of amalgamating inventories from within one echelon of a hierarchical system. The work presented in this part of the thesis continues in that tradition and provides a critique of previously published work. This research, then, is relevant to multi-echelon systems but not to situations where the *structure* of the echelon system requires analysis. Models to address this issue represent a further stage of research.

### 1.3.14 A classification of inventory amalgamation models

If the assumptions discussed above are adopted, then there are five factors which should be used to classify amalgamation models :

1. Single-item / multi-item
2. System objective.
3. Ordering rule.
4. Character of lead-time.
5. Correlation of demand between depots.

Queueing models are generally classified using a triple such as M/M/r to indicate the arrival process, service process and number of servers. Using a similar style of presentation, the following classification notation is proposed for inventory amalgamation models:

Aggregation / objective / ordering rule / lead-time / correlation.

For example, a cost-minimisation model for a single SKU controlled by an (s,q) system with constant lead-time and no correlation of demand between depots would be classified as :

SKU / COST / (s,q) / LT(const) / CORR(0)

This classification notation has the advantage over a purely numeric system that the model is immediately recognisable. The notation will be used in the literature review of the following chapters.

## 1.4 Conclusions

Three classification systems for fixed-location inventory models have been reviewed and Chikán's system was analysed in some detail. A number of weaknesses were identified in Chikán's system, principally its inability to deal with multi-echelon systems and correlated demand.

A classification system for amalgamation models has been developed, based on Chikán's system and guided by a principle of parsimony. Additional factors may be included in the system in the future if a sufficiently strong case can be made for their inclusion. However, it is considered that the five factor system is sufficiently succinct for general use in practice.

Three areas were identified for further research : 'Gross Models', 'Multiple Criteria Decision Models' and 'Multi-Echelon Models'. Although these issues are outside the scope of this thesis, they have emerged from the development of the classification system as aspects of inventory amalgamation modelling which may merit further investigation.

In chapters 2, 3 and 4, amalgamation models will be classified using the new system. It will be shown that some mis-classifications have occurred in the literature, which can be easily highlighted and corrected using the classification system presented in this chapter.

# CHAPTER 2

## *Safety Stock Centralisation Models*

## 2.1 Introduction

### 2.1.1 Overview

A number of papers have been published on two issues related to safety stock centralisation:

* Under what conditions does centralisation yield benefits ?
* What effect does centralisation have on stock-holding, assuming that service remains unchanged ?

Most of the literature has concentrated on the second question, working on the assumption that service levels may always be maintained after centralisation. This assumption has been challenged by Chen and Lin (1989) who claimed that it is possible to give a worse service, with the same total inventory, at a centralised depot. This claim will be analysed in chapter 6, where the circumstances under which disbenefits occur are examined. In this chapter, the second question will be addressed: what safety-stock saving will result from depot centralisation ?

### 2.1.2 The nature of the critique

A critique of previous work on safety stock centralisation models is presented. The models reviewed are varied but a number of consistent themes emerge which have not always been adequately addressed in the literature :

1. What are the assumptions underpinning the model ?

2. How robust is the model to deviations from the assumptions ?

3. How can the variables in the model be estimated ?

4. Are the results mathematically sound ?

5. What are the significance of the model results ?

To answer the first question, the assumptions of each model are analysed and some omissions are rectified. Based on the full set of assumptions, the models are categorised using the classification system developed in chapter 1. Robustness issues are also explored.

The issues of measurement and estimation are often neglected in the centralisation literature. In this chapter, the variables which may cause estimation difficulties are highlighted for later evaluation.

Most of the results in this field have rigorous proofs and are mathematically sound. There are some exceptions, however, and in these cases the necessary amendments are made. The significance of the model results are assessed for individual SKUs and for stock-holdings aggregated over a set of SKUs. Particular emphasis is placed on those aspects not highlighted by the authors themselves.

## 2.2 Square root models for service-driven inventory systems with uncorrelated demand between depots

### 2.2.1 Maister's claims

Maister (1976) postulated a square root law of the following form for a single stock-keeping unit : *"the total inventory in a system is proportional to the square root of the number of locations at which a product is stocked"* (eg, total inventory in three depots is 1.732 ($\sqrt{3}$) times the inventory at a single depot). This 'law' has an unwritten assumption of *ceteris paribus*: it holds if the total demand on an inventory system remains constant, and all other relevant variables, such as lead-times, remain constant.

Maister claimed that the square root law may be used for both cycle and safety stocks. The implications for cycle stocks will be addressed in chapter 4. The formula for safety stocks is as follows :

$$\frac{\Sigma_j SS_j}{SS_0} = \sqrt{n}$$

where $SS_j$ is the safety stock at the $j$th decentralised depot
$SS_0$ is the safety stock at the centralised depot (depot 0).
$n$ is the number of decentralised depots.

Maister also provided a 'proof' of this formula. However, Das (1978) showed Maister's proof to be inadequate. He gave a rigorous proof of the following theorem (using T to represent the ratio of decentralised to centralised safety stocks) : *"For the case of safety stock, $T < \sqrt{n}$ unless the variability of demand is the same at all locations in which case T reaches its maximum, $\sqrt{n}$, and the square root law applies"*.

### 2.2.2 The significance of Maister's claims

The claims made by Maister have implications for stock holdings aggregated over all stock-keeping-units. Suppose that the square root law holds for *all* items in a range:

$$\frac{\Sigma_j CS_{ij}}{CS_{i0}} = \frac{\Sigma_j SS_{ij}}{SS_{i0}} = \sqrt{n} \quad \text{for all i.}$$

where : $CS_{ij}$ is the cycle stock for item i at depot j
$SS_{ij}$ is the safety stock " "

and : depot 0 denotes the centralised depot.

Then the ratio must also apply to the aggregate cycle and safety stocks :

$$\frac{\Sigma_j CS_j}{CS_0} = \frac{\Sigma_j SS_j}{SS_0} = \sqrt{n}$$

where : $CS_j$ is the total cycle stock at depot j ($CS_j = \Sigma_i CS_{ij}$)
$SS_j$ is the total safety stock at depot j ($SS_j = \Sigma_i SS_{ij}$).

Since the law holds for both aggregate safety stocks *and* cycle stocks :

$$\frac{\Sigma_j (CS_j + SS_j)}{CS_0 + SS_0} = \sqrt{n}$$

Therefore, if Maister's claim is that the 'square root law' holds for the safety and cycle stocks of each individual SKU, then it must follow that it also holds for aggregate total stocks under the assumptions to be examined in the next sub-section.

### 2.2.3 Classification of Maister's model

Maister's assumptions for the cycle stock square root law as follows :

1. Each location has the same proportion of total system demand (for the result to hold strictly),
2. Cycle stocks are controlled using an 'Economic Order Quantity' rule,
3. All locations have the same fixed cost per order,
4. All locations have the same unit inventory-holding costs,
5. Total demand, before and after centralisation, remains constant,

and the assumptions for the safety stock square root law are given as :

6. The variance of demand is the same at all locations (for the result to hold strictly),
7. All locations (including the centralised location) use the same safety stock multiple,
8. Demands at decentralised locations are un-correlated.

On the basis of these assumptions, Maister's model may be categorised using the classification system proposed in chapter 1. By the arguments of the previous sub-section, the model is applicable to 'AGGREGATE' stock-holdings. The seventh assumption (constant safety stock multiple) indicates the use of a service criterion over a fixed lead-time. If lead-times were variable, then a constant multiple approach would not be valid (see Clark (1957) for the required model formulation). The use of an EOQ approach for cycle stocks and the constant safety stock multiple indicate that an (s,q) ordering rule is in operation. Hence, the model may be classified as :

AGGREGATE / SERV / (s,q) / LT(const) / CORR(zero).

### 2.2.4 Estimation of demand variance

If demand variabilities are unequal, then the variance of demand at each depot must be estimated. This may present difficulties, since there are large sampling errors associated with variance estimation, particularly for small sample sizes which will arise for slow-moving SKUs. Also, variances may be over-estimated if there has been a step-change in mean demand.

This is an important issue in the practical implementation of centralisation models which has been virtually ignored in the papers reviewed. Although not addressed in this chapter, this issue will be the focus of chapters 12 and 13.

## 2.3 Square root models for cost-driven inventory systems with correlated demand between depots

### 2.3.1 Eppen's approach

Eppen (1979) concentrates on modelling single SKUs which are 'faster moving items' (items with demands which may be represented by normal distributions). He drops the assumption of uncorrelated demands at the decentralised locations. Eppen shows that if stock orders are calculated by minimising the expected sum of holding and penalty costs, then :

* Expected holding and penalty costs in a decentralised system exceed (are greater than or equal to) those in a centralised system.
* The magnitude of saving depends on the correlation of demand.
* If demands are identical and uncorrelated, then the costs increase as the square root of the number of locations.

### 2.3.2 Classification of Eppen's model

The full set of Eppen's assumptions is given below :

1. Demand occurs at the beginning of each scheduling period immediately after the inventory has been raised to the recommended level.
2. Demand is normally distributed at location j with mean $\mu_j$ and variance $\sigma_j^2$.
3. The coefficient of variation of demand, $\sigma_j/\mu_j$, at each location is sufficiently small that the probability of negative demand is negligible.
4. Demand is correlated between locations with covariances $\sigma_{jk}$ and correlation coefficients $\rho_{jk}$.
5. Stock orders are calculated on the basis of minimising total costs (ie holding and penalty costs).

6. Unit holding and penalty costs are identical at each location.

7. Holding and penalty costs are linear.

The fifth assumption is not identified as such by Eppen but it underpins the proofs of the above results. According to this assumption, the model should be categorised as 'COST' based. Since the model is of a 'newsboy' situation, the ordering rule must be of the (t,S) form (fixed intervals between orders, variable order quantities). In newsboy problems, the lead-time is taken to be zero since it is not relevant to the inventory requirement and so may be disregarded. According to the fourth assumption, demand correlations may not be taken to be zero. Hence, using the classification system proposed in the previous chapter, Eppen's model may be categorised as :

SKU / COST / (t,S) / LT(zero) / CORR(non-zero).

2.3.3 Eppen's results and their significance

Using the assumptions given above, Eppen showed that :

$$\frac{TC_D}{TC_C} = \frac{K \Sigma \sigma_i}{K [\, \Sigma \sigma_i^2 + 2 \Sigma\Sigma_{i<j} \sigma_i \sigma_j \rho_{ij} \,]^{1/2}}$$

where $TC_D$ is the total cost relevant to stocks in the decentralised system
$TC_C$ is the total cost relevant to stocks in the centralised system
K is a factor determined by the unit holding and penalty costs.

In the above formula, the safety factor is assumed to be identical in decentralised and centralised systems. For any newsboy problem, with linear holding and penalty costs, it may be shown (eg Naddor (1966)) that minimisation of total costs leads to an order level at which the cumulative distribution function attains a value of p/(p+h), where p and h are unit penalty and holding costs respectively. Since p and h are identical at all locations (by assumption 6), the K parameters are also identical and cancel in the expression given above.

Eppen's three results (summarised at the beginning of this section) follow immediately from the above equation. In the case of the square root law it is clarified that 'identical demands' denote demands with identical standard deviations.

Thus, Eppen extends the conditions under which the 'square root law' applies to cost-driven systems. Moreover, suitable adaptations to the square root law formulation are found to cover those situations where the demand between depots is correlated.

### 2.3.4 Estimation of demand covariance

The problems of estimating demand variance were discussed earlier. There are similar difficulties with the estimation of demand covariance, as required for Eppen's model. There are large sampling errors associated with the estimation of covariance, particularly for slow moving SKUs, for which demand data is sparse.

In chapter 12, an approach to variance estimation will be presented which takes covariance into account without requiring estimates of demand covariance between each pair of depots.

## 2.4 Square root models for service-driven inventory systems with correlated demand between depots

### 2.4.1 Zinn, Levy and Bowersox model

Zinn, Levy and Bowersox (1989) sought to extend the work of Eppen by deriving a formula for the safety stock ratio which takes into account the relative sizes of the standard deviations as well as the correlation coefficient values. They define a 'Magnitude' (M) ratio as follows :

$$M = \frac{\sigma_1}{\sigma_2} \qquad \text{for } \sigma_1 \geq \sigma_2 \text{ and } \sigma_2 > 0.$$

Then the following formula is derived for a two-depot problem :

$$\frac{SS_0}{SS_1 + SS_2} = \frac{(M^2 + 1 + 2M\rho_{12})^{1/2}}{M + 1}$$

where $\rho_{12}$ is the correlation of demand between depots 1 and 2.

The above formula cannot be readily extended to the general n-depot problem (n>2), as conceded by Zinn, Levy and Bowersox (henceforth ZLB). However, a portfolio effect (PE) matrix, similar in form to a correlation matrix, is suggested by the authors. The (i,j) element of this matrix shows the saving in safety stock to be obtained by amalgamating the *i*th and *j*th depots.

The formula is applicable to the single SKU case only. An illustrative example for a single stock-keeping-item is given in ZLB's paper to show how the most promising opportunities for centralisation may be identified using this approach.

### 2.4.2 Classification of the Zinn, Levy and Bowersox model

The assumptions used to derive ZLB's formula are as follows:

1. Inventory transfers between stocking locations may be ignored.
2. The standard deviation of lead-time is zero (ie *constant* lead times are assumed).
3. Customer service level, as measured by inventory availability, does not vary between stocking locations.
4. Demand at each stocking location is normally distributed.

The second assumption should be extended to state that the constant lead-times are taken to be identical at all stocking locations. This is necessary, together with the third assumption given above, to argue that the safety factors are equal at all stocking locations.

On inspection of Zinn, Levy and Bowersox's paper, it is clear that their results are readily derivable from Eppen's under the assumptions stipulated by Eppen. However, ZLB's assumptions are quite different from Eppen's. They assume an 'assurance of service' criterion instead of minimisation of costs and 'constant lead times' instead of a newsboy problem.

The full classification of ZLB's model is :

SKU / SERV / (s,.) / LT(const) / CORR(non-zero)

and *not*, as claimed by ZLB, the following classification of Eppen's model :

SKU / COST / (t,q) / LT(zero) / CORR(non-zero)

Note : (s,.) has been used rather than (s,q) or (s,S) for the ZLB model, since the method of determining the order quantity is not specified in the paper.

Zinn, Levy and Bowersox's work should, therefore, be seen as an extension of the conditions under which a safety stock ratio may be calculated rather than as a major extension of the formula itself.

### 2.4.3 Zinn, Levy and Bowersox's illustrative example

The authors gave an example of their formula applied to real data, to show its applicability. The product line chosen was men's white Jockey underwear, size 36, with data from four stores analysed over a two year period, 1985 to 1987. This product was presented as a 'typical' SKU, but this description should be challenged. Such a product is only 'typical' of goods which are fast-moving and unlikely to experience step-changes of demand. While there are many such items in a range, there are also many others for which the estimation of demand variance and covariance may be expected to be more difficult.

### 2.4.4 Debate between Ronen and Zinn, Levy and Bowersox

Ronen (1990) criticised the paper by Zinn, Levy and Bowersox on two grounds :

1. The probability of running out of stock *during lead time* is used as the availability measure. Ronen argued in a previous paper (Ronen (1982)) that this is an inferior measure of product availability.

2. If an Economic Order Quantity is assumed, then centralisation of inventory changes the number of lead-times per year, making a comparison of 'running out of stock during lead-time' meaningless.

Ronen shows that if 'overall fraction of demand supplied from stock' is used as the availability measure, then to maintain the same service level for two decentralised depots and also the centralised depot, the following conditions must apply :

$$\frac{E(k_1)\,\sigma_1}{S_1^{1/2}} = \frac{E(k_2)\,\sigma_2}{S_2^{1/2}} = E(k_0)\left(\frac{\sigma_1^2 + \sigma_2^2 + 2\sigma_1\sigma_2\rho_{12}}{S_1 + S_2}\right)^{1/2}$$

where $E(k) = \int_k^\infty (x-k)\, f(x)\, dx$

$f$ is the probability density function of demand
$S_1$ is the size of demand at the first depot
$S_2$ is the size of demand at the second depot.

This condition holds for each of the two cases :

* back-ordering in the event of a stock-out
* lost sales in the event of a stock-out.

Since the above conditions do not yield closed solutions, numerical methods must be used to find the appropriate values of $k_0$, $k_1$ and $k_2$.

If the lead-times are identical at all stocking locations, then the safety stock ratio may be found using the formula :

$$\frac{SS_0}{SS_1 + SS_2} = \frac{k_0\,(\sigma_1^2 + \sigma_2^2 + 2\sigma_1\sigma_2\rho_{12})^{1/2}}{k_1\,\sigma_1 + k_2\,\sigma_2}$$

after the values $k_0$, $k_1$ and $k_2$ have been found using a suitable computer program.

### 2.4.5 Comments on the debate between Zinn, Levy and Bowersox and Ronen

The debate between ZLB and Ronen concerns the question of which service level measure is superior : 'probability of running out of stock during lead-time' or 'overall fraction of demand filled from stock'. However, in the context of centralisation modelling, the main concern is that the model reflects the service level measure which is currently in use or which is intended to be used. Ronen's model enables the scope of centralisation models to be extended to include a different service-driven system, one in which 'fraction of demand filled from stock' is the criterion. It can be argued, though, that Ronen's objection does not go far enough. A number of other service level measures are in common use and need to be catered for in centralisation models. Some research into this question is presented in chapter 7. The findings allow a wider range of service level measures to underpin amalgamation models.

## 2.5 Square root models for service-driven inventory systems with correlated demand and variable lead-times

### 2.5.1 Tallon's model and its classification

Tallon (1993) sought to extend the results of Zinn, Levy and Bowersox (1989) to the variable lead time case. He gives an 'extended Portfolio Effect' formula for the two-depot case which takes three factors into account :

* ratio of the demand standard deviations at the two (decentralised) depots. This is called the 'magnitude' (M) after Zinn, Levy and Bowersox.
* correlation between demand at the two depots
* mean and variance of lead time at the two depots.

The assumptions are identical to those of Zinn, Levy and Bowersox, except that the lead-time may be stochastic. Hence, the classification of Tallon's model is:

SKU / SERV / (s,.) / LT(var) / CORR(non-zero)

### 2.5.2 Tallon's results

Tallon gives the following formula for the 'Portfolio Effect' :

$$\frac{SS_0}{SS_1 + SS_2} = \frac{(M^2 + 1 + 2M\rho_{12})^{1/2}}{M + 1}$$

$$\text{where } M = \left(\frac{\mu_{L1}\,\sigma_{d1}^2 + \mu_{d1}^2\,\sigma_{L1}^2}{\mu_{L2}\,\sigma_{d2}^2 + \mu_{d2}^2\,\sigma_{L2}^2}\right)^{1/2} = \frac{\sigma_{dL1}}{\sigma_{dL2}}$$

for $\sigma_{dL1} \geq \sigma_{dL2}$ and $\sigma_{dL2} \neq 0$.

In the above equations, d indicates 'demand', L indicates 'lead-time' and the 1 and 2 suffix values indicate the first and second depots. Tallon gives an illustrative example of two depots with known first and second moments of lead time and demand. The two depots both operate to a desired stock-out probability of 3% and the correlation of demand between the two depots is known. Based on these assumptions, an estimate of the safety stock savings due to centralisation is made.

### 2.5.3 Generalisation of Tallon's results

Tallon's results apply under restrictive conditions :

1. The desired probability of stockout is the same at the two decentralised depots and is unchanged by centralisation.
2. The mean lead-times of products will not change after centralisation.
3. The variance of lead-times will not change after centralisation.

The first assumption is not problematic as Tallon's approach may be easily adapted to allow for different desired probabilities of stockout. However, relaxation of the final two assumptions requires greater model adaptation. A less restrictive and more realistic formulation may be obtained by including the following variables explicitly:

* mean lead time to centralised depot
* standard deviation of lead time to centralised depot.

Note : For clarity of exposition, it will be assumed that the 'desired stock-out probabilities' are the same at each decentralised location and remain unchanged after centralisation. The approach may be easily adapted for the case of unequal stock-out probabilities.

Assuming equal safety factors at each of the locations,

$$\frac{\Sigma\, SS_i}{SS_0} = \frac{\Sigma\, \sigma_i}{\sigma_0}$$

Using the Clark (1957) equations for the variance of demand in variable lead-time,

$$\frac{\Sigma\, SS_i}{SS_0} = \frac{\Sigma\, [\mu_{li}\, \sigma_{di}^2 + \mu_{di}^2\, \sigma_{li}^2]^{1/2}}{[\mu_{l0}\, \sigma_{d0}^2 + \mu_{d0}^2\, \sigma_{l0}^2]^{1/2}}$$

where

$$\mu_{d0} = \Sigma\, \mu_{di}$$

$$\sigma_{d0}^2 = \Sigma\, \sigma_{di}^2 + 2\, \Sigma\Sigma_{1<j}\, \sigma_{di}\sigma_{dj}\, \rho_{ij}.$$

In all, n 'magnitude' values will be used: the *i*th magnitude value ($M_i$) compares the standard deviation of demand during lead time at the *i*th decentralised depot with the standard deviation of lead time demand at the centralised location.

$$M_i = \frac{[\mu_{li}\, \sigma_{di}^2 + \mu_{di}^2\, \sigma_{li}^2]^{1/2}}{[\mu_{l0}\, \sigma_{d0}^2 + \mu_{d0}^2\, \sigma_{l0}^2]^{1/2}}$$

It is important to note that :

* $\mu_{d0}$ and $\sigma_{d0}$ can be derived from data on the decentralised locations
* $\mu_{l0}$ and $\sigma_{l0}$ cannot be so derived.

The ratio of decentralised to centralised safety stocks may be written :

$$\frac{\Sigma\, SS_i}{SS_0} = \Sigma\, M_i$$

This simple formulation makes it clear that 'portfolio effect' savings do not happen automatically (as implied by Tallon's result). In fact, there are three cases :

* $\Sigma\, M_i > 1$ savings through centralisation
* $\Sigma\, M_i = 1$ indifferent
* $\Sigma\, M_i < 1$ savings through decentralisation.

### 2.5.4 Estimation of the variance of lead-time and the variance of demand

Tallon's model requires, in addition to the estimation of demand variance, the estimation of the variance of lead-time. The same difficulties apply. Large sampling errors and problems in estimation if there is a step-change in mean level are also relevant to lead-time estimation. This may be exacerbated by problems of small samples if items are re-ordered infrequently. Lead-time variability also compounds the errors in estimating the variance of demand which were noted in sub-section 2.2.4.

## 2.6 Conclusions

In this chapter, it has been shown that Maister's square root law for safety stocks is potentially very powerful, since it may be classified as 'AGGREGATE' and applied over any aggregated set of SKUs. As a similar form of model is available for cycle stocks, it was noted that a total stock model, aggregated over SKUs may be formulated.

A review and critique of safety stock centralisation models has been presented. Specific criticisms and enhancements have been made to two models :

* It has been shown that Zinn, Levy and Bowersox's model has been mis-categorised. It has been re-categorised as : SKU / SERV / (s,.) / LT(const) / CORR(non-zero).

* Tallon's model holds under very restrictive conditions. A more realistic formulation has been presented.

Some more general points have also emerged. Firstly, the question of parameter estimation has been neglected in the centralisation literature. Demand variance and covariance, in particular, are key components of a number of models but, as yet, inadequate consideration has been given to their estimation. Secondly, the question of the most appropriate service level measure, although debated by Ronen and Zinn, Levy and Bowersox, has received little attention from other authors. Since the measure actually used in practice should be reflected in the centralisation model, this aspect of inventory modelling merits further study.

# CHAPTER 3

# *Safety Stock Consolidation Models*

## 3.1 Introduction

### 3.1.1 The distinction between 'centralisation' and 'consolidation'

In the previous chapter, the focus was on the effect of centralising inventories from a number of depots to *one* depot. In this chapter, the focus will shift to the more general problem of consolidating inventories to a smaller number of depots. In practice, the issue of consolidation is important, since it is often impractical or economically undesirable (because of transport costs, for example), to centralise stocks to one depot.

### 3.1.2 Overview of literature on consolidation

Consolidation models had received very little attention in the literature until recent years. In this chapter, two papers which address the issue of consolidation will be reviewed. It will be shown that the findings of the first paper to be reviewed (Evers and Beier (1993)) are limited by some practical considerations not discussed by the authors.

The second paper to be reviewed, by Mahmoud (1992), offers an alternative approach by introducing the measures of 'portfolio quantity effect' and 'portfolio cost effect'. Two of Mahmoud's conclusions regarding the application of the 'portfolio quantity effect' to consolidation are questioned and alternative analyses are presented.

## 3.2 Consolidation modelling based on the portfolio effect

### 3.2.1 The model of Evers and Beier and its classification

Evers and Beier (1993) sought to extend the work of Zinn, Levy and Bowersox (1989). Four of ZLB's assumptions were relaxed, namely :

* Same safety stock multiple is used at all locations.
* No lead-time uncertainty.
* All average lead-times are the same.
* Depots are centralised to *one* location.

Evers and Beier assume that the safety stock multiples may vary between depots and may differ before and after consolidation ($k_{aj}$ at each consolidated location j; $k_{bi}$ at each unconsolidated location i). The lead-times may be uncertain (standard deviations $\sigma_{Laj}$ and $\sigma_{Lbi}$ for lead-times to consolidated and unconsolidated locations respectively). The mean lead-times may differ between depots ($L_{aj}$ and $L_{bi}$ for mean lead-times to consolidated and unconsolidated locations). Finally, the model presented is a *consolidation model* : it is assumed that safety stock is consolidated from n depots to m depots, where m may be greater than unity.

Under the above assumptions, the model of Evers and Beier may be classified as :

SKU / SERV / (s,.) / LT(var) / CORR(non-zero)

The classification is the same as for the model of Tallon (1993), reviewed in section 2.5, but the application has changed from centralisation to consolidation.

### 3.2.2 The general formulation of Evers and Beier

Evers and Beier's formulation rests on the same definition of the 'portfolio effect' as adpoted by Tallon (1993), namely "*the percentage reduction in aggregate safety stock made possible by consolidation of inventory from multiple locations into one location*". Evers and Beier present the following formula for the portfolio effect (PE), based on the assumptions discussed in the previous sub-section :

$$PE = 1 - \frac{\Sigma_{j=1}^{m} k_{aj} [L_{aj} (\Sigma_{i=1}^{n} W_{ij}^2 \sigma_{Dbi}^2 + 2 \Sigma_{i=1}^{n} \Sigma_{l=1}^{i-1} W_{ij} W_{lj} \delta_{Dbil}) + \sigma_{Laj}^2 (\Sigma_{i=1}^{n} W_{ij} D_{bi})^2]^{1/2}}{\Sigma_{j=1}^{n} k_{bi} [L_{bi} \sigma_{Dbi}^2 + d_{bi}^2 \sigma_{Lbi}^2]^{1/2}}$$

where the notation is as described above and :

$W_{ij}$ is the proportion of mean demand during one time period transferred from unconsolidated location i to consolidated location j (where $0 \leq W_{ij} \leq 1$ for all i and j and $\Sigma_{j=1}^{n} W_{ij} = 1$ for all i).

$D_{aj}$ is the mean demand during one time period at consolidated location j
$D_{bi}$ is the mean demand during one time period at unconsolidated location i.

$\sigma_{Daj}$ is the standard deviation of demand during lead-time at consolidated location j
$\sigma_{Dbi}$ is the standard deviation of demand during lead-time at unconsolidated location i.

$\delta_{Dbil}$ is the covariance of demand between unconsolidated locations i and l

$L_{aj}$ is the mean lead-time to consolidated location j
$L_{bi}$ is the mean lead-time to unconsolidated location i

$\sigma_{Laj}$ is the standard deviation of lead-time to consolidated location j
$\sigma_{Lbi}$ is the standard deviation of lead-time to unconsolidated location i.

Evers and Beier's formula is sound, but depends on the estimation of a number of parameters which are difficult to estimate, as previously discussed. For the formula to be useful in practice, these difficulties must be addressed.

### 3.2.3 A critique of Evers and Beier's challenge to Maister's model

Evers and Beier revisited the problem of consolidating from n depots to m depots, based on the assumptions underlying the investigation by Maister (1976). They claim that the portfolio effect indicated by Maister is non-optimal. Maister showed that the portfolio effect may be calculated as:

$$PE = 1 - \frac{\sqrt{m}}{\sqrt{n}}$$

However, Evers and Beier claim that *all* of the safety stock benefits of centralising to one depot may be achieved by consolidating to m depots (m > 1) and the portfolio effect may be calculated as :

$$PE = 1 - \frac{1}{\sqrt{n}}$$

This result depends on re-allocating all the demands to depots in such a way that demands at all locations are perfectly correlated. If this happens, then there are no further safety stock savings by centralising from m depots to just one depot. Consequently, all the benefits are attained by the first consolidation and re-allocation of demand.

While this is a powerful result, Evers and Beier's discussion of its practicality is rather limited. In principle, it would be possible for a computer system to re-allocate demand in the way they anticipate, but two major problems remain :

1. For slow moving items, no more than one item held in stock at each of the m consolidated depots may be required to attain the necessary inventory service-

levels. If demand is sufficiently low, then fewer than m items at the single centralised depot may also be enough to meet service level requirements. In this case, a further stock saving may be made, whether or not demand is correlated. The conditions under which centralisation yields benefits for slow-moving stock are discussed further in chapter 6.

2. Depots are often sited so as to serve a defined geographical region (eg Scotland, north of England etc). Evers and Beier's 'optimal solution' is only optimal with respect to safety stock holding; when transport costs are taken into account, the solution may be far from optimal.

## 3.3 Consolidation modelling based on the Portfolio Quantity Effect and the Portfolio Cost Effect

### 3.3.1 Mahmoud's approach and its classification

Mahmoud (1992) defines a 'consolidation scheme' of a set of stocking locations, I, to be a set of consolidations $I_1, I_2, ..., I_z$ satisfying the two conditions :

$$I_1 \cup I_2 \cup ... \cup I_z = I \qquad \text{(mutually exhaustive)}$$

$$I_i \cap I_j = \emptyset \quad \forall\, i \neq j \qquad \text{(mutually exclusive).}$$

The first condition ensures that each facility of I belongs to *at least* one consolidation; the second condition ensures that each facility of I belongs to *at most* one consolidation.

In section 2.4, it was noted that Zinn, Levy and Bowersox (1989) had suggested using the 'portfolio effect' (PE), the percentage reduction in safety stock for those facilities which are consolidated, to compare different consolidation schemes. A formal definition of PE was given in sub-section 3.2.2. However, Mahmoud points out that, "*Since PE is defined as a reduction ratio, not as a reduction quantity, a higher PE value does not necessarily mean a greater quantity reduction in safety stock*".

Mahmoud gives an example of this criticism by replicating ZLB's illustrative example of the effect of consolidating stocks of men's underwear from four stores. Taking the unconsolidated stock at the fourth store as a base (stock = 100), the stocks at the four depots are 75, 174, 94 and 100. The portfolio effect (PE) matrix is shown in Table 3.1 :

TABLE 3.1
Mahmoud's PE Matrix (1992)
(from an example by Zinn, Levy and Bowersox (1989))

| | *1* | *2* | *3* | *4* |
|---|---|---|---|---|
| *1* | | .274 | .406 | .321 |
| *2* | | | .136 | .058 |
| *3* | | | | .116 |
| *4* | | | | |

The matrix shows the proportionate saving by combining any pair of depots. The table may give the misleading impression that if there is to be a consolidation to two depots, then the best pairings are depots 1 & 3 and 2 & 4, since this gives the highest total PE value. However, this would be mistaken, as the following calculations illustrate:

Saving by consolidating to (1,3) and (2,4) = [.406 * (75+94)] + [.059 * (174+100)]
= 68 + 16
= 84

Saving by consolidating to (1,4) and (2,3) = [.321 * (75+100)] + [.136 * (174+94)]
= 56 + 36
= 92

These calculations show that proportions may be misleading without reference to the total values of which they are a proportion. To overcome this problem, Mahmoud defines the 'portfolio quantity effect' (PQE) as follows:

$$PQE_I = PE_I \Sigma_{i\epsilon I} SS_i$$

$$PE_I = \frac{\Sigma_i SS_i - SS_0}{\Sigma_i SS_i}$$

where I is the set of depots to be consolidated, $PE_I$ is the portfolio effect, $SS_i$ is the safety stock at the $i$th depot and $SS_0$ is the safety stock at the centralised depot. Thus, the PQE measure is defined as the quantity reduction in safety stock.

Mahmoud also defines a 'portfolio cost effect' (PCE) composed of:

* The portfolio holding cost effect,
* The portfolio transportation cost effect,
* The portfolio investment cost effect,
* The portfolio procurement cost effect.

and argues that the portfolio cost effect measure is superior to the portfolio effect since it assesses the total logistics cost saving due to depot centralisation.

As summarised above, it is clear that Mahmoud is replacing one criterion (portfolio effect) by alternative criteria (portfolio quantity effect and portfolio cost effect). In the same way as the portfolio effect may be used for different model assumptions, Mahmoud's criteria can be used for a wide range of classifications of the centralisation model. However, in the examples which Mahmoud uses to illustrate his approach, the assumptions made by Zinn, Levy and Bowersox have been maintained and the model may be classified as :

SKU / SERV / (s,.) / LT(const) / CORR(zero).

### 3.3.2 Mahmoud's analysis of sub-consolidations and super-consolidation

Mahmoud uses the portfolio quantity effect approach to examine the conditions under which 'super-consolidation' (centralisation to one depot) is preferable to 'sub-consolidations' (consolidations to more than one depot). His conclusions are as follows :

1. Assuming identical safety stock factors at all locations, the conditions under which a total centralization is better than, worse than, or equivalent to, several sub-consolidations cannot be explicitly derived.

2. Under the simple assumptions of uniform demand variations and safety stock levels for all locations, for the case of four locations, the preference of super-consolidation (1,2,3,4) versus two sub-consolidations, say (1,2) and (3,4), depends on the relative magnitude of the sum of the *inter*-correlation coefficients (correlations between depots in different consolidation sets) and the *intra*-correlation coefficients (correlations between depots in the same consolidation set).

The first conclusion will be referred to as relating to the 'general case' and the second as relating to the 'special case' (of course, the general case is not fully general since equal safety stock factors are assumed). Both of the above conclusions will be shown to be false and alternative conclusions will be stated and proved.

### 3.3.3 Preference between super and sub-consolidations when demand variances are unequal (general case)

Suppose that two options are to be assessed :

1. Consolidate all the depots in the set S into one depot.

2. Consolidate to p depots :

$$S = S(1) \cup S(2) \cup \ldots \cup S(p)$$
$$S(i) \cap S(j) = \varnothing \quad \text{for all } i \neq j.$$

It may be shown that there is indifference between the two options if the following condition holds:

$$\Sigma\Sigma_{i<j\,(\mathrm{INTER})}\, \sigma_i\sigma_j\rho_{ij} = \Sigma\Sigma_{i<j}\, [C(i)\, C(j)]^{1/2}$$

whereas super-consolidation is better if :

$$\Sigma\Sigma_{i<j\,(\mathrm{INTER})}\, \sigma_i\sigma_j\rho_{ij} < \Sigma\Sigma_{i<j}\, [C(i)\, C(j)]^{1/2}$$

and sub-consolidations are better if :

$$\Sigma\Sigma_{i<j\,(\mathrm{INTER})}\, \sigma_i\sigma_j\rho_{ij} > \Sigma\Sigma_{i<j}\, [C(i)\, C(j)]^{1/2}$$

where $\Sigma\Sigma_{i<j\,(\mathrm{INTER})}\, \sigma_i\sigma_j\rho_{ij} = \Sigma\Sigma_{i<j;i,j\in S}\, \sigma_i\sigma_j\rho_{ij} - \Sigma_m\, [\Sigma\Sigma_{i<j;i,j\in S(m)}\, \sigma_i\sigma_j\rho_{ij}]$

$$C(i) = \Sigma_{m\in S(i)}\sigma_m^{\;2} + 2\Sigma\Sigma_{m<n;m,n\in S(i)}\sigma_m\sigma_n\rho_{mn}$$

and $\sigma_i$ is the standard deviation of demand at the *i*th depot

$\rho_{ij}$ is the correlation coefficient of demand between the *i*th and *j*th depots.

The proof of the above results is given in Appendix 3.1. The formulae demonstrate that Mahmoud's first conclusion is false. It *is* possible to derive explicit rules for preference between super-consolidation and several sub-consolidations.

3.3.4 Numerical example (general case)

The data summarised by Zinn, Levy and Bowersox on underwear sales at four stores will be used as a numerical illustration of the general case. Mahmoud inferred the following standard deviations of sales at the four stores (without loss of generality, the standard deviation at the fourth store is taken to be unity).

$$\sigma_1 = 0.746,\ \sigma_2 = 1.740,\ \sigma_3 = 0.943,\ \sigma_4 = 1.000$$

In order to compare super-consolidation of all four stores with two sub-consolidations ((1,2) and (3,4)), the above values are substituted into the general formula to give the following indifference relation :

$$0.703\rho_{13}+0.746\rho_{14}+1.641\rho_{23}+1.740\rho_{24} = [(3.5841+2.596\rho_{12})(1.889+1.886\rho_{34})]^{1/2}$$

The indifference relation shows the *inter*-correlations on the left (between depots from different sub-consolidations) and the *intra*-correlations on the right (between depots from the same sub-consolidation).

Some special cases are of interest :

1. If all the correlations are zero, then super consolidation is preferable, since the left-hand expression is less than the right-hand expression. This result holds generally: whenever all correlations are zero, super-consolidation is preferable. The result is an immediate corollary of the general condition proved in Appendix 3.1.

2. If all the *intra*-correlations are zero, then the indifference relation is linear:

$$0.703\rho_{13}+0.746\rho_{14}+1.641\rho_{23}+1.740\rho_{24} = (3.5841*1.889)^{1/2} = 2.602.$$

The result is a general one : the linear form of the indifference relation applies whenever all the *intra*-correlations are zero. Again, it is a direct corollary of the general result proved in Appendix 3.1.

3. If all the *inter*-correlations are zero, then super-consolidation is preferable if and only if :

$$(3.5841+2.596\rho_{12})(1.889+1.886\rho_{34}) > 0$$

In this case, it is clear that the above expression is positive over the entire range $-1 \leq \rho_{12} \leq 1$ and $-1 \leq \rho_{34} \leq 1$ and so super-consolidation is always preferable to the sub-consolidations (1,2) and (3,4).

### 3.3.5 Preference between super and sub-consolidations when demand variances are equal (special case)

In this case, if equal safety factors are assumed, then equal safety stocks have also been assumed. The general condition simplifies to :

$$\Sigma\Sigma_{i<j\,(\mathrm{INTER})}\,\rho_{ij} = \Sigma\Sigma_{m<n}\,[(s_m+2\Sigma\Sigma_{i<j;i,j\in S(m)}\rho_{ij})(s_n+2\Sigma\Sigma_{i<j;i,j\in S(n)}\rho_{ij})]^{1/2}$$

where $s_m$ is the number of depots in the sub-consolidation S(m). The proof is given in Appendix 3.1.

### 3.3.6 Numerical example (special case)

For the four depot case, Mahmoud gives, without proof, the following rule (corrected for a typographical error) to compare consolidating two pairs of depots to consolidating all depots :

$$\rho_{13} + \rho_{14} + \rho_{23} + \rho_{24} - \rho_{12} - \rho_{34} > 2, \qquad \text{sub-consolidations better}$$

$$\rho_{13} + \rho_{14} + \rho_{23} + \rho_{24} - \rho_{12} - \rho_{34} = 2, \qquad \text{indifferent}$$

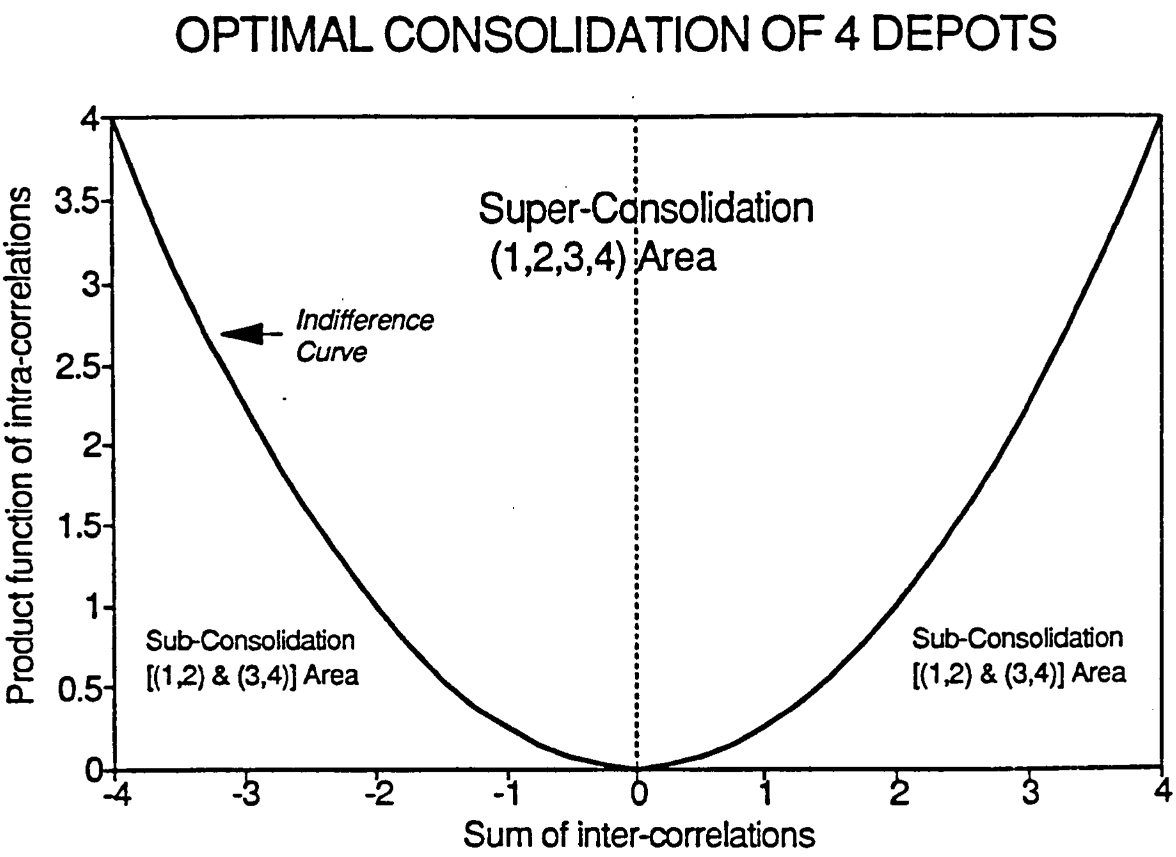
$$\rho_{13} + \rho_{14} + \rho_{23} + \rho_{24} - \rho_{12} - \rho_{34} > 2, \qquad \text{super-consolidation better}$$

However, using the example followed by Mahmoud, this will be shown to be false. If variances are equal at all depots, then the indifference relation between super-consolidation (1,2,3,4) and two sub-consolidations ((1,2) and (3,4)) may be derived. It is given by the following quadratic relation between a product function of intra-correlations and the sum of the inter-correlations :

$$(2 + 2\rho_{12})\ (2 + 2\rho_{34}) \ = \ (\rho_{13} + \rho_{14} + \rho_{23} + \rho_{24})^2$$

The result is shown graphically in Figure 3.1 :

Figure 3.1
Optimal consolidation of four depots




The above diagram shows that the form of the indifference relation is quadratic. The indifference relation is not linear, as suggested by Mahmoud's second conclusion, except when all the *intra*-correlations are zero.

## 3.4 Conclusions

Evers and Beier's model was examined and found to be mathematically sound. However, their challenge to Maister's consolidation model was found to depend on ignoring transport costs and the savings which may accrue for slow-moving items at the centralised depot.

The safety stock economies of scale due to depot-consolidation have been analysed using the portfolio quantity effect. It has been shown that a general condition for preference of super-consolidation over several sub-consolidations exists; this condition has been presented and proved. The form of the indifference relationship is shown to be quadratic; the relationship is linear when all of the *intra*-correlations (correlations between depots within the same sub-consolidation) are zero. Another consequence of the general relationship is that if all of the correlations are zero, then super-consolidation is always preferable.

# CHAPTER 4

## *Cycle Stock Centralisation Models*

### 4.1 Introduction

Cycle stock models have been somewhat neglected in the inventory centralisation literature. Since cycle stocks may amount to a substantial proportion of an organisation's stock-holding, this neglect is surprising. A comprehensive analysis of centralisation cannot be conducted without assessing both safety and cycle stocks.

In this chapter, a review and critique will be presented of the papers which have been published on cycle stock centralisation models. The issue of the robustness of cycle stock models has received some attention in these papers. However, it will be shown that the treatment is inadequate since the results do not permit the analyst to distinguish those cases where models are a good approximation and cases where the approximation is poor. New results are presented which allow this distinction to be made.

The work in this chapter is based on the Economic Order Quantity (EOQ) approach. A number of criticisms have been made of this approach for fixed location models. These criticisms will be reviewed in the next chapter, with special consideration of their relevance to inventory amalgamation models. Although it will be shown in chapter 5 that the results presented in this chapter are robust to some deviations from EOQ assumptions, if the EOQ approach is not adopted at centralised or decentralised depots, then the results no longer apply.

## 4.2 Maister's square root law

### 4.2.1 Introduction

In section 2.2, the paper by Maister (1976) was reviewed. Maister argues that the square root law may be used for stock centralisation. Using a definition of cycle stock as '*those stocks held to achieve economies of scale in ordering and production*' (and, hence, excluding safety stocks), Maister posits the following formula for the cycle stock quantity determined by an Economic Order Quantity rule :

$$\frac{\Sigma_j CS_j}{CS_0} = \sqrt{n}$$

where $CS_j$ is the cycle stock at the *j*th decentralised depot
$CS_0$ is the cycle stock at the centralised depot (depot 0).
n is the number of decentralised depots.

Maister concedes that the result is not new and has often been applied to safety stocks. His contribution is to meet three objectives :

* explore the necessary assumptions for the law to hold
* provide a formal proof of the square root law
* extend the square root law result beyond its normal usage.

In meeting the final objective, Maister claims that the square root law holds for *cycle stocks*, as well as safety stocks.

As shown in sub-section 2.2.3, Maister's model may be categorised as :

AGGREGATE / SERV / (s,q) / LT(const) / CORR (zero).

4.2.2 Maister's derivation of the square root law for cycle stocks

Maister's cycle stock proof is in three steps :

*Step 1* : It is shown that if $d_{ij}$ denotes the proportion of the demand for item i which is made on depot j, then the cycle stock ($CS_{ij}$) for item i at depot j may be expressed as a function of the demand at each depot. Moreover, the total decentralised cycle stocks ($\Sigma_j CS_{ij}$) exceed the centralised cycle stock ($CS_{i0}$):

$$\Sigma_j CS_{ij} = k \Sigma_j (d_{ij})^{1/2} > k (\Sigma_j d_{ij})^{1/2} = CS_{i0}$$

subject to the assumptions of Maister's model. In the above expression, $CS_{i0}$ denotes the cycle stock for item i at the centralised depot (depot 0) and k is a constant term derived, in the usual manner, from ordering costs, costs of holding inventory and total demand. It is assumed that these values are unchanged when moving from a decentralised system to a centralised system or vice-versa.

*Step 2* : It is shown that the cycle stock savings due to centralisation are maximised when the proportions of demand at each location are identical, ie :

$$d_{i1} = d_{i2} = \dots = d_{in} = 1/n$$

and then :

$$\frac{\Sigma_j CS_{ij}}{CS_{i0}} = \sqrt{n}$$

*Step 3* : Using an argument based on the binomial expansion, Maister claims that the result proved in Step 2 also holds approximately when the proportions $d_{ij}$ are unequal. Step 3 of the proof is reviewed in more detail in sub-section 4.3.2.

## 4.3 Criticisms of Maister's square root law

### 4.3.1 The challenge to Maister's cycle stock results by Das

Das (1978) challenges Maister's results by questioning the criterion used for deriving the cycle stock ratio. In particular, he queries Maister's approach in Step 2 of the proof for the cycle stock square root law. Das presents an alternative approach which assumes that demands may be allocated to depots by the organisation subject only to the minimum proportions stipulated ($r_{i1}, r_{i2}, \ldots, r_{in}$) for each item i at depots 1,...n. Das suggests the following formulation based on finding the allocation of demands which minimises total decentralised inventory.

$$\begin{aligned} &\text{Minimise} \qquad && \Sigma \sqrt{d_{ij}} \\ &\text{subject to} && \Sigma\, d_{ij} = 1 \\ & && d_{ij} \geq r_{ij} \;, j = 1,2,\ldots,n \end{aligned}$$

where $d_{ij}$ is the demand for item i at depot j

and, without loss of generality it is assumed that,

$$r_{i1} \leq r_{i2} \leq \ldots \leq r_{in}$$

Das gives the solution to this constrained minimisation problem as :

$$d_{ij} = r_{ij} \;,\; j = 1,2,\ldots,n\text{-}1$$

$$d_{in} = 1 - (r_{i1} + r_{i2} + \ldots \; r_{i,n-1})$$

and, consequently :

$$\frac{\Sigma_j\, CS_{ji}}{CS_{i0}} = \Sigma_j^{*}\, r_{ij}^{1/2} + (1 - \Sigma_j^{*} r_{ij})^{1/2}$$

where $\Sigma_j^{*}$ denotes summation over j from j=1 to j=n-1.

Hence, in this case, equal allocations are sub-optimal for the decentralised depots and the centralised: decentralised cycle stock ratio is over-stated.

### 4.3.2 Critique of Das's challenge to Maister's cycle stock model

Das' formulation is more general than Maister's, and the result given above shows why the decentralised to centralised cycle stock ratio is less than $\sqrt{n}$ unless $r_{ij} = 1/n$ for each depot j. However, Maister does not suppose the square root law to be generally applicable exactly but claims it to hold approximately. Das gives examples of specific values of $r_{ij}$ which demonstrate that the square root law may not be a good approximation in certain circumstances. He does not show *why* the square root law may be a poor approximation. Also, he provides no demarcation between cases where the square root law is a good and a poor approximation. In this sub-section, the question of why the root law breaks down will be addressed; in the next section, the robustness of the law is examined in more detail.

On re-examination of Step 3 of Maister's proof, it will become clear why the square root law may provide a poor approximation for cycle stocks. Suppose that the system begins with equal proportions of demand in each of the n locations and that the proportion of demand for product i at location j is perturbed by a percentage $\delta_{ij}$. It is supposed that these perturbations do not affect total demand and so the condition that $\Sigma_j \delta_{ij}=0$ is assumed to hold.

The perturbation leads to the following proportions of demand :

$$d_{ij} = (1/n)(1 + \delta_{ij}) \qquad \text{for } j = 1, \dots, n$$

Since the proportions sum to unity ($\Sigma_j\, d_{ij} = 1$), the cycle stock ratio may be expressed:

$$\frac{\Sigma_j\, CS_{ij}}{CS_{i0}} = \Sigma_j\, d_{ij}^{1/2}$$
$$= (1/\sqrt{n})\, \Sigma_j\, (1 + \delta_{ij})^{1/2}$$

And, using the binomial expansion :

$$\frac{\Sigma_j\, CS_{ij}}{CS_{i0}} = (1/\sqrt{n})\, \Sigma_j\, [1 + 1/2\, \delta_{ij} - 1/8\, \delta_{ij}^2 + O(\delta_{ij}^3)]$$

Since the perturbations sum to zero ($\Sigma_j\, \delta_{ij} = 0$), this may be simplified to :

$$\frac{\Sigma_j\, CS_{ij}}{CS_{i0}} = \sqrt{n} - [(1/8)/\sqrt{n}]\, \Sigma_j\, \delta_{ij}^2 + O(\delta_{ij}^3)$$

In Maister's paper, it is argued that the higher order terms, $O(\delta_{ij}^3)$, may always be ignored since $\delta_{ij}$ is a *percentage*. Maister began by expressing the perturbation as a percentage, but went further and implicitly assumed the percentage to be less than a hundred (and, so, the perturbation to be less than one). Now, if $\delta_{ij}$ is less than one, then the higher order terms become small and may be neglected. However, percentages may be greater than one hundred, and correspondingly $\delta_{ij}$ may be greater than one. Then the higher order terms may not be ignored. Hence for these larger perturbations, the approximation may become poor.

There is a further difficulty with Maister's argument. Even if the higher order terms are small, it is not obvious that the second term, $- [(1/8)/\sqrt{n}] \Sigma_j \delta_{ij}^2$, can always be ignored. For a small number of depots, n, this second term may be significant. This factor complicates the identification of perturbations beyond which the approximation becomes poor. In such cases, more detailed calculations are required.

## 4.4 The robustness of the cycle stock square root law

The discussion in the previous section indicated that the square root law may not always be a good approximation to the true saving in cycle stock through centralisation. However, the point at which the approximation ceases to be 'good' was not quantified. In this section, some results are given to help identify those situations when the square root law will be 'good' and when it will be 'poor'.

### 4.4.1 Robustness for the two-depot case

Firstly, a measure must be defined by which an approximation may be assessed. From Das' work, it is known that the square root law gives an *upper bound* on the amount of cycle stock savings. It would be natural, then, to calculate the proportion (P) of the maximum savings which will accrue by depot centralisation. Defined formally :

$$P = \frac{1 - [CS_{i0} / (CS_{i1} + CS_{i2})]}{1 - (1/\sqrt{2})}$$

where P = proportion of maximum cycle stock savings which will occur

and the other notation is unchanged.

The cycle stock savings are a function of the mean demands at the two decentralised depots. Denoting the ratio of mean demands by R (mean demand at larger depot divided by mean demand at smaller depot), the expression for P may be re-written as follows :

$$P = \frac{1 - [(R + 1)^{1/2} / (R^{1/2} + 1)]}{1 - (1/\sqrt{2})}$$

This result is a special case of the more general expression presented in the next sub-section.

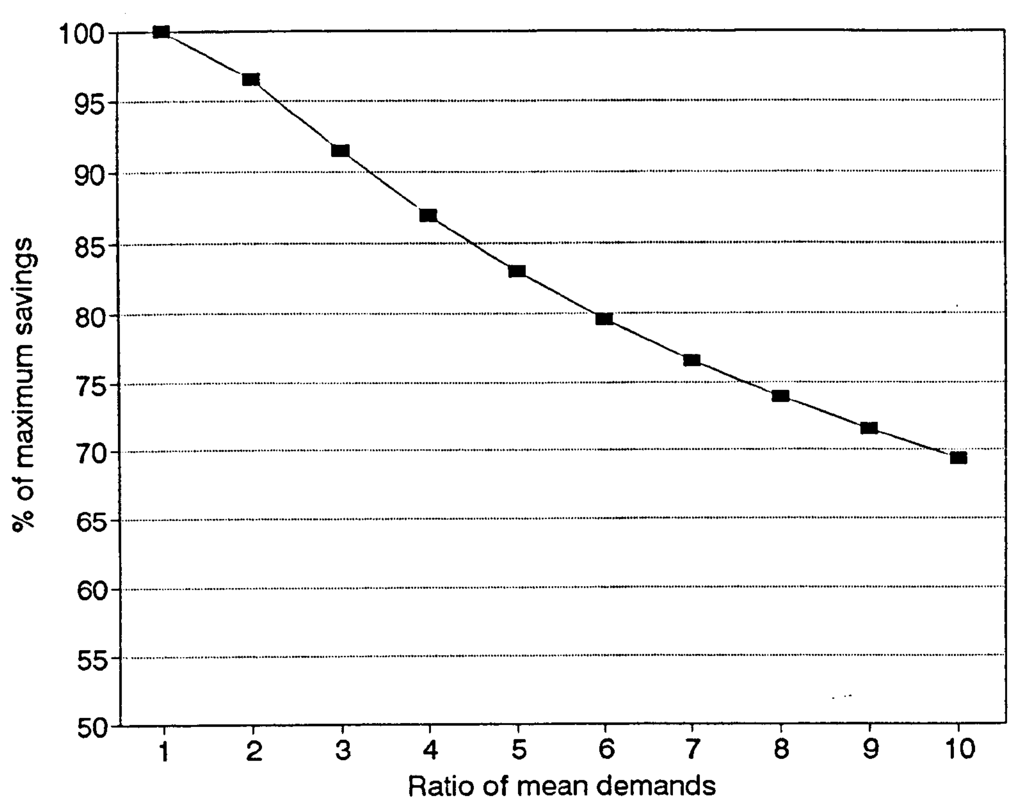
Using this simple algebraic expression, the sensitivity of the cycle stock square root law may be readily investigated for the case of two decentralised depots. The proportions of maximum saving have been calculated for a range of mean demand size ratios from unity to ten, and are presented in Table 4.1.

TABLE 4.1
Proportion of maximum cycle stock savings
if the two decentralised mean demands are not identical

| **Ratio of Mean demands** | **Percentage of Maximum Savings** |
|---|---|
| 1 | 100.00 |
| 2 | 96.47 |
| 3 | 91.48 |
| 4 | 86.94 |
| 5 | 82.99 |
| 6 | 79.55 |
| 7 | 76.54 |
| 8 | 73.88 |
| 9 | 71.50 |
| 10 | 69.37 |

This is shown graphically in Figure 4.1 :

Figure 4.1
Proportion of maximum cycle stock savings attainable




The above table and graph indicate that, when centralising from two depots to one, the square root law for cycle stocks is indeed robust but the robustness requires qualification. The previous examination of Maister's proof showed that the square root law approximation may become poor when the perturbation exceeds one (or the size ratio is more than two). Table 4.1 confirms this finding and quantifies the degree of approximation.

The relation between P (proportion of maximum saving) and R (size ratio) may be re-expressed to give R as a function of P :

$$R = \left( \frac{Q^2 + (2Q^2 - 1)^{1/2}}{1 - Q^2} \right)^2 \qquad \text{where} \qquad Q = 1 - \left( 1 - \frac{1}{\sqrt{2}} \right) P$$

The above formula yields the following results for the two-depot case, summarised below in Table 4.2 :

TABLE 4.2
Size ratios above which predicted savings are not attained

| **Percentage of maximum savings** | **Ratio of Mean demands** |
|---|---|
| 95 | 2.29 |
| 90 | 3.31 |
| 85 | 4.47 |
| 80 | 5.85 |

Table 4.2 shows that if 95% accuracy in estimating savings is adequate, then the cycle stock square root law is adequate if the ratio of mean demands is less than 2.29. This confirms that the approximation is adequate if the ratio is less than two, as shown by the analysis of Maister's proof.

### 4.4.2 Robustness for the multi-depot case

In the previous sub-section, a result for the 'proportion of maximum saving' was given, when centralising from two depots to one :

$$P = \frac{1 - [CS_{i0} / (CS_{i1} + CS_{i2})]}{1 - (1/\sqrt{2})}$$

This result generalises easily to the case of centralising from n depots to one :

$$P = \frac{1 - [CS_{i0} / \Sigma_j CS_{ij}]}{1 - (1/\sqrt{n})}$$

If the assumptions are maintained that holding and interest costs remain unchanged and that total demand also stays constant, then this may be re-expressed as :

$$P = \frac{1 - (1 / \Sigma_j d_{ij}^{1/2})}{1 - (1/\sqrt{n})}$$

Hence, for any known combination of demands, it is possible to calculate the proportion of cycle stock savings predicted by the square root law which may be achieved. Straightforward substitution in the above formula is all that is required.

If the exact split of demands between depots is difficult to estimate, or if it is undergoing rapid change, then it would be useful if *lower bounds* on the cycle stock savings may still be estimated (under the unchanged proviso that total demand remains unaltered). In the following, a simple method for identifying such bounds is outlined.

The above expression for P is minimised when $\Sigma_j d_{ij}^{1/2}$ is minimised. However, it is known from Das' argument that if each of the total demand proportions, $d_{ij}$, is constrained to be greater than r (say), then $\Sigma d_{ij}^{1/2}$ is minimised when :

$$d_{i1} = r,\ d_{i2} = r,\ \ldots,\ d_{i,n-1} = r,\ d_{in} = 1 - (n-1)r$$

Therefore, if the minimum proportion, r, is known with confidence, then a lower bound on the percentage of maximum savings may be estimated by using the above proportions. A simple analytical expression may be used for calculation purposes :

$$P = \frac{1 - [(R + n - 1)^{1/2} / (R^{1/2} + n - 1)]}{1 - (1/\sqrt{n})}$$

where R is the ratio of the largest demand to the (common) smaller demands :

$$R = \frac{1 - (n-1)\,r}{r}$$

The earlier result for two depots is a special case of the above formula.

The consequences of this result are tabulated below for up to five depots, although the table is easily extendable to more than five depots.

TABLE 4.3 Proportion of maximum cycle stock savings attainable if mean demands are not identical (multi-depot case)

| Minimum proportion at one depot | Percentage of maximum savings by centralisation of 2 depots | 3 depots | 4 depots | 5 depots |
|---|---|---|---|---|
| 0.05 | 56.50 | 67.10 | 74.43 | 79.77 |
| 0.10 | 71.50 | 81.64 | 87.98 | 92.20 |
| 0.15 | 80.65 | 89.76 | 94.93 | 97.98 |
| 0.20 | 86.94 | 94.84 | 98.69 | 100.00 |
| 0.25 | 91.49 | 98.00 | 100.00 | |
| 0.30 | 94.80 | 99.67 | | |
| 0.35 | 97.17 | | | |
| 0.40 | 98.77 | | | |
| 0.45 | 99.70 | | | |
| 0.50 | 100.00 | | | |

As in the two-depot case analysed in the previous section, Table 4.3 provides some support for the robustness of the cycle stock square root law, but a support which requires qualification. For example if the demand is split 10 : 10 : 80 between decentralised depots, then only 81.64 % of the savings predicted by the square root law are achievable. However, if the lowest proportion of demand is 10 % in one of five decentralised depots, then at least 92.20 % of square root law savings may be achieved, no matter what the split of demand between the four remaining decentralised depots.

To identify more precisely the point at which the square root law approximation becomes poor, it may be helpful to re-express the above results in a further table :

TABLE 4.4 Size ratios above which given percentages of savings predicted by the cycle stock square root law are not attained

| **Percentage of maximum savings** | **Ratio of larger demand to (common) smaller demands** | | | |
|---|---|---|---|---|
| | **2 depots** | **3 depots** | **4 depots** | **5 depots** |
| 95 | 2.29 | 2.95 | 3.64 | 4.35 |
| 90 | 3.31 | 4.58 | 5.95 | 7.41 |
| 85 | 4.47 | 6.49 | 8.71 | 11.12 |
| 80 | 5.85 | 8.82 | 12.13 | 15.76 |

Table 4.4 indicates that :

1. As the number of depots to be centralised increases, so the robustness of the cycle stock square root law improves. For example, for the 5-depot case, a size ratio of beyond 7.41 : 1 is required for less than 90% of square root law savings to be achieved (this is equivalent to a 64.9 : 8.75 : 8.75 : 8.75 : 8.75 split of demand).

2. The approximate rule that size ratios greater than two are needed before any serious inaccuracies are observed is confirmed. No size ratios below 2.29 are required for less than 95% of square root law savings to be achieved.

## 4.5 Conclusions

In this chapter, extensive but not unqualified support has been given to the square root law for cycle stocks. Maister (1976) overstated the robustness of the cycle stock square root law, as shown by Das (1978). In section 4.3, the reasons why the square root law may become a poor approximation were identified. It was shown that, for theoretical reasons, when the ratio of the largest to the (common) smaller demands exceeds two, then the square root law approximation may become poor. Since such ratios are commonly observed in practice, the qualification is an important one.

Although the theory was developed for the situation when (n-1) of the n depots have the same minimum demands, the results are also applicable to the general case as a lower bound. For example, if 95% of the savings predicted by the square root law are attainable when the split of demands between depots is 20 : 20 : 60, then *at least* 95% of the savings will be achieved by any combination of demands at three depots with a minimum value of 20%.

Further analysis of centralising from more than two depots to one showed that the robustness of the cycle stock square root law improves as the number of depots to be centralised increases. If the ratio of the smallest to the largest demand sizes is less than two (strictly 2.29), then at least 95% of the square root law savings may be achieved. Tables have been presented to give more precise estimates of the accuracy of the square root law in a range of situations which may be observed in practice.

# CHAPTER 5

## *Criticisms of the Economic Order Quantity : Implications for Cycle Stock Centralisation Models*

### 5.1 Introduction

In the previous chapter, cycle stock centralisation models based on the principles of the Economic Order Quantity (EOQ) were examined. However, the EOQ approach has been subject to a number of criticisms over the years, for example by Eilon (1964) and Woolsey (1988). The relevance of these criticisms to cycle stock centralisation models will be analysed in this chapter.

An alternative approach is afforded by the maximisation of return on investment (ROI). Recently, Trietsch (1995) has derived closed analytic solutions to models based on ROI maximisation. The implications of these results for inventory centralisation will be examined.

### 5.2 An overview of criticisms of the EOQ approach

An early critique of the Economic Order Quantity approach was published by Eilon (1964). He made a number of criticisms :

* The model is built on unrealistic assumptions; in particular, a multi-product approach may be more appropriate than a single-product approach.

* Determination of the cost parameters is problematic.

* The 'minimum cost' criterion may not coincide with company objectives (eg maximising return).

* In implementing such a model, an 'optimal range' is more appropriate than an 'optimal point'.

The above criticisms have been advanced from within the management science discipline. Other criticisms have been made from outside the discipline. Firstly, it has been argued that the EOQ approach is not understood by practitioners (Woolsey (1988)). Secondly, Burnham and Mohanty (1990) have suggested that the EOQ rule has been interpreted too parochially in different links of the distribution chain and a 'Unified Order Quantity' (UOQ) approach is needed. Finally, many authors have argued (eg Weiss (1990)) that the EOQ rule has encouraged managers to concentrate on optimising the order quantity, rather than managing reductions in set-up costs.

If an EOQ approach is not adopted at the decentralised depots, then it is clearly not appropriate for determining cycle stock savings. In the following sections, criticisms of the EOQ approach will be reviewed, assuming it is used at decentralised depots and for the assessment of cycle stock savings from centralisation.

## 5.3 Economic Order Quantity assumptions

### 5.3.1 Overview of EOQ assumptions

Eilon (1964) summarised the EOQ model assumptions as follows :

1. The total cost function is a *continuous* function of the order size.
2. The total cost function is of the form :

$$\text{Total Cost} = c + \frac{s}{Q} + K\,Q$$

where c is the unit cost
Q is the order quantity
s is the set-up cost
K is a carrying-cost parameter.

3. The parameters c, s and K are independent of the order quantity, Q.

4. Each product is considered on its own.

However, this is not a complete list of the EOQ model assumptions. A more comprehensive account may be provided by incorporating the assumptions identified by Kingsman (1983) and Silver and Peterson (1985). An integrated set of assumptions is given below :

1. Demand is constant and deterministic.
2. The cost function is continuous and the order-quantity may be non-integral.
3. The unit variable cost (c) does not depend on the order quantity (Q).
4. Set-up (s) and carrying-cost (K) do not depend on the order quantity (Q).
5. The cost parameters do not vary over time (in particular, inflation is ignored).
6. Each item is treated independently of other items.
7. The lead-time is of zero duration.
8. No shortages are permitted.
9. The entire order quantity is delivered.
10. Payment for stock is made on arrival of the order.

Each of these assumptions will be examined and the main variations in EOQ models arising from relaxation of assumptions will be discussed. Since so many variants have been developed, only the most important are considered in this review.

### 5.3.2 Constant demand

In his paper on the EOQ formula, Woolsey (1988) commented that the assumption of constant demand is unsustainable : *"I have now worked and taught on five continents and I have yet to see an example of constant demand. We know perfectly well that the demand pattern looks more like a graph of earthquakes measured on the Richter scale ..."*. It cannot be denied that highly variable demand patterns are often observed in practice. Furthermore, it has long been known that, for a dynamic system, an (s,S) model is optimal for variable demand (see Scarf (1960) and Ingelhart (1963) for the demonstration of optimality).

Roberts (1962) showed that the optimal order quantity, Q, for an (s,S) model satisfies the relation :

$$Q = \sqrt{\{ (2 s \mu) / K \}} + o(Q)$$

where s and K are the set-up and carrying costs and the order notation, $f(Q) = o(Q)$ denotes that a function, f, has the property that $f(Q)/Q \to 0$ as $Q \to \infty$.

Roberts' approximation of the EOQ to the optimal order quantity motivated the 'power approximation' approach of Ehrhardt (1979). The revised 'power approximation' (Ehrhardt and Mosier (1984)) is of the following form :

$$Q = 1.30 \mu^{0.494} (s / K)^{0.506} (1 + \sigma_L^2/\mu^2)^{0.116}$$

where $\mu$ is mean lead-time demand
$\sigma_L^2$ is the variance of lead-time demand
s,K are the set-up and carrying costs.

Ehrhradt and Mosier claimed an excellent fit to their simulated data using the revised power approximation. Recently, Sani and Kingsman (1995) simulated the performance of a number of (s,S) methods against real data. Although relatively simple, the revised power approximation was found to perform as well as the other methods tested and to perform better in some few cases.

Assuming that parameters are unchanged by centralisation and demand is uncorrelated between decentralised depots, the cycle stock ratio is given by :

$$\frac{\text{Total Decentralised Cycle Stock}}{\text{Centralised Cycle Stock}} = n^{1-0.494} \frac{[\,1 + \sigma_L{}^2/\mu^2\,]^{0.116}}{[\,1 + \sigma_L{}^2/n\mu^2\,]^{0.116}}$$

$$\approx \sqrt{n}$$

The accuracy of the approximation depends on the number of decentralised depots (n) and the variance to mean-squared ratio of lead-time demand ($\sigma_L{}^2/\mu^2$). The above result quantifies the robustness of the square root law for cycle stocks to a deviation from the assumption of constant demand for an (s,S) model.

Research has also been conducted on the robustness of the constant demand assumption for an order quantity /re-order point system (an (s,q) system in the notation of chapter 1). Zheng (1992) showed that, if the inventory position in steady state is uniformly distributed on (s,s+q] and independent of the lead-time demand, then the proportionate increase in the cost function of using the EOQ formulae instead of the optimal solution is no more than 1/8 (12.5%). Examination of numerical examples showed no relative error greater than 2.9%. Axsäter (1994) sharpened the upper bound to ($\sqrt{5}$-2)/2 (11.8% approximately).

### 5.3.3 Continuity and non-integral solutions

The EOQ formulation assumes that total costs are a continuous function of the order quantity, Q, whereas, in practice, integral values (or multiples of pack sizes) must be used. Silver and Peterson (1985) commented that, "*... it can be shown mathematically ... that the best integer value of Q has to be one of the two integers surrounding the best (non-integer) solution ... rounding produces trivial percentage cost penalties except possibly when the EOQ is in the range of one or two units*". Hence, except for the slowest moving items, the ratio of EOQ values should provide a good approximation to the cycle stock ratio.

### 5.3.4 Quantity discounts

Quantity discounts introduce discontinuities (break-points) in the total cost function. Although it is possible that the EOQ will give the 'global minimum', it does not do so necessarily. In the special case that the global minimum (decentralised stocks) is at the EOQ *and* the global minimum (centralised stocks) is also at the EOQ, the square root law for cycle stocks may be used. However, in all other cases, the EOQ formula cannot be used and the square root law does not apply.

### 5.3.5 Independence of cost parameters and order quantity

The EOQ formula assumes that the parameters s (set-up cost) and K (carrying-cost parameter) do not depend on Q (order quantity). Relatively little has been written on this matter, but Aucamp (1982) examined the case where the set-up cost, s, depends on the volume of inventory to be transported. He postulated a set-up cost function of the form $C_0 + nF$, where $C_0$ is the ordering or set-up cost, F is the fixed cost per transportation vehicle ('car') and n is the number of cars required. He assumed that n could be found by calculating $Q / Q_0$ where Q is the quantity to be transported and $Q_0$ is the vehicle capacity of one car.

Under the assumption of constant demand, Aucamp found the EOQ formula to be robust. He remarked, *"What may be surprising at first glance is, no matter what the fixed charge F, $Q^*$ [the optimal order quantity] falls somewhere within a partial carload of the EOQ value".* Hence, if order-shipments are small (low number of carloads), then the ratio of EOQ values may be inaccurate, but otherwise it serves as a good approximation to the ratio of optimal cycle stocks.

5.3.6 Inflation

Buzacott (1975) investigated the effect of inflation on the EOQ. He showed that, if price is subject to the same inflation as costs, then the optimal time between placing orders, T, may be found from the following equation :

$$T = \left( \frac{2S_0 (1 + kT)}{DC_0 (r-k)} \right)^{1/2}$$

where $S_0$ and k are the parameters in the set-up cost inflation model :

$$S(t) = S_0 e^{kt}$$

$C_0$ and k are the parameters in the purchase cost inflation model :

$$C(t) = C_0 e^{kt}$$

D is the constant demand per unit time

r is the inventory carrying cost in £/£/unit time.

The square root relationship does not apply, since T appears in both the right and left hand sides of the above equation. However, Buzacott notes that for most k and T, the factor $(1 + kT)^{1/2}$ is close to unity and has only a minor effect on the EOQ.

In the context of centralisation models, if the factor $(1 + kT)^{1/2}$ is close to unity before and after centralisation, then the ratio of the decentralised to centralised factor will also be close to one. Hence, if inflation factors are assumed to be equal at all depots and not affected by centralisation, the 'square root law' for cycle stocks is not sensitive to inflation effects.

### 5.3.7 Single-item and co-ordinated replenishments

In practice, order costs may be reduced by ordering for a family of items (eg from the same supplier) at the same time. Since the EOQ approach ignores such interaction between items, the formula may give severely sub-optimal results if not adapted for this situation. However, the problem is not intractable. Assuming constant demand, constant cost parameters and an infinite planning horizon, Goyal (1974) developed a procedure to find the optimal order quantity for joint replenishment of items. Silver (1975) derived an approximate procedure which he claimed gave results at or near the optimal solution. The procedure has four steps, summarised below:

1. Order the items in ascending values of $a_i / D_i v_i$

2. Let $m_1 = 1$

   Calculate $m_i = \left( \frac{a_i}{D_i v_i} \frac{d_1 v_1}{A + a_1} \right)^{1/2}$

   and round to the nearest integer greater than zero.

3. Calculate $T^* = \left( \frac{2 (A + \Sigma a_i / m_i)}{r \Sigma m_i D_i v_i} \right)^{1/2}$

4. Determine the order quantities for items i = 1,2, ..., n by :

   $$Q_i = D_i T^*$$

where $D_i$ is the (constant) demand for item i
$v_i$ is the unit variable cost of item i
A is the major set-up cost for the family of items
$a_i$ is the minor set-up cost for item i
r is the inventory carrying cost in £/£/unit time.

The effect of centralising stocks, assuming equal demand, unit variable costs and minor set-up costs at all decentralised depots will now be considered. Since the ascending ordering of items, determined by step 1 above, is the same at each decentralised depot, then it will be unchanged at the centralised depot. Similarly, in step 2, the ratio $d_1 / d_i$ is unchanged by centralisation and so the $m_i$ values are unchanged. Hence, applying steps 3 and 4, it follows that :

$$\frac{\text{Total Decentralised Cycle Stock}}{\text{Centralised Cycle Stock}} = \sqrt{n}$$

So, when demand is constant (and equal at all depots), the square root law holds. The law is approximate, since Silver's procedure does not guarantee optimality.

### 5.3.8 Lead-time of zero duration

If demand is constant, extension of the EOQ to a known non-zero delivery lead-times presents no problems. The case of variable demand and non-zero delivery times has been discussed in sub-section 5.3.2.

### 5.3.9 Shortages not permitted

If demand is constant and, in contravention to the assumption, shortages are permitted, the point at which an order is made ('re-order level') must be specified as well as the order quantity. Naddor (1966) showed that, for continuous demand, the optimal order quantity is directly proportional to the square root of demand in both the 'back-orders' and 'lost sales' cases. As before, for discrete demand, the solution for the continuous case is a good approximation unless demand is very low.

### 5.3.10 Partial supplies

Silver (1976) investigated the situation where the amount supplied does not necessarily match the amount ordered. There are a number of reasons why this may occur in practice. The suppliers may have inadequate stock, despatch errors may have been made or goods may have been damaged or lost in transit. Silver defines a quantity called 'bias' to quantify the proportion of an order which is received :

$$\text{Bias (b)} = \frac{\text{Expected amount received}}{\text{Expected amount ordered}}$$

Since suppliers who under-supply rarely do so consistently, the standard deviation of the amount received, given the amount ordered ($\sigma_{y|Q}$), must also be estimated. Silver distinguishes between two cases :

1. Standard deviation of amount received is independent of the amount ordered.
2. Standard deviation of amount received is proportional to the amount ordered.

In the first case (when $\sigma_{y|Q} = \sigma_y$), the optimal order quantity, $Q^*$, is given by :

$$Q^* = \frac{1}{b}\left(\frac{2\,A\,D}{v\,r} + \sigma^2\right)^{1/2}$$

where the normal EOQ $= \left(\frac{2\,A\,D}{v\,r}\right)^{1/2}$

and $\sigma = \sigma_{y|Q} = \sigma_y$.

Hence, in the special case of having under-supply ($b < 1$) but no supply variability ($\sigma^2 = 0$), it is clear that the order quantity is proportional to the square root of demand. If it is assumed that the bias factor is unchanged by centralisation, then the square root law holds exactly. However, if there is bias and variability, then the ratio of decentralised to centralised cycle stocks is given by :

$$\frac{\text{Decentralised CS}}{\text{Centralised CS}} = \sqrt{n} \left( \frac{1 + R}{1 + R/n} \right)^{1/2}$$

where n is the number of decentralised depots
$R = \sigma^2 / EOQ^2$

In this case, the square root law no longer holds exactly, since it is perturbed by the second factor in the right hand side of the above equation.

In the second case, where the standard deviation of the amount received is proportional to the amount ordered, the optimal order quantity is given by :

$$Q^* = \left( \frac{2 A D}{vr\,(b + \sigma_1^2)} \right)^{1/2}$$

where the parameters are as defined above, except for $\sigma_1$, the scalar factor relating the standard deviation of the amount received to the amount ordered (ie, $\sigma_{y|Q} = \sigma_1 Q$). In this case, the optimal order quantity is proportional to the square root of demand, D. As before, if all parameters are unchanged by centralisation, then the square root law holds exactly.

In summary, following Silver's analysis, if the standard deviation of amount received is independent of the amount ordered and the EOQ parameters are unchanged by centralisation, then :

* If there is bias, the bias factor is unchanged by centralisation and the standard deviation of the amount supplied is zero, then the square root law holds exactly.

* If the standard deviation is greater than zero and independent of the amount ordered, then the square root law is no longer exact but holds approximately.

If the standard deviation of amount supplied is proportional to the amount ordered and the bias factor, scalar factor and EOQ parameters are unchanged by centralisation, then the square root law holds exactly.

5.3.11 Payment terms

Kingsman (1983) notes that the EOQ model assumes the cost of stock to be directly proportional to the average stock level over time. In practice, this assumption does not always hold. For example, if payment is at a set day (eg 15th) of the following month, then the increases and decreases in capital investment in stock do not mirror the physical stockholding. Kingsman formulates a general model to solve this problem, assuming a constant demand rate. The best policy is to order in lots of an integer number of months' demands, with a minimum order quantity of one month's supply. Kingsman shows that the EOQ, rounded to the nearest integer of months' supply, gives a good approximation to the optimal result.

Carlson and Rousseau (1989) criticise Kingsman's model on the grounds that the financing and other variable inventory costs should be treated separately. Kingsman used a single interest rate, i, but Carlson and Rousseau suggest using two rates: $i_1$ (cost of capital) and $i_2$ (variable holding costs other than capital). This is necessary since the cost of capital is a function of the period of investment, but the other costs (eg damage, insurance etc) are a function of the period of possession. The authors show that if 'day terms' apply (eg payment within 30 days of receipt), then the classical EOQ result applies. If 'date terms' apply (eg payment by the last day of the following month), then the classical result does not hold. Carlson and Rousseau propose a search procedure to find the EOQ and conclude that 'multiple EOQs' are possible in some circumstances.

Kingsman (1991) agrees that the separation of the two rates of interest is necessary. However, he shows that if the investment component is at least 30% of the total holding costs then, under date-terms, his approximate formula (Kingsman (1983)) remains valid. The '30% condition' will almost always hold in practice. Hence, under 'day terms' the EOQ result applies exactly. Under 'date-terms', the nearest integral multiple of a month's demand to the EOQ will give a good approximation to the optimal result.

### 5.3.12 Robustness to EOQ assumptions

The examples in the previous sub-sections show that, in a wide range of situations, the EOQ formula is robust to departures from model assumptions. Consequently, calculations based on the EOQ rule, with parameters suitably amended for the particular situation, are good approximations to the optimal order quantity. Of course, the above review is not comprehensive. Many models, representing different deviations from the assumptions, have not been discussed; nor have combinations of departures from model assumptions (eg multi-item models for variable demand). However, the review does illustrate that cycle stock rules based on EOQ formulae are relevant to many situations which arise in practice.

## 5.4 Determination of cost parameters

Woolsey's (1988) main line of criticism of the EOQ method is that the variables required for the formula cannot be estimated with any reliability. Similar criticisms were advanced by Eilon (1964) almost twenty five years earlier. Woolsey argues, further, that the traditional defence of insensitivity of the EOQ formula is flawed; he suggests that inaccuracies in parameter estimation may have serious consequences. In this section, these arguments will be reviewed and the relevance of the criticisms to centralisation models will be assessed.

### 5.4.1 Difficulties in parameter estimation

There is no doubt that estimation of each of the parameters presents its own difficulties in practice. EOQ theory assumes that prices are known and fixed over time. In practice, the prices of some goods, especially raw materials, depend on the day the order is placed and the time of order-delivery.

Eilon (1964) commented on the difficulty in estimating the cost of ordering: the marginal cost is required but, in practice, the average cost is often used. Selen and Wood (1987) recommended using regression methods (orders processed against relevant expenditures) to estimate the marginal cost of ordering.

Eilon argued that the 'interest rate' is particularly difficult to estimate. Each of the possible methods (rate for borrowing money, present rate of return on capital, return on opportunity investment, average industry return, 'target' return on capital) has its own problems associated with it.

### 5.4.2 Sensitivity of the EOQ formula

The well-known criticisms of the EOQ method which were briefly summarised in the previous sub-section are usually countered by the claim that the EOQ result is not sensitive to inaccuracies in the parameters. Woolsey (1988) remarks that as the 'interest rate' rises, so the EOQ result becomes more sensitive to errors. Using an example with a high 'interest rate', he shows that the total cost curve may become steep.

### 5.4.3. Sensitivity of the square root law

Although the estimation of parameters may be problematic in the implementation of EOQ models, square root law models for cycle stocks are not affected under certain conditions. If it is assumed that demand is identical at each decentralised depot and the parameters are the same at each decentralised depot and unchanged by centralisation, then :

$$\frac{\text{Total Decentralised Cycle Stock}}{\text{Centralised Cycle Stock}} = \sqrt{n}$$

and since the ratio is not a function of any of the parameters, they do not need to be estimated. A similar result applies when demand is not identical at all decentralised depots. The ratio is independent of the parameters.

## 5.5 Maximisation of return as an alternative criterion

### 5.5.1 Alternative criteria to cost-minimisation

Eilon (1964) criticised equating 'economic performance' with 'minimum costs' and proposed four possible criteria :

1. Minimum cost,
2. Maximum profit per batch,
3. Maximum return (ratio of the profit to the cost of producing a unit),
4. Maximum rate of return (proportional to the return divided by the time taken for the batch to be consumed).

However, Tate (1964) objected to the second and third criteria on the grounds that they "*ignore the times over which the different profits were realized*". Tate also claimed that, if the cost of capital tied up in stock is equal to the rate of interest obtained, then maximising the rate of return (Eilon's fourth criterion) is equivalent to minimisation of costs. This claim has recently been challenged by Trietsch (1995), who shows that different results are obtained from the two criteria. Trietsch also claims that maximising the rate of return does not necessarily lead to a square root relationship with demand. In this section, Trietsch's work will be reviewed and its implications for cycle stock centralisation models will be assessed.

### 5.5.2 Trietsch's return on investment maximisation model

Trietsch (1995) defines 'return on investment' (ROI) as the ratio of profit before tax to total equity, including inventory. He denotes the optimal order quantity under the criterion of ROI maximisation as 'ROQ'. Trietsch concedes that, if inventory is *not* included in the equity, then ROI maximisation is equivalent to cost minimisation and ROQ = EOQ. However, if inventory is included in the equity, it is claimed that : ROQ $\leq$ EOQ.

Trietsch expresses the ROI definition in algebraic terms :

$$ROI = \frac{\text{Annual Profit}}{\text{Equity}} = \frac{D\,(M - S/Q) - F - (J - I)(CQ / 2)}{L + CQ}$$

where :

| | |
|---|---|
| D | is annual demand |
| M | is marginal profit per item |
| S | is set up or re-order cost |
| Q | is re-order quantity |
| F | is annual fixed costs |
| J | is annual holding rate per £ invested in inventory, excluding interest charges |
| I | is the rate of return earned by free funds |
| C | is the unit cost |
| L | is the owners' equity, excluding inventory. |

and obtains the following result for a single-item model:

$$ROQ = \frac{LDS}{\{\,(CDS)^2 + CDLS\,[DM - F + (J\text{-}I)\,L/2]\,\}^{1/2} - CDS}$$

This may also be expressed :

$$ROQ = \frac{CDS + \{\,(CDS)^2 + CLDS\,[DM - F + (J\text{-}I)\,L/2]\}^{1/2}}{C\,[\,DM - F + (J\text{-}I)\,L/2\,]}$$

Trietsch concludes that the ROQ is not a simple square root function of D. Indeed, in the special case that F = (J-I)L/2, the ROQ is not a function of D at all.

### 5.5.3 Implications of Trietsch's model for cycle stock centralisation

In the following discussion, it will be useful to distinguish between two possible situations. In the first situation, some inventories are centralised but none of the original depots are closed. In the second situation, some depots are closed. This distinction is important since, in the second situation, the total equity of the firm (excluding inventory) is affected by centralisation as well as the stock-holding. In the first situation, centralisation does not necessarily have an effect on equity.

In the first situation, if it is assumed that all parameters are identical at decentralised depots and unchanged by centralisation, then the following general inequality applies:

$$\sqrt{n} \quad < \quad \frac{\text{Total Decentralised Cycle Stock}}{\text{Centralised Cycle Stock}} \quad < \quad n$$

The proof of this inequality is given in Appendix 5.1.

The result is useful since it shows that the 'square root law' is a lower bound for the cycle stock reduction by inventory centralisation if maximisation of return, as defined above, is the objective.

In the second situation, when at least one depot is closed, the equity and fixed costs cannot be assumed to be unchanged. Simple lower and upper bounds are not available in this situation.

## 5.6 The 'optimal range' of order quantities

Eilon (1964) recommended that the EOQ result should not be taken too literally and that an 'optimal range' of possible order quantities is more useful to managers than a single 'optimal point'.

The use of an 'optimal range' highlights the approximate nature of cycle stock centralisation models. For any individual stock-keeping unit, an order quantity above or below the EOQ may be more convenient in practice, for a variety of reasons. However, with the exception of the slowest moving items, there is no reason to suppose the EOQ formula to be biased. For example, if a non-integer value is returned by the formula, the next integer value above or below the EOQ may be used; similarly, the next multiple of the pack size below or above the EOQ may be used. In the case of slow-moving stocked items, if a value of less than one (or less than the pack size) is returned, then of course the EOQ must be rounded up. Hence, with the exception of the slowest moving items, the order quantity should be approximated reasonably well by the EOQ.

## 5.7 Criticisms of the EOQ from a managerial perspective

In this section, the three 'managerial' criticisms presented at the beginning of the chapter will be reviewed and their relevance to cycle stock centralisation models will be assessed. The criticisms are :

* The EOQ approach is not understood by practitioners.
* A Unified Order Quantity is more appropriate than an Economic Order Quantity.
* The EOQ approach has encouraged managers to optimise order quantities rather than manage reductions in set-up costs.

### 5.7.1 Understanding of the EOQ approach

One of the virtues of the EOQ formula is its simplicity. The simple graphical representation of order costs, cycle costs and total costs make it easy to teach and use. However, Woolsey (1988) commented on the difficulty he had explaining the method to a US Air Force supply sergeant. When the sergeant asked how the EOQ came from the total cost formula, he was, not surprisingly, mystified by Woolsey's explanation involving first and second derivatives.

Woolsey's anecdote misses the point. Almost all analytical models may be criticised for being something of a 'black box' to practitioners. It can be argued, though, that if practitioners understand the assumptions of the model and how to interpret the results, then the approach may be applied successfully.

In the case of cycle stock centralisation models, similar arguments apply. For practitioners versed in the necessary mathematics, a derivation may be readily given.

For all practitioners, the assumptions should be well understood in order that the results are interpreted correctly.

### 5.7.2 Unified Order Quantity

Burnham and Mohanty (1990) commented that independently determined EOQs for each link in the distribution chain cannot be superior to order quantities developed across the entire system using a Unified Order Quantity (UOQ) approach. If the EOQs are determined independently, then the purchase order quantity, the make order quantity, the transport order quantity, the receiving order quantity, the assembly order quantity and the delivery order quantity may not integrate well. Alternatively, a Unified Order Quantity approach would be based on optimising the total costs across the chain, with each link using its own UOQ (an integer multiple of a common 'basic conversion unit' (BCU)) but determined with reference to all the links in the chain.

Burnham and Mohanty's arguments were inspired by their experiences in a manufacturing environment. In a warehousing / wholesaling environment, there will be fewer links in the logistic chain to consider but the unified order approach remains valid. The determination of pack-sizes, in particular, should take into account costs of shipment, inbound and outbound, and of handling.

If a Unified Order Quantity approach is used, then inventory centralisation may lead to the use of a different 'basic conversion unit' (BCU), using Burnham and Mohanty's nomenclature. This will make the square root law approach more approximate. However, following the arguments on the 'optimal range', the rounding should not give biased results, overall, except for the slowest moving items.

### 5.7.3 Optimisation instead of management

Weiss (1990) argued that the EOQ approach has encouraged Western managers to take a 'tunnel vision' approach to inventory control, focusing entirely on the order quantity and ignoring the potential benefits of set-up cost reductions. Japanese managers, on the other hand, have reaped the benefits of dramatic reductions in manufacturing set-up times and, hence, costs.

In a warehousing/ wholesaling environment, however, order costs have decreased over recent decades in the West as computerised inventory systems have been implemented. Managers have understood that fixed ordering costs are not a recommendation of the EOQ but merely an assumption.

In the context of inventory centralisation models, the square root law may be used in two ways :

1. If ordering cost reduction is not at issue, then assume it remains unchanged by centralisation and use the familiar square root law.

2. If order costs may be reduced, for example by the introduction of an improved management information system, then use the following form of the square root law :

$$\frac{\text{Total Decentralised Cycle Stocks}}{\text{Centralised Cycle Stocks}} = [\, n\,(A_D \,/\, A_C)]^{1/2}$$

where $A_D$ set-up (order) cost at decentralised depots
$A_C$ set-up (order) cost at centralised depot.

In this context, if the first model is used, ignoring ordering costs, when the second model is appropriate, then it is the use of the square root law which should be criticised and not the model itself.

## 5.8 Conclusions

In this chapter, the EOQ model assumptions have been examined. It has been shown that, with the exception of 'quantity discounts', inventory amalgamation models are robust to departures from each model assumption. However, departures from combinations of model assumptions have not been investigated.

Although the determination of cost parameters may be problematic in the EOQ approach, it has been shown that, provided the parameters remain unchanged by inventory amalgamation, the parameters cancel and, therefore, do not present difficulties for amalgamation models.

If the criterion of cost minimisation is replaced by return on investment maximisation (following Trietsch (1995)), then the EOQ model no longer applies, and the square root form is no longer valid. Although the square root law cannot be applied in this case, lower and upper bounds have been derived for centralisation models, assuming that the decentralised depots are not closed and all parameters are identical at decentralised depots and unchanged by inventory centralisation.

The 'managerial' criticisms of the EOQ formula have been shown to be surmountable for inventory amalgamation models if the formula is applied with an appreciation of the whole supply chain and the importance of ordering cost reduction.

# PART II

# INVENTORY

# SERVICE MODELS

## Summary of Part II

In this part of the thesis, issues of operationalising inventory service models are examined. Firstly, in chapter 6, the literature on the effect of centralisation on service is reviewed. The findings reported in the literature are not useful to practitioners since only one example of disbenefits is given and the conditions under which the example applies are unlikely to occur in practice. Some results in ascertaining general conditions for inventory service benefits / disbenefits are presented.

The results given in chapter 6 are conjectural, since full mathematical proofs have not yet been derived. However, partial proofs have been obtained and extensive simulations have not revealed any counter-example to the general condition that :

*The probability of stock-out never deteriorates by centralisation if the total stock is unchanged, provided that the mean demand at a decentralised depot is less than $m + ln(2)$, where $m$ is the number of items in stock at each decentralised depot. For such mean demand values, the 'probability of stock-out' is less than 1/2.*

Since the maximum acceptable risk of a stock-out is almost always less than 50% in practice, concerns about deterioration in service through centralisation are unfounded if this conjectural result is valid.

The second aspect of operationalisation considered in this part of the thesis is the measurement of service level. Models presented in the literature generally assume a 'probability of stock-out' constraint, but many different measures are used in practice. Indeed, there are strong arguments that more than one measure should be used, to stop the system being wilfully abused, as in the example cited in sub-section 0.1.1. In a centralisation context, measures may differ between decentralised depots or the measure may change after centralisation. Results are given in chapter 7, establishing algebraic relationships between six of the most commonly used measures.

# CHAPTER 6

## *Benefits and Disbenefits of Inventory Centralisation*

### 6.1 Introduction

At the beginning of chapter 2, it was noted that two questions have been addressed in the inventory literature :

* Under what conditions does centralisation yield benefits ?
* What effect does centralisation have on stock-holding, assuming that service remains unchanged ?

The literature on the second question has been discussed. In this chapter, the first question is addressed. A further issue which may be addressed arises from slow moving SKUs becoming faster-moving SKUs after centralisation. This results in statistical benefits of estimating parameters of faster-moving products. However, this is an issue which is outside the scope of this thesis.

From a managerial viewpoint, it is important to identify the criteria against which centralisation benefits may be assessed. Using the classification developed in chapter 1, 'benefits' may be evaluated according to service criteria ('SERV' models), or cost criteria ('COST' models). Hence, the two models will be treated separately in this chapter.

The results in the literature on the conditions under which benefits arise from centralisation, using cost-based criteria, depend on assumptions of concave cost-functions. In this chapter, the situations for which such assumptions hold are discussed. The conditions under which benefits arise from centralisation, using service criteria, are delineated by finding 'break-points' for benefits / disbenefits.

## 6.2 Review of cost-based models

### 6.2.1 Eppen's model for normal demand

As discussed in section 2.3, for normally-distributed demand and a cost-function with holding and penalty components, Eppen (1979) showed that :

$$\frac{TC_D}{TC_C} = \frac{K \, \Sigma \, \sigma_i}{K \, [\, \Sigma \, \sigma_i^2 + 2 \, \Sigma\Sigma_{i<j} \, \sigma_i \sigma_j \rho_{ij} \,]^{1/2}}$$

where $TC_D$ is the total cost relevant to stocks in the decentralised system
$TC_C$ is the total cost relevant to stocks in the centralised system
$K$ is a factor determined by the unit holding and penalty costs.

The full set of model assumptions required for this result is given in sub-section 2.3.2. The list includes the assumptions identified by Eppen and a further assumption, omitted by Eppen, that the K parameters are calculated on the basis of minimising total costs (ie holding and penalty costs). Naddor (1966) showed that minimising total costs leads to an order level at which the cumulative distribution function attains a value of p/(p+h) where p and h are the penalty and holding costs. Since it is assumed, in Eppen's model, that p and h are identical at all locations, the K parameters are also identical and cancel in the above equation, yielding :

$$\frac{TC_D}{TC_C} = \frac{\Sigma \, \sigma_i}{[\, \Sigma \, \sigma_i^2 + 2 \, \Sigma\Sigma_{i<j} \, \sigma_i \sigma_j \rho_{ij} \,]^{1/2}} \, .$$

It follows that the total costs in the decentralised system are greater than or equal to the costs in the centralised system, equality being attained if $\rho_{ij} = 1$ for all i and j.

### 6.2.2 Chen and Lin's model for concave costs

Chen and Lin (1989) extended the work of Eppen (1979) by presenting a model of inventory centralisation which applies to concave cost-functions and any probability distribution of demands. Four of Eppen's assumptions - given in full in sub-section 2.3.2 - are retained, namely :

1. Demand is assumed to occur at the beginning of each scheduling period immediately after the inventory has been raised to the recommended level.

2. Demand is correlated between locations with covariances $\sigma_{jk}$ and correlation coefficients $\rho_{jk}$.

3. Safety factors are calculated on the basis of minimising total costs (ie shortage and stockholding costs).

4. Holding and shortage (penalty) costs are identical at each location.

Eppen's assumptions regarding a normal distribution are dropped. His assumption of linear holding and penalty costs is replaced by :

5. Holding and penalty costs are concave (first derivative greater than or equal to zero and second derivative less than or equal to zero for all demands greater than or equal to zero).

The demand distributions are defined generally :

6. The demand at the ith location has a probability distribution $f_i(x)$ where $f_i(x) = 0$ for all $x < 0$.

7. The demand at the centralised location has a joint probability distribution $f(x_1, x_2, ..., x_n)$ where $f(x_1, x_2, ..., x_n) = 0$ if some $x_i < 0$.

Under the above assumptions, Chen and Lin prove two results :

<u>*Chen and Lin's results*</u>

*Theorem 1* The expected holding and penalty costs in a decentralised system exceed (are greater than or equal to) those in a centralised system.

*Theorem 2* If the holding and penalty costs are linear and the correlation coefficients between the demands at the decentralised locations are all equal to one, then the expected holding and penalty costs in a decentralised system are equal to those in a centralised system.

The first theorem shows that, under the assumptions specified, cost-disbenefits are not possible. This is significant since the same authors have claimed that service disbenefits are possible; this work will be reviewed later in the chapter. Both theorems are limited by lack of consideration of transport costs. However, this limitation is addressed in a subsequent paper by Chang and Lin (1991) to be reviewed in sub-section 6.2.4.

### 6.2.3 Discussion of the assumption of concave penalty and holding costs

Chen and Lin (1989) criticise Eppen's (1979) model on the grounds that "*sufficient reasons are not given to convince us that the holding cost must vary linearly with the exceeded demand*". However, no arguments are given by Chen and Lin to support their use of concave cost-functions.

It is possible to make an *a priori* case for concave penalty costs if demand is 'captive' (no lost sales, all unmet demand is back-ordered). In this case, as the amount of goods to be expedited increases, so the unit cost of expediting may be expected to decrease. If demand is not captive, then the effect of lost sales must be taken into account. Although immediate lost revenue may be assumed to vary linearly with unmet demand, the longer term effects are more difficult to estimate. It is possible that larger quantities of unmet demand may be linked with a convex loss function of future revenue, since the chance of a customer switching to an alternative supplier is likely to increase as unmet demand increases.

Holding costs have been taken as linear by most authors (see, for example, Lambert and Stock (1993)). However, it is possible that some cost elements, such as insurance, may show economies of scale. On the other hand, the cost of obsolescence may be linked with a convex function since the risk of obsolescence increases as excess stock rises.

In conclusion, the assumption of concave cost-functions may be valid for captive demand items which are not at risk of obsolescence (eg non-deteriorating items rarely subject to design modification). Otherwise, the concave cost formulation may not necessarily hold.

### 6.2.4 Chang and Lin's model incorporating transport costs

Chang and Lin (1991) extended the work of Chen and Lin (1989) by comparing the decentralised system (inventory transfers between locations not permitted) with a 'centralised system' with the same deployment of stock at decentralised locations but with inventory transfers permitted. It should be noted that Chang and Lin's usage of the term 'centralised' differs from the usage in the remainder of this thesis.

The same assumptions were adopted as in the work of Chen and Lin, with the additional assumption of concave transport costs. Under these assumptions, Chang and Lin prove three results (where x is demand, t(x) is transport cost, h(x) is holding cost, p(x) is penalty cost, $s_C$ is the optimal stock for the centralised depot, with expected costs $H_C(s_C)$ and $s_D$ is the vector of optimal stocks for the decentralised depots, with expected costs $H_D(s_D)$.

*Chang and Lin's results*

*Theorem 1* If h(x) and p(x) are concave functions with $t(x) \le \min\{h(x),t(x)\}$, then $H_C(s_C) < H_D(s_D)$.

*Theorem 2* If h(x) and p(x) are linear functions, then for any function t(x), $t(x) < h(x) + p(x)$ if and only if $H_C(s_C) < H_D(s_D)$.

*Theorem 3* If h(x) and p(x) are linear functions and all the correlation coefficients are equal to one, then for any function t(x), $H_C(s_C) = H_D(s_D)$.

These results show that cost-disbenefits are possible from 'centralisation' when transport costs are included in the formulation. The second theorem specifies the conditions for benefits / disbenefits exactly for linear cost-functions. The first theorem gives a necessary, but not sufficient, condition for cost-benefits for general concave cost-functions.

### 6.2.5 Discussion of Chang and Lin's transport cost-function

As in the previous paper by Chen and Lin (1989), Chang and Lin (1991) make no attempt to justify the assumption of concave cost functions.

Transport costs may be considered as a step-function of demand: when demand reaches certain levels, another (or larger) vehicle is required. Since such a function is discontinuous, it is not differentiable and cannot be defined as concave. However, in practice, cost-functions are often approximated by such methods as linear interpolation (see, for example, the case-study by Thornley (1978)). As throughput rises, there should be greater scope for economies of scale through the use of larger vehicles and higher vehicle utilisation rates. Therefore, it would seem sensible to assume that, although transport functions are not concave in a strict mathematical sense, they may be approximated by concave functions.

The transport-cost function is of the form t(x): transport-costs depend on throughput. However, a more realistic function would be t(x,d), including distance (d) between the two locations to estimate the transport cost. Indeed, if distance is excluded, and the locations are geographically dispersed (as would often be the case in practice), then ignoring the distance between locations may lead to large errors.

## 6.3 Review of service-based models

### 6.3.1 Stulman's model

Stulman (1987) seeks to extend the work of Eppen (1979) by replacing the minimisation of costs criterion by an 'assurance of service' constraint at each location. Stulman addresses a multi-centre newsboy problem with first come, first served allocation. The assumptions of Stulman's model are as follows :

1. An assurance of service constraint is used at each location to determine inventory levels.
2. Orders are filled on a first come, first served basis.
3. Demand at each location, j, is Poisson distributed with parameter $\mu_j$.
4. The demands at each location are independent of each other.
5. Each individual in the entire population served by the set of depots has the same probability of making a demand.
6. The time when an order is placed by an individual is distributed independently and identically for all individuals in the entire population.

Stulman uses the term 'location j' to denote customers of the jth depot prior to centralisation. After centralisation, 'stockouts at j' indicates the stockouts experienced by customers who used to be served by depot j before centralisation. Using this nomenclature, Stulman shows that :

$$P(\text{no stock-outs at } j \mid Y) = P(TD \leq Y) + (\mu_{TD} / \mu_{RD})^Y \; e^{-\mu(j)} \; P(RD \geq Y + 1)$$

where j designates the *j*th location
Y is the amount of stock at the centralised location
TD is the total demand

RD is the remaining demand when the demand at the *j*th location has been removed from the total demand

$\mu_{TD}$ is the mean total demand

$\mu_{RD}$ is the mean remaining demand

$\mu(j)$ is the mean demand at the *j*th location.

The above formula shall be called 'Stulman's theorem'. The first term represents the probability that there will be no stock-outs at any location. The second term represents the probability that there will be no stock-outs at the *j*th location, although some stock-outs will occur at other locations. The proof of the theorem is given in the appendix of Stulman's paper.

### 6.3.2 Stulman's result for normal demand

Stulman's theorem is based on Poisson demand. Stulman also showed that if $\mu(j)$ is large enough to allow the use of the normal approximation, and if the same probability of stock-out constraint applies at each location, then the centralised inventory is less than the total inventory in the decentralised system.

### 6.3.3 Stulman's conjecture for Poisson demand

Stulman was unable to prove that centralised inventory is always less than the decentralised total in the case of Poisson demand which cannot be approximated by the normal distribution. Since the mathematical analysis proved intractable, he simulated a large number of examples with varying values of k (number of decentralised locations), $\alpha$ (probability of stock-out constraint) and $\mu$ (mean demand at each location, assumed to be equal). In each of the 2000 simulated examples, the centralised inventory was less than or equal to the total inventory in the decentralised system, with equality being observed in just one case. However, since no analytical proof could be found, the assertion that the centralised inventory is no greater than the total decentralised inventory shall be called 'Stulman's conjecture'.

### 6.3.4 Chen and Lin's counter-example to Stulman's conjecture

The reason why a proof could not be found for Stulman's conjecture became clear when Chen and Lin (1990) presented a counter-example, thus showing that the conjecture does not hold for all possible assurance of service constraints. Suppose that there are two depots in the decentralised system serving equal mean demands. Suppose further that each depot holds one item in stock of a product with mean demand $\mu$, where $3 < \mu < 4$ and $\mu$ satisfies the equation :

$$\mu = 4 \sum_{j=2}^{\infty} \frac{e^{-\mu} \mu^{j}}{j\,!} = 4 \left[ 1 - ( e^{-\mu} + \mu e^{-\mu} ) \right]$$

The solution of this equation in the range (3,4) is 3.42, to two decimal places. Chen and Lin showed that, for a mean demand rate of 3.42, the probability of a stock-out at a decentralised location is 0.86 but the probability of a stock-out at the centralised location, with two items in stock, is 0.88. Thus, there are service disbenefits of centralisation.

Chen and Lin's counter-example is rather artificial. It is difficult to imagine an organisation which could prosper when the probability of a stock-out is as high as 86%. However, the counter-example to Stulman's conjecture shows that it is not universally valid.

## 6.4 A more general approach to service-based models with one in stock at each decentralised depot

### 6.4.1 Motivation for a more general approach

Chen and Lin's counter-example to Stulman's conjecture, discussed in the previous section, leaves open the question of the general conditions under which service benefits and disbenefits occur through centralisation. The counter-example is artificial since the probability of stock-out is very high. However, it is possible that disbenefits may occur under more realistic conditions. In this section, the conditions for service benefits are investigated further.

### 6.4.2 General condition for service disbenefits

The starting point for a more general approach is to establish a theorem which provides a demarcation between benefits and disbenefits :

*Theorem*

If demand is Poisson distributed with parameter $\mu$ at each decentralised depot and the conditions for Stulman's theorem are obeyed, then disbenefits arise from centralisation if and only if :

$$e^{-(k-1)\mu} \sum_{j=0}^{k-1} \frac{(k\mu)^j}{j!} \left\{ \left( \frac{k}{k-1} \right)^{k-j} - 1 \right\} \quad > \quad \left( \frac{k}{k-1} \right)^{k} - 1 - \mu$$

where k is the number of decentralised depots

m is the number of items held in stock at each decentralised depot.

A proof of this result is given in Appendix 6.1.

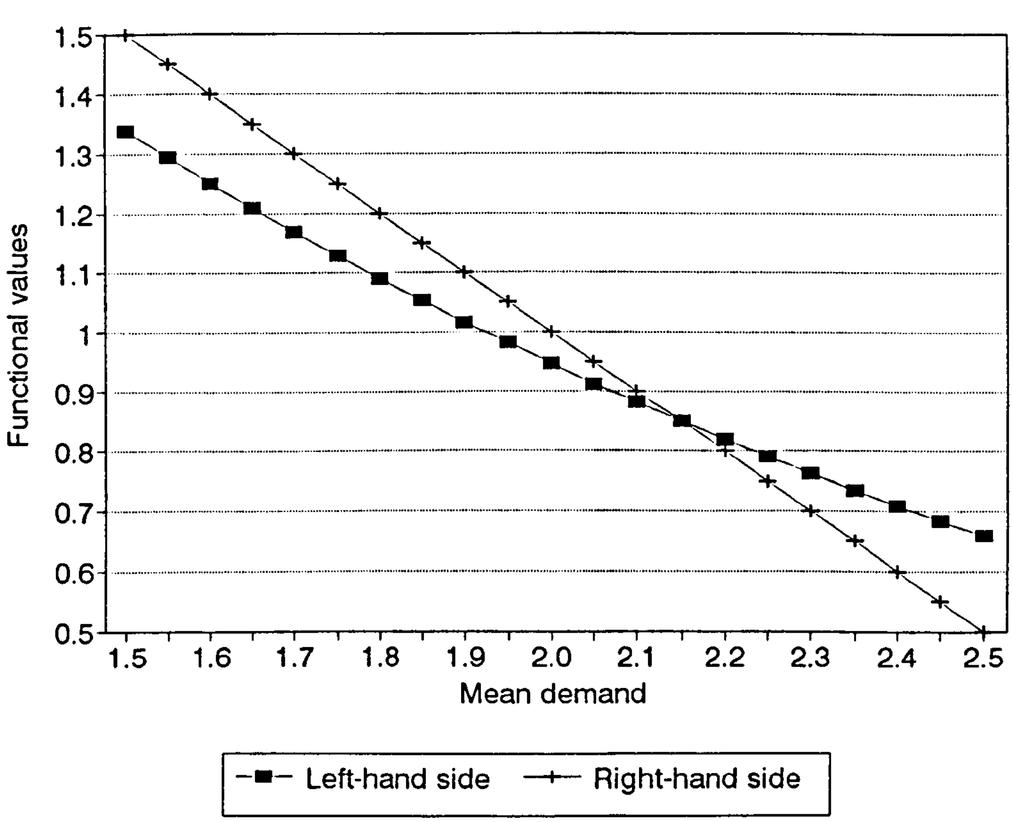
### 6.4.3 Application of the general condition to centralisation of two inventory locations

Chen and Lin's counter-example concerned the centralisation of two inventory locations. To find the complete range of mean demands over which service disbenefits occur, k = 2 is substituted in the general condition established in the theorem of the previous sub-section :

$$(2\mu + 3)\, e^{-\mu} \quad > \quad -\mu + 3$$

The functions on the left-hand and right-hand sides of the inequality are shown in graphical form in Figure 6.1 :

Figure 6.1
Benefits / disbenefits for two-depot case with one item in stock



corresponding probability of a stock-out at a decentralised location is 0.63. Hence, service disbenefits occur in the two-depot problem (with equal mean demands) if and only if 'probability of stock-out' constraints of 0.63 or greater are acceptable.

The graph also shows that a rough 'upper bound' on the break-point is provided by the point at which the function $-\mu + 3$ crosses the axis, namely at $\mu = 3$. This observation will prove useful when considering the more general case.

The counter-example of Chen and Lin, at $\mu = 3.42$, is thus shown to be an example of the general rule for two depot centralisation with equal demands :

*For Poisson demand, there are service disbenefits (higher risk of stock-out) if and only if the mean demand is greater than 2.149 (calculated to three decimal places).*

### 6.4.4 Application of the general condition to centralisation of more than two inventory locations

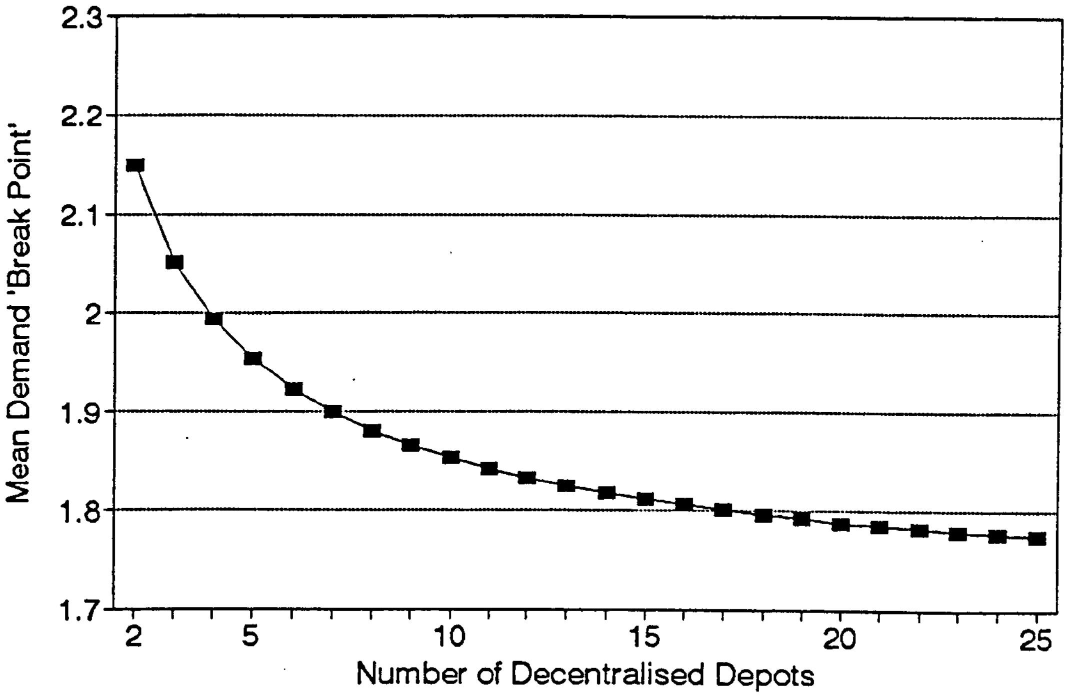
Break-points have been calculated for the centralisation of more than two depots, as shown in Table 6.1 :

TABLE 6.1
Mean demand break points for service benefits

| *Number of depots* | *Mean Demand 'Break point'* | *Prob (stock-out) 'Break point'* |
|---|---|---|
| 2 | 2.149 | 63.38 % |
| 3 | 2.053 | 60.82 % |
| 4 | 1.994 | 59.24 % |
| 5 | 1.953 | 58.11 % |
| 6 | 1.923 | 57.28 % |
| 7 | 1.900 | 56.63 % |
| 8 | 1.881 | 56.08 % |
| 9 | 1.866 | 55.65 % |
| 10 | 1.853 | 55.37 % |
| 11 | 1.842 | 54.95 % |
| 12 | 1.833 | 54.69 % |
| 13 | 1.825 | 54.46 % |
| 14 | 1.817 | 54.22 % |
| 15 | 1.811 | 54.05 % |
| 16 | 1.806 | 53.89 % |
| 17 | 1.801 | 53.75 % |
| 18 | 1.796 | 53.60 % |
| 19 | 1.792 | 53.48 % |
| 20 | 1.788 | 53.36 % |
| 21 | 1.785 | 53.27 % |
| 22 | 1.782 | 53.18 % |
| 23 | 1.779 | 53.09 % |
| 24 | 1.777 | 53.03 % |
| 25 | 1.774 | 52.94 % |

This table shows that mean demand and probability of stock-out 'break points' decline as the number of inventory locations to be centralised increase. The decline in the mean demand 'break point' is shown graphically in Figure 6.2 :

Figure 6.2
Mean demand break-points for service benefits




The graph suggests the possibility of a 'flattening off' of the curve to an asymptote at a limiting value. In the next sub-section, this limiting value will be identified.

### 6.4.5 Identification of a limiting 'break-point'

The break-points are found by replacing the inequality in the 'general condition' by equality, giving the following equation :

$$e^{-(k-1)\mu} \sum_{j=0}^{k-1} \frac{(k\mu)^j}{j!} \left\{ \left( \frac{k}{k-1} \right)^{k-j} - 1 \right\} = \left( \frac{k}{k-1} \right)^{k} - 1 - \mu$$

The left hand side of the above equation is always positive but the right hand side can be negative. This was observed graphically in sub-section 6.4.3 for the special case of two depots. Hence an upper bound for the critical value is given by :

$$\mu = \left( \frac{k}{k-1} \right)^{k} - 1$$

In the limit, this upper bound becomes :

$$\mu = e - 1$$

If it can be shown that the left hand side tends to zero as the number of depots (k) tends to infinity, then the upper bound of e-1 will also be the limiting value. Analysis using standard results of convergence theory shows that the left hand side does tend to zero (see Appendix 6.2 for a formal proof). Hence the limiting break points for mean demand and probability of stock-out are given by:

*Limiting mean value break point*

$$\mu = e - 1 = 1.718...$$

*Limiting probability of stock-out break point*

$$P(s/out) = 1 - e^{2-e} = 51.24...\ \%$$

### 6.4.6 A new conjecture

Using the results of the previous sections, Stulman's conjecture may be modified for the special case when the mean demands and assurance of service constraints are identical at each decentralised location (and the assurance of service constraint is unchanged by centralisation) and the number in stock at each decentralised location is one.

*Conjecture*

*The probability of stock-out never deteriorates by centralisation if the total stock is unchanged provided that the mean demand at a decentralised depot is less than e-1. For such mean demand values, the 'probability of stock-out' is less than* $1 - e^{2-e}$.

This conjecture differs from Stulman's in two respects. Firstly, its scope is restricted to situations where the mean demands are equal at decentralised depots. Secondly, it overcomes counter-examples such as Chen and Lin's by restricting the application to a limited range of mean values.

It is important to note that the above result is *conjectural* for the same reason that Stulman's original result was conjectural: a rigorous mathematical proof has not been found to substantiate the claim. Such a proof would require three propositions to be established :

*Proposition 1* There exists a unique 'break-point', $\mu^*(k)$, for each $k \geq 2$, such that service benefits occur for all $0 < \mu < \mu^*$ and service disbenefits occur for all $\mu > \mu^*$.

*Proposition 2* The sequence $\{\mu^*(k)\}$ is monotonic decreasing.

*Proposition 3* As $k \rightarrow \infty$, $\mu^*(k) \rightarrow e - 1$.

The third proposition has been proved mathematically (see Appendix 6.2). However, neither of the first two propositions have yet been proven. The difficulty lies in the analytic intractability of the mathematical expressions, as noted by Stulman (1987) himself.

6.4.7 Evaluation of the 'general condition' inequality to test the new conjecture

Since the new conjecture has not been proved mathematically, it would seem prudent to check it for a range of mean values and numbers of decentralised depots. A short program was written to evaluate the left hand side of the general condition :

$$e^{-(k-1)\mu} \sum_{j=0}^{k-1} \frac{(k\mu)^j}{j!} \left\{ \left( \frac{k}{k-1} \right)^{k-j} - 1 \right\}$$

and the right hand side of the general condition :

$$\left( \frac{k}{k-1} \right)^k - 1 - \mu$$

The functions were evaluated for $\mu = 0.001$ to $\mu = 2.718$, using an increment of 0.001. This evaluation was performed for each value of k running from $k = 2$ to $k = 25$. In each evaluation, the right hand side was found to be greater than the left hand side, indicating service benefits from centralisation.

## 6.5 Extension of the general approach to service-based models with more than one in stock at each decentralised depot

### 6.5.1 Extension of the general condition for service disbenefits

There is a greater risk of stock-outs from a centralised location (m * k items in stock, where the asterisk denotes multiplication) than at a decentralised depot with m items in stock (one of k depots) if and only if the following condition holds :

$$e^{-(k-1)\mu} \sum_{j=0}^{mk-1} \frac{(k\mu)^j}{j!} \left\{ \left( \frac{k}{k-1} \right)^{mk-j} - 1 \right\} \quad > \quad \left( \frac{k}{k-1} \right)^{mk} - \sum_{j=0}^{m} \frac{\mu^j}{j!}$$

where $\mu$ is the mean demand at each decentralised depot.

This result will be called the 'extended general condition'. A proof is given in Appendix 6.1.

### 6.5.2 Application of the 'extended general condition'

As in section 6.4, the 'break-points' have been calculated for benefits / disbenefits from centralisation. The calculations have been performed for a range of decentralised depots (k = 2,..,10) and a range of number of items in stock at each decentralised depot (m = 2,..,5). The results for 'mean demand break points' are presented in Table 6.2 and the results for 'probability of stock-out break points' are presented in Table 6.3 :

TABLE 6.2
Mean demand break points for service benefits from centralisation

| | *Number of items in stock at each depot* | | | |
|---|---|---|---|---|
| *Number of depots* | *2* | *3* | *4* | *5* |
| 2 | 3.166 | 4.177 | 5.185 | 6.190 |
| 3 | 3.078 | 4.095 | 5.106 | 6.116 |
| 4 | 3.023 | 4.042 | 5.056 | 6.067 |
| 5 | 2.984 | 4.005 | 5.020 | 6.032 |
| 6 | 2.955 | 3.976 | 4.992 | 6.005 |
| 7 | 2.931 | 3.953 | 4.970 | 5.983 |
| 8 | 2.912 | 3.934 | 4.951 | 5.964 |
| 9 | 2.896 | 3.918 | 4.935 | 5.949 |
| 10 | 2.883 | 3.904 | 4.921 | 5.935 |

TABLE 6.3
Probability of stock-out break points for service benefits from centralisation

| | *Number of items in stock at each depot* | | | |
|---|---|---|---|---|
| *Number of depots* | *2* | *3* | *4* | *5* |
| 2 | 61.29% | 60.03% | 59.13% | 58.43% |
| 3 | 59.40% | 58.49% | 57.79% | 57.28% |
| 4 | 58.19% | 57.47% | 56.93% | 56.50% |
| 5 | 57.32% | 56.75% | 56.30% | 55.94% |
| 6 | 56.66% | 56.18% | 55.81% | 55.51% |
| 7 | 56.12% | 55.73% | 55.42% | 55.16% |
| 8 | 55.68% | 55.35% | 55.09% | 54.85% |
| 9 | 55.31% | 55.03% | 54.80% | 54.61% |
| 10 | 55.01% | 54.75% | 54.55% | 54.38% |

Recalling the analysis of one item in stock, similar patterns may be observed. The 'mean demand break point' declines as the number of decentralised depots rises, but at rates which seem to be slowing down. The same applies to the 'probability of stock out break point'.

### 6.5.3 Identification of a limiting 'break-point'

As before, the break-points are found by replacing the inequality in the 'general condition' by equality, giving the following equation :

$$e^{-(k-1)\mu}\sum_{j=0}^{mk-1}\frac{(k\mu)^j}{j!}\left\{\left(\frac{k}{k-1}\right)^{mk-j}-1\right\} = \left(\frac{k}{k-1}\right)^{mk}-\sum_{j=0}^{m}\frac{\mu^j}{j!}$$

The left hand side of the above equation is always positive but the right hand side can be negative. Hence an upper bound for the critical value is given by the positive solution to the following equation:

$$\sum_{j=0}^{m}\frac{\mu^j}{j!} = \left(\frac{k}{k-1}\right)^{mk}$$

It may be asked whether it is meaningful to speak of *the* positive solution to a polynomial equation of degree m; there may be more than one positive solution or, indeed, none at all. However, since the left hand side of the equation takes a value of unity at $\mu=0$ (less than the right hand side) and is monotonic increasing, tending to infinity as $\mu$ tends to infinity, it is clear that there is one, and only one, positive solution to the equation.

In the limit, as k tends to infinity (fixed value of m) the equation for the upper bound becomes :

$$\sum_{j=0}^{m}\frac{\mu^j}{j!} = e^m$$

If it can be shown that the left hand of the 'extended general condition' equation tends to zero as the number of depots (k) tends to infinity, then the upper bound equation shown above will also be the equation for the limiting value. Analysis using standard results of convergence theory shows that the left hand side does tend to zero (see Appendix 6.2 for a formal proof).

The equation for the limiting value has one positive solution; this may be shown using identical arguments to those given above. It may also be shown that the root of the limiting value equation is greater than m + ln(2) (see Appendix 6.3 for the proof).

If it is known that the break-point is greater than m + ln(2), it follows that :

$$\sum_{j=0}^{m} \frac{[m + \ln(2)]^j}{j!} \leq e^m$$

and, hence, the probability of not having a stock-out satisfies the inequality :

$$\sum_{j=0}^{m} \frac{[m + \ln(2)]^j \; e^{-[m+\ln(2)]}}{j!} \leq 1/2$$

or, equivalently, the probability of a stock-out is greater than 50%.

### 6.5.4 A generalisation of the new conjecture

Using the results of the previous sections, the new conjecture presented earlier in the chapter may be generalised, maintaining the assumption that the mean demands and assurance of service constraints are identical at each decentralised location and the assurance of service constraint is unchanged by centralisation. The generalisation consists in dropping the assumption that the number in stock at each decentralised location is one, replacing it with the more general assumption of m units in stock at each location.

The generalised conjecture is :

*The 'probability of stock-out' never deteriorates by centralisation if the total stock is unchanged, provided that the mean demand at a decentralised depot is less than m + ln(2), where m is the number of items in stock at each decentralised depot. For such mean demand values, the 'probability of stock-out' is less than 1/2.*

Again, it is important to note that the above result is *conjectural* : a rigorous mathematical proof has not yet been found. Such a proof would require three propositions to be established :

*Proposition 1* There exists a unique 'break-point', $\mu^*(k,m)$, for each $k \geq 2$ an $m \geq 1$, such that service benefits occur for all $0 < \mu < \mu^*$ and service disbenefits occur for all $\mu > \mu^*$.

*Proposition 2* $\mu^*(k,m) \;\geq\; \lim_{k\to\infty}\{\mu^*(k,m)\}$ for all $k \geq 2$ and $m \geq 1$.

*Proposition 3* As $k \to \infty$, $\mu^*(k,m) \to m + \ln(2)$.

The third proposition has been proved mathematically (see Appendix 6.3). However, neither of the first two propositions have yet been proven.

6.5.5 Evaluation of the 'extended general condition' inequality

Since the generalised conjecture has not been proved mathematically, it will be checked for a range of mean values ($\mu$), numbers of decentralised depots (k) and number in stock at each depot (m). A short program was written to evaluate the left hand side of the general condition :

and the right hand side of the general condition :

$$\left(\frac{k}{k-1}\right)^k \quad - \quad \sum_{j=0}^{m} \frac{\mu^j}{j!}$$

The functions were evaluated for $\mu = 0.001$ to $m + \ln(2)$, using an increment of 0.001. This evaluation was performed for each value of k running from $k = 2$ to $k = 25$ and each value of m running from $m = 2$ to $m = 5$. In each evaluation, the right hand side was found to be greater than the left hand side, indicating service benefits from centralisation.

## 6.6 Conclusions

In this chapter, research on the benefits of inventory centralisation has been reviewed. Chen and Lin (1989) proved that cost-reduction benefits will occur if the cost-functions are concave. This is a useful result as it covers the special case of linear cost-functions. However, benefits do not necessarily occur if there are convex cost-functions. Some arguments have been presented to show that such functions may arise quite naturally in practice. The extended model of Chang and Lin (1991) covered transport functions which did not take distance into account. Hence, for any situation where there are some unequal distances between the depots, the formulation may be unrealistic.

It had been conjectured by Stulman (1987) that for service driven models (service being defined by the probability of a shortage at any given time), disbenefits could not occur from centralisation. Chen and Lin (1990) showed this conjecture to be false. In this chapter, two new conjectures have been presented :

1. For the case of one item in stock at each depot, there will be no service disbenefits if the mean demand at each decentralised depot is less than e-1 (probability of stock-out greater than $1 - e^{e-2}$).

2. For the case of more than one item in stock at each depot (m in stock), there will be no service disbenefits if the mean demand is less than $m + \ln(2)$, with a probability of stock-out greater than 1/2.

Detailed calculations have shown no counter-examples to these two conjectures, but they still await mathematical proof.

If the two new conjectures are true, then Stulman's original conjecture holds for all mean values likely to hold in practice. A probability of stock-out of greater than 50% is too high for organisations to sustain. Hence, in practice, service disbenefits will rarely, if ever, occur if the mean demand at each decentralised location is identical. Further research is required to analyse the situation when there are unequal mean demands.

# CHAPTER 7

## *Relationships Between Service Level Measures*

### 7.1 Introduction

In chapter 6, it was assumed that a 'probability of stock-out' criterion is used to measure the service level to customers. This measure is mathematically convenient but, in practice, many organisations use other measures. Indeed, the service level can be measured in many different ways. If the same measure is used at all stock-holding depots and does not change over time, then comparisons can be made using the common measure. If, however, different measures are used at the various stock-holding depots then comparison of performance is more difficult. Also, the impact of centralisation or decentralisation of depots on stock-holding is difficult to assess if a common service measure is not used. If the same measures are used at all depots but the inventory system changes from being based on one measure to another, similar difficulties arise. In this chapter, relationships between six of the most widely used service measures are presented, enabling each of these measures to be converted to a common measure. Data requirements for using the relationships are considered and, where appropriate, alternative relationships are presented which depend on data which may be more readily available in practice.

---

*Note : This chapter is based on the article by Boylan and Johnston (1994), entitled 'Relationships between Service Level Measures for Inventory Systems' (Journal of the OR Society, Vol 45, pp 838-844).*

## 7.2 Definition of measures

Many measures are used in practice and it would be impractical to analyse them all. In this chapter, some of the most commonly used measures will be analysed. Measures may be based at one of three levels : orders, order lines or units. Some measures count an order line as 'unfilled' if the line is not completely filled. For example, if a customer orders ten of a particular SKU, and only eight are delivered, the order line is considered to be 'unfilled'. Using other measures, such a line is considered to be 'partially filled' and some allowance is made for this in the definition of the measure.

Although 'partial fill' is a relevant concept at the levels of orders and order lines, it does not apply to units since no further dis-aggregation is possible at this level. Instead of basing the calculation on the volume of units supplied, the value of the units can also be used.

On the basis of the above considerations, six service level measures will be considered in detail. The six measures are given in Table 7.1 :

<u>TABLE 7.1</u>
<u>Six service level measures</u>

| | |
|---|---|
| $O_c$ | - share of orders filled completely |
| $O_p$ | - share of orders filled (at least) partially |
| $L_c$ | - share of order lines filled completely |
| $L_p$ | - share of order lines filled (at least) partially |
| U | - share of units (quantity) filled from stock |
| V | - share of value (£,$ etc) filled from stock. |

The measures given above are relevant to any inventory system in which lost sales and/or back-orders are recorded. Hill (1990) argued that this data may not always be available in a retailing environment. In such situations, the relationships requiring lost sales or back-order data given in this chapter may not be applied.

## 7.3 Previous research on service level measures

### 7.3.1 Unit fill-rate over lead-time

Ronen (1982) restated a result derived by Brown (1967), which shows that the unit fill-rate over lead-time ($U_{LT}$) is related to the overall fill rate (U) as follows :

$$U \quad = \quad 1 - \frac{L\,S}{Q}\,(1 - U_{LT})$$

where L is the lead time (years)
S is the total demand in a year
Q is the order quantity

and it is assumed that $LS < Q$ and unfilled demand is back-ordered. If unfilled demand is lost, rather than back-ordered, Ronen showed (using the same notation as above) :

$$U \quad = \quad \frac{1}{\frac{L\,S}{Q}\,(1 - U_{LT}) + 1}$$

In the analysis which follows, the period of measurement for each of the measures is fixed as one year and is not taken as the supplier lead-time. As Ronen argues, the overall measure is more meaningful to customers than its lead-time equivalent.

### 7.3.2 The use of a range of measures

Ronen further argued that an organisation should not restrict itself to using just one service measure. A single measure may be manipulated to give a distorted picture. For example, an inventory manager may buy very large quantities of low-cost fast-moving items to inflate a unit fill-rate based on volume; a different picture would emerge if a value-based measure were also in use. An example of this was given in sub-section 0.1.1. Thus, Ronen's advice is sound and, if taken at all stock-holding depots, eliminates the need for a set of relationships between different measures.

In practice, it may be difficult for an organisation to ensure that all stock-holding depots use a common set of measures. An example would be a company selling its products through a number of national sales companies, each acting autonomously or semi-autonomously. Moreover, even if all depots do use the same set of measures, a new inventory management system may be introduced which requires the use of a different (possibly overlapping) set of measures. In such circumstances, the relationships presented in this chapter would prove to be of some benefit.

## 7.4 Relationships between measures

### 7.4.1 Relationships between measures based on 'complete fill'

Relationships between the measures based on complete order and line fill ($O_c$ and $L_c$ respectively) and share of units and value filled (U and V respectively) are given below. In each case, the relationship follows almost immediately from the definitions of the measures.

$$\frac{L_c}{O_c} = \frac{l_f}{l_d}\frac{o_d}{o_f} = \frac{l_f}{o_f}\frac{o_d}{l_d} = \frac{\text{Average number of lines per order filled}}{\text{Average number of lines per order placed}}$$

$$\frac{U}{L_c} = \frac{u_f}{u_d}\frac{l_d}{l_f} = \frac{u_f}{l_f}\frac{l_d}{u_d} = \frac{\text{Average number of units per order line filled}}{\text{Average number of units per order line placed}}$$

$$\frac{V}{L_c} = \frac{v_f}{v_d}\frac{l_d}{l_f} = \frac{v_f}{l_f}\frac{l_d}{v_d} = \frac{\text{Average value of order lines filled}}{\text{Average value of order lines placed}}$$

where $o_d$ = number of orders demanded
$o_f$ = number of orders filled

$l_d$ = number of order lines demanded
$l_f$ = number of order lines filled

$u_d$ = number of units demanded
$u_f$ = number of units filled

$v_d$ = value of units demanded
$v_f$ = value of units filled.

Given the above relationships, other relationships between any two of the four measures may be obtained.

In most inventory systems, it should be straightforward to find some of the data required, namely : average number of lines per order *placed*, average number of units per order line *placed* and average value per order line *placed*. However, it may be more difficult to find the values of the averages per order *filled* or per order-line *filled*. Although past order files may be available on disk or tape, it is less likely that files of filled orders will be available. This obstacle may be overcome if the organisation keeps files of unfilled orders (in the case of 'captive demand', back-order files may be used for this purpose). The relationships have been re-worked to allow calculations when averages for unfilled orders are available but averages for filled orders are not:

$$L_c = 1 - (1 - O_c)\,\frac{l_u}{o_u}\,\frac{o_d}{l_d}$$

$$U = 1 - (1 - L_c)\,\frac{u_u}{l_u}\,\frac{l_d}{u_d}$$

$$V = 1 - (1 - L_c)\,\frac{v_u}{l_u}\,\frac{l_d}{v_d}$$

where $o_u$ - number of orders unfilled
$l_u$ - number of order lines unfilled
$u_u$ - number of units unfilled
$v_u$ - value of units unfilled

In this case, the relationships do not follow immediately and proofs are given in Appendix 7.1.

## 7.4.2 Relationships between measures based on 'partial fill'

In the case of order lines being partially filled, a new definition is required for a service level. A wide variety of measures are used in practice and the following 'general' definition is proposed :

$$L_p = \frac{l_{cf} + \alpha\, l_{pf}}{l_{cf} + (\alpha+\beta)\, l_{pf} + l_u}$$

where $l_{pf}$ - number of order lines partly filled
$l_{cf}$ - number of order lines completely filled
$\alpha,\beta$ - weighting parameters.

This definition, involving the weightings $\alpha$ and $\beta$, has been designed to cover a wide range of possible conventions. For example, for the purpose of calculating $L_p$, one partially filled line may be taken to be :

* one line, unfilled ($\alpha$=0, $\beta$=1)
* two half-lines, one filled and one unfilled ($\alpha$=0.5, $\beta$=0.5)
* two full-lines, one filled and one unfilled ($\alpha$=1, $\beta$=1)
* one line, completely filled ($\alpha$=1, $\beta$=0).

Any pair of non-negative values of $\alpha$ and $\beta$ may be used with the choice of values depending on the context; if partially filled lines are of no use to customers, then it would be appropriate to use the $\alpha$=0, $\beta$=1 convention, for example. It is not necessary that the sum of $\alpha$ and $\beta$ must be unity, as shown by the third example. This example may be visualised as the partially filled line splitting and one line being sent to a 'filled order line file' and another line to an 'unfilled order line file'. Such a measure has been experienced by the author in an internal consultancy role.

In the case of orders being partially filled, the following definition is used :

$$O_p = \frac{o_{cf} + \Theta\, o_{pf}}{o_{cf} + (\Theta+\phi)\, o_{pf} + o_u}$$

where $o_{pf}$ - number of orders partly filled
$o_{cf}$ - number of orders completely filled.

$\Theta$ and $\phi$ are chosen according to the desired weighting of partially filled orders.

The following relationship between $O_p$ and $L_p$ has been derived :

$$\frac{L_p}{O_p} = \frac{r_{cf} + \alpha\, r_{pf}}{s_{cf} + \Theta\, s_{pf}} \; \frac{1 + (\Theta+\phi-1)\, s_{pf}}{1 + (\alpha+\beta-1)\, r_{pf}}$$

where :

$$r_{cf} = l_{cf} / (l_{cf} + l_{pf} + l_u) = l_{cf} / l_d$$
$$r_{pf} = l_{pf} / (l_{cf} + l_{pf} + l_u) = l_{pf} / l_d$$

$$s_{cf} = o_{cf} / (o_{cf} + o_{pf} + o_u) = o_{cf} / o_d$$
$$s_{pf} = o_{pf} / (o_{cf} + o_{pf} + o_u) = o_{pf} / o_d$$

In the case where $\alpha+\beta = \Theta+\phi = 1$, the expression simplifies to:

$$\frac{L_p}{O_p} = \frac{r_{cf} + \alpha\, r_{pf}}{s_{cf} + \Theta\, s_{pf}}$$

The proof is given in Appendix 7.1.

As may be expected, the data requirements for relationships between 'partial fill' measures are more onerous than for 'complete fill' measures. Clearly, the weightings used to define the measures ($\alpha,\beta$ and $\Theta,\phi$) are needed. Also, the split between completely filled, partially filled and unfilled orders and order lines are needed to establish the relationships. This is in contrast to the 'complete fill' measures, where the split between filled and unfilled orders is not required.

### 7.4.3 Relationships between 'partial fill' and 'complete fill' measures

The relationships between the equivalent 'complete fill' and 'partial fill' service level measures are given below :

$$\frac{L_c}{L_p} = \frac{r_f}{r_{cf} + \alpha r_{pf}} \; [\, 1 + (\alpha+\beta-1)\, r_{pf} \,]$$

$$\frac{O_c}{O_p} = \frac{s_f}{s_{cf} + \Theta s_{pf}} \; [\, 1 + (\Theta+\phi-1)\, s_{pf} \,]$$

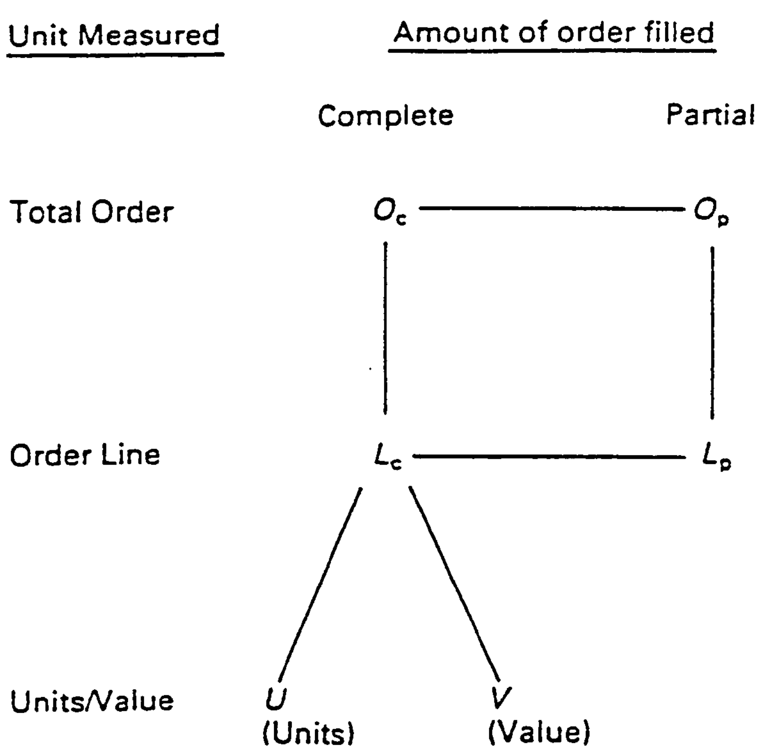
where $r_f = l_f / l_d$ and $s_f = o_f / o_d$.

If $\alpha+\beta = 1$ and $\Theta+\phi = 1$, then the expressions simplify as before. The proofs are given in Appendix 7.1. Data requirements are again more onerous than for relationships between 'complete fill' measures.

### 7.4.4 A network of relationships

A network of relationships between the measures has now been established, as summarised in Figure 7.1 :

Figure 7.1
Network of relationships between service-level measures




By using the relevant relationship-links in the diagram, it is possible to find a relationship between any two of the six measures.

## 7.5 Implementation issues

If the service measures differ between depots but the categorisation rules are the same, then the relationships may be applied for each stock category. Difficulties may arise if the categorisation rules are inconsistent between depots.

To assess the effects of centralisation for one category of stock, it would be necessary to analyse all depots according to the same categorisation system. However, it may not be possible to split data from all depots by the same categories since there may have been no need for this capability in the past. Similarly, to assess the effects of a change in service measure when the categorisation rules also change, it would be necessary to analyse old data according the new categorisation system. Again, this may not be possible because of the lack of need for such analysis previously.

In the case of inconsistent categorisation rules, it would be unwise to assume that the ratios used in the relationships for the old categories of stock will also apply to the new set of categories. Sensitivity analyses are required to show the differences between ratios for different stock categories. If the ratios are found to be insensitive to stock boundaries, then approximate relationships may be obtained; if the ratios are sensitive, then any such relationships would be unreliable.

## 7.6 Illustrative examples

The relationship between share of units filled from stock (U) and share of order lines filled completely from stock ($L_c$) has been shown to be :

$$U = 1 - (1 - L_c) \frac{u_u}{l_u} \frac{l_d}{u_d}$$

The key ratio in this equation is $u_u \, l_d / l_u \, u_d$ . For any given value of $L_c$, substitution of this ratio gives the unit-fill service measure (U) immediately. Using a hypothetical example, suppose that $L_c = 0.90$, and $u_u \, l_d / l_u \, u_d = 0.6$. Then,

$$U = 1 - (0.1 * 0.6) = 0.94.$$

This simple example, showing a difference of 4% between the two measures, highlights the significant differences that may arise in practice.

The next example is based on results from a study conducted by the author (Boylan (1986)) for Unipart Group Ltd. At the time of the study, the company served three clients using the same service measure ($L_c$) for each, with a common target of 92%. The company was considering changing over to the unit-fill measure, U. The ratio $u_u \, l_d / l_u \, u_d$ had been estimated for the three clients (1,2 and 3) and for three stock-movement categories (A,B and C). The ratios are given in Table 7.2 :

TABLE 7.2

Key ratio ($u_d l_u / l_u u_d$) for each client group and movement class

| | Stock - Movement Category | | |
|---|---|---|---|
| Client | A | B | C |
| 1 | 0.74 | 0.82 | 0.82 |
| 2 | 0.93 | 0.93 | 0.79 |
| 3 | 0.56 | 0.63 | 0.50 |

Sensitivity analyses for movement categories and clients were conducted, assuming that $L_c$ = 0.92 and using the ranges of the ratio $u_u\ l_d\ /\ l_u\ u_d$. The results of these analyses are presented in Tables 7.3 and 7.4 and show the potential effect of inconsistencies of movement categories and client groups. Such inconsistencies may occur because of differences in categorisation between decentralised depots or changes to categorisation after centralisation.

TABLE 7.3

Sensitivity to movement category for each client group

| Client | Max Ratio | Min Ratio | U (max ratio) | U (min ratio) | Difference in U |
|---|---|---|---|---|---|
| 1 | 0.82 | 0.74 | 93.44% | 94.08% | 0.64% |
| 2 | 0.93 | 0.79 | 92.56% | 93.68% | 1.12% |
| 3 | 0.63 | 0.50 | 94.96% | 96.00% | 1.04% |

TABLE 7.4
Sensitivity to client group for each movement category

| Movement Category | Max Ratio | Min Ratio | U (max ratio) | U (min ratio) | Difference in U |
|---|---|---|---|---|---|
| A | 0.93 | 0.56 | 92.56% | 95.52% | 2.96% |
| B | 0.93 | 0.63 | 92.56% | 94.96% | 2.40% |
| C | 0.82 | 0.50 | 93.44% | 96.00% | 2.56% |

The final column of Table 7.3 shows that, if accuracy to within 1% is sufficient, then approximate relationships may be used even if the boundaries of the movement classes change. However, suppose there were to be a change in the definition of 'client' (eg some customers move from one client group to another). In this case, Table 7.4 shows that it would be possible to be accurate to 3%. This is of limited value since safety stocks are highly sensitive to small changes in unit-fill rate targets when the targets exceed 90%.

This example demonstrates that if relationships between service level measures are required by category of stock, and if the categories are subject to inconsistencies, then a sensitivity analysis is required to identify the potential errors which may result.

## 7.7 Conclusions

Relationships between six of the main aggregate service level measures have been obtained. The application of these results to inventory centralisation has been examined. It should be noted that there are two further application areas : comparison of service between stock-holding depots and estimation of new service levels when the service measure is changed. These applications areas are discussed in the paper by Boylan and Johnston (1994), on which this chapter is based.

The relationships may be applied to inventory as a whole or to specific categories of inventory. However, if there are inconsistencies in stock categorisation, then it would be unwise to use the relationships unless it is possible to perform a sensitivity analysis.

# PART III

# DEMAND

# DISTRIBUTION FAMILIES

## Summary of Part III

In this part of the thesis, demand distributions are considered from theoretical and empirical perspectives. Two families of distributions, the compound Poisson and compound binomial, are also examined. The aim of part III of the thesis is to identify a suitable family of distributions which can be used as the basis for demand variance laws, the subject of part IV.

In chapter 8, three distributions, the log-zero-Poisson (lzP), the negative binomial (NBD) and the gamma distribution are assessed in terms of the underlying mechanisms which may justify their use as demand models. It is found that the NBD and lzP may be justified on the basis of Poisson and Bernoulli streams of orders (demand incidence). Using results from insurance mathematics, it is shown that the compound Poisson family of distributions - of which the NBD is a special case - can model a situation in which the mean order-rate is varying and the size of orders is also variable.

In chapter 9, alternative models of inter-purchase times are examined. The geometric and exponential distributions, which correspond to the compound binomial and compound Poisson families of demand distributions, are discussed. The theoretical and empirical evidence supporting the Erlang-2 inter-purchase time distribution is reviewed. The inverse Gaussian is also examined, but the empirical support for this distribution is shown to depend on a comparison with another distribution whose parameters are inappropriately estimated.

In chapter 10, the bias and efficiency properties for various estimation methods of demand and demand incidence distributions are examined. In some cases, it is found that use of the wrong estimation method may result in strong bias or very inefficient estimates.

In chapter 11, demand incidence and demand distributions are tested for their goodness-of-fit on a sample of 230 SKUs from Pillar Engineering Supplies Ltd. Strong support for Poisson demand incidence and NBD demand (using maximum likelihood estimation) is found. This backing for the compound Poisson family of distributions is the main conclusion of part III of the thesis.

# CHAPTER 8

## *Demand Distribution Models*

### 8.1 Introduction

Many probability distributions have been used to represent demand. In this chapter, three distributions will be reviewed: log-zero-Poisson (lzP), negative binomial (NBD) and gamma. Other distributions have been suggested in the literature, but most have been justified by abstract arguments alone and have little or no empirical support. Examples of such distributions include the logarithmic Poisson gamma (Nahmias and Demmy (1982)), the Hermite (Bagchi, Hayya and Ord (1983)) and the truncated normal-gamma mixture (Ord and Bagchi (1983)). The three distributions chosen have been tested empirically in previous research projects and shown to give reasonable fits to the data. Goodness of fit (empirical evidence) is not the only criterion by which demand distributions can be assessed. It is proposed that three criteria should be employed:

1. *A priori* grounds for modelling demand.
2. The flexibility of the distribution to represent different types of demand.
3. Empirical evidence.

To satisfy the first criterion, a distribution must be explainable in terms of an underlying mechanism. If a distribution may be justified on these grounds, then it will be said to *model* demand. If either or both of the other criteria are met, but the first is not, the distribution will be said to *represent* demand.

The second criterion was suggested by Katti (1966) in the context of biometric modelling. Katti proposed that distributions should be assessed by their 'flexibility', by which was meant the scope of skewness and kurtosis attainable by the distribution. Katti and Rao (1970) extended the concept of 'flexibility' to include the range of values attainable by the variance to mean ratio. Kwan (1991) applied this extension of the concept to demand modelling, arguing that if a distribution is to be used to represent demand for a broad range of SKUs, then the widest range of variance to mean ratios is required.

Published empirical evidence in support of the three distributions is reviewed in this chapter. A detailed analysis of data from Pillar Engineering Supplies Ltd will be conducted in chapter 11.

It is shown that the lzP distribution is a member of the compound binomial family of distributions and that the NBD is a member of the compound Poisson family of distributions. Linkages between these two families are highlighted, since the compound binomial is the discrete analogue of the compound Poisson family. The behaviour of the compound Poisson distribution under non-stationary demand is examined, as a precursor to the analysis of demand variance in chapter 12.

## 8.2 The log-zero-Poisson distribution

### 8.2.1 Definition of the log-zero-Poisson distribution

The density function of the log-zero-Poisson has three parameters ($\tau$, $\theta$ and p) and may be defined by the following recursive formulae :

$$P(0) = \theta$$

$$P(i+1) = \frac{\left(\dfrac{\tau p(1-\theta)\,\pi_i(\tau)}{\ln(q - pe^{-\tau})}\right) + p\,\Sigma_{j=1}^{i}\, j\,P(j)\,\pi_{i+1-j}(\tau)}{(i+1)\,(q - pe^{-\tau})}$$

where $q = 1+p$
$\pi_i(\tau)$ is the Poisson probability of i given a mean of $\tau$.

Kwan (1991) comments that the parameters $\tau$ and p do not have natural interpretations.

### 8.2.2 The flexibility of the log-zero-Poisson distribution

Kwan's main argument to justify the use of the lzP rests on its flexibility: it can represent demand patterns where the variance is greater than the mean, equal to the mean or less than the mean. The NBD, on the other hand, is restricted to demand patterns where the variance is greater than the mean, and the Poisson is restricted to those cases where the variance and mean are equal. The flexibility of the lzP in terms of the variance to mean ratio was demonstrated by Katti and Rao (1970) who showed that the variance is less than the mean, if and only if :

$$1 + p < \frac{(1 - \theta)\,p}{\log\,(q - p\,e^{-\tau})}$$

Since the above inequality may be satisfied or not according to the values of the parameters, the lzP distribution may be used in situations in which the variance is less than, equal to or greater than, the mean demand.

Moreover, Katti (1966) showed that the lzP distribution has a wider range of skewness and kurtosis values than a broad range of 'contagious' distributions, including the Neyman Type A, the negative binomial and the Poisson binomial.

### 8.2.3 Empirical evidence to support the log-zero-Poisson distribution

Empirical evidence to support the lzP distribution of demand over a fixed periods of time has been provided by Kwan (1991). She analysed two data sets, one consisting of 29 spare parts for a vehicle manufacturing company and the other consisting of 56 spare parts for a steel manufacturing company. Seven of the first set of parts and three of the second set of parts had demand distributions with variances less than the means. Since the NBD does not allow the variance to be less than the mean, it may be expected that the lzP would outperform the NBD.

The analysis showed a higher proportion of SKUs to be well-fitted by the lzP (82%) than by the NBD (75%), where 'well-fitted' means that the distribution was not rejected by a chi-square test with a significance level of 5%. If the NBD is adjusted, by inflating the sample variance estimate to up to 5% more than the sample mean (for those SKUs with sample variance less than the sample mean), then the results change. The adjusted NBD fitted the demands of 91% of SKUs. However, as Kwan notes, it is difficult to find a theoretical justification for such an adjustment.

### 8.2.4 A biometric analogy

Kwan gave no *a priori* justification for the lzP as a model for the distribution of demand. The lack of such a theoretical justification limits the insights which may be gained into the components of demand variance. In this sub-section and the next, two possible *a priori* models are advanced and their limitations are assessed.

Katti and Rao (1970) demonstrated the relevance of the lzP to modelling the number of larvae on a stalk. The model was justified in three stages, as follows:

1. Define $X_1 = 1$ if a moth settles on a stalk
   $= 0$ otherwise.

   The density function of $\Sigma X_1$ may be modelled by a binomial distribution.

2. Let $X_2$ represent the number of egg masses the moth lays on the stalk.

   The density function of $X_2$ per stalk may be represented by a logarithmic distribution.

3. Let $X_3$ be the number of larvae hatched from an egg mass.

   The density function of $X_3$ per egg mass may be modelled by the Poisson distribution.

The first distribution is termed a *model* since its justification rests on the assumed Bernoulli process. The second distribution is posited because its shape is similar to observed distributions of egg masses per stalk. It is not a model, in the sense outlined in the beginning of this chapter, since no underlying mechanism has been presented to justify its use. Hence, the logarithmic distribution is a *representation* of the number of egg masses per stalk. The third distribution is justified by Katti and Rao by assuming a Bernoulli process on a large number of eggs with a low probability of success. Therefore, the Poisson distribution is a *model* of the number of larvae per egg mass. On the basis of these three distributions, Katti and Rao derived the distribution of the total number of larvae per stalk to be a log-zero Poisson distribution.

This model may be translated to the context of demand modelling to give the *first a priori model*. The analogy required for this translation is shown in Table 8.1 :

TABLE 8.1
First analogy between the hatching of larvae and the ordering of items

| | |
|---|---|
| Moth settles on stalk $j$ | Customer places order in time period $j$ |
| Moth lays number of egg masses | Order consists of order-lines for several SKUs |
| Larvae hatch from egg mass | Order-line for each SKU is for several items |

In the above analogy, each of the three levels analysed in the previous chapter - orders, order-lines and items - are included. Orders are represented by the arrival of a moth on stalk $j$ (or, the arrival of a customer in time period $j$); order-lines are analogous to egg masses and items are analogous to larvae.

There is a serious weakness in the above analogy: the lzP is a model of the distribution of the total number of items ordered. Hence, the lzP models the number of items ordered for a collection of SKUs over time. It is not a model of demand for an individual SKU over time, as required in inventory applications.

### 8.2.5 The log-zero-Poisson as a 'compound binomial' distribution

To overcome the weakness of the first model, discussed above, a *second a priori model* is proposed. The analogy required is shown in Table 8.2 :

TABLE 8.2
Second analogy between the hatching of larvae and the ordering of items

| | |
|---|---|
| Moth settles on stalk $j$ | Customer places order for a particular SKU in time period $j$ (customer places an order-line) |
| Moth lays number of egg masses<br>Larvae hatch from egg mass | Order-line is for several items |

In this analogy, orders are no longer included. Items are analogous to larvae, as before, and the placing of an order-line is represented by the arrival of a moth. The lzP is a model of the distribution of the total number of items ordered for a particular SKU over time.

The number of items per order-line is represented by the logarithmic-Poisson distribution. It should be noted that this distribution is distinct from the Poisson-logarithmic. Jones and Mollison (1948) used the Poisson-logarithmic distribution, modelling the number of colonies of soil micro-organisms per field as Poisson and the number of bacteria per colony as logarithmic. Quenouille (1949) showed that a Poisson-logarithmic process yields a negative binomial distribution. No such result has been found for the logarithmic-Poisson distribution.

In this second model, the binomial distribution of the number of customer arrivals may be justified by the assumed Bernoulli process. However, the Poisson-logarithmic representation of items per order does not seem natural: there is no obvious reason why this distribution should be used rather than any other. If the assumption of the Poisson-logarithmic is dropped, then there is a binomial distribution of the total number of order-lines (for a particular SKU) and a general distribution, not specified, of the number of items per order-line. Such a distribution will be called a 'compound binomial' distribution.

To summarise the argument of this section, the lzP is supported by its wide range of variance to mean ratio, skewness and kurtosis, thus ensuring great 'flexibility' for representing demand. Kwan (1991) has presented some empirical evidence, based on 85 SKUs, showing over 80% of the SKUs being well fitted by an lzP distribution. Attempts to provide an *a priori* justification for the distribution have been only partially successful. The binomial distribution of the number of orders may be justified by a Bernoulli process; there is no such model to support the logarithmic-Poisson distribution of the number of items per order. A strong *a priori* case may be made for the family of 'compound binomial' distributions but not for the lzP, a member of this family.

## 8.3 Negative binomial distribution

### 8.3.1 Representation of demand by the negative binomial distribution

Taylor (1961) argued for the use of the negative binomial distribution (NBD) to model demand. The density function is of the form :

$$f(x) = \frac{(n-1)!}{(r-1)!\,(n-r)!}\, p^r q^{n-r}$$

where $q = 1 - p$.

Taylor conceded that there were practical difficulties in using the NBD : "*tables of the negative binomial distribution are not easy to find (if, indeed, they exist at all)*". To overcome this problem, Prichard and Eagle (1965) gave a method for calculating probabilities based on recursive formulae for the NBD :

$$P(0) = \left(\frac{1}{z}\right)^{\mu/(z-1)}$$

$$P(x) = \frac{[\mu/(z-1)]+x-1}{x}\ \frac{z-1}{z}\ P(x-1) \qquad x=1,2,...$$

where $\mu$ is the mean and z is the variance to mean ratio. The proof of the formulae was omitted by Prichard and Eagle ; this omission was rectified by Kwan.

Kwan (1991) compared the NBD with four other distributions (Poisson, Hermite, Laplace and log-zero-Poisson) and concluded that the best distribution for the two data sets is the NBD and that the log-zero-Poisson is the next best. However, this result depends upon an adjustment to the sample variance for some spare parts when using the NBD model, as mentioned in sub-section 8.2.3.

## 8.3.2 *A priori* justification of the NBD demand model

The negative binomial distribution (NBD), suggested by many authors for demand modelling, has been shown to have a number of theoretical bases. Ehrenberg (1959, 1972) demonstrated how the distribution of demand across customers may be derived from a 'mixing over customers model' :

* **Mixing over Customers Model.** The demand from the *i*th customer is assumed to be Poisson distributed about its stationary mean $\mu_i$. The means are assumed to be gamma distributed across the population of customers. Demand is then NBD distributed across customers.

Wright (1991) suggested an almost identical model for the distribution of demand across SKUs :

* **Mixing across SKUs Model.** The demand for the *i*th SKU is assumed to be Poisson distributed about its stationary mean $\mu_i$. The means are assumed to be gamma distributed across the population of SKUs. Demand is then NBD distributed across SKUs.

Taylor (1961) showed that the NBD could be used for modelling demand over a variable lead-time. Kwan's empirical evidence showed support for the NBD over fixed lead-time. There are two quite different models which can be used to justify the NBD distribution of demand over a fixed lead-time, namely 'compounding' and 'mixing over time' :

* **Compounding Model.** Order occasions are assumed to be Poisson distributed. When an order is placed, the order size is not fixed but follows a logarithmic distribution. The total demand is then NBD distributed over time.

* **Mixing over Time Model.** Order occasions are assumed to be Poisson distributed about a mean $\mu$. The mean varies over time and it is assumed that $\mu$ follows a gamma distribution. Demand is then NBD distributed over time.

Sherbrooke (1992) argued for the necessity of considering movement of the mean demand over time. He commented that *"From earlier demand-prediction studies we know that, over very short periods, demands for most items tends to be Poisson, and that the variance to mean ratio for a particular item tends to increase as the period of observation increases"*. This finding is consistent with a Poisson process with the mean value varying over time, but not with the Poisson distribution having a stationary mean.

The NBD has strong theoretical support, then, since it may be justified by two distinct models. However, commenting on this matter in the context of 'mixing across SKUs', Wright (1991) sounded a note of caution: *"Whether the same models can be applied to those situations where both compounding and mixing occur simultaneously is not known"*. This issue will be discussed further in section 8.6.

## 8.4 The gamma distribution

### 8.4.1 Representation of demand by a gamma distribution

Burgin and Wild (1967) argued for the use of the gamma distribution to represent demand. The density function is of the form :

$$f(x) = \frac{\alpha^k x^{k-1} e^{-\alpha x}}{\Gamma(k)} \qquad \text{for } x \geq 0,$$

where $k > 0$ and is known as the 'shape' parameter; $\alpha > 0$ and is known as the 'scale' parameter; and $\Gamma$ denotes the complete gamma integral.

Burgin (1975) listed five characteristics of the gamma distribution which make it suitable for representing demand in an inventory control context. Three of the characteristics make the distribution attractive from a purely practical viewpoint :

* The gamma distribution is easily convolvable.
* The probability integral of the distribution is well tabulated.
* The distribution is generally mathematically tractable in its inventory control applications (see Burgin and Wild (1967), Johnston (1973a, 1973b) and Burgin (1975)).

Two further characteristics of the gamma distribution which make it attractive for representing demand are as follows :

* The distribution ranges from a monotonic decreasing function, through unimodal distributions skewed to the right to normal type distributions.
* It is defined for non-negative values only.

Burgin comments that the concept of negative demand is *'unrealistic'* for inventory control work, though he notes that customer returns is a possible interpretation.

Empirical evidence to support the use of the gamma distribution for demand modelling has been reported in the literature. Johnston and Milne (1972) commented that, *"[Burgin and Wild] indicated that a gamma distribution fitted on the first two moments represented actual demand patterns better than the corresponding normal distribution. This view is supported by one of the present authors who has investigated the apparent demand distribution of many items in different industries and found a gamma distribution far superior"*. Johnston (1980) analysed 784 SKUs and noted that this led to *"the adoption of the gamma family in lieu of any better alternative"*.

8.4.2 *A priori* justification of the gamma demand model

No *a priori* justification of the gamma distribution as a demand model is given by Burgin. A derivation from first principles was proposed by Goodhardt and Chatfield (1973) for the gamma as a distribution of mean demand values across customers. However, this derivation is not relevant to the distribution of demand over time.

If it is assumed that demand is discrete, then the gamma can be only an approximation to the distribution of demand. Feller (1957) noted that the negative binomial distribution (NBD) is the discrete analogue of the gamma. The theoretical arguments underpinning the NBD as a demand model were reviewed in sub-section 8.3.2. Thus, although not having *a priori* support, the gamma is related to a distribution which has its own theoretical justification.

## 8.5 The compound Poisson distribution for stationary demand

### 8.5.1 *A priori* justification of the compound Poisson demand model

In the justification of the 'compound binomial' distribution of sub-section 8.2.5, there were three implicit assumptions:

1. Demand is discrete.
2. Demand in time period $j$ is independent of previous periods.
3. Demand is stationary.

In most inventory systems, the first assumption holds. It is not difficult to think of exceptional products (eg sand) but packaging and pricing conventions mean that goods such as sand are usually sold in standard measures. The final assumption is more restrictive. It is possible that the mean demand level will vary stochastically through time. This more general situation is examined in section 8.6.

The second assumption may be modified. In the development of the 'compound binomial' as a demand distribution, time periods are treated as discrete entities. If, instead, time is treated as a continuous variable, then the second assumption may be adapted and generalised to time-intervals which are not necessarily of equal length :

* Demands in any two non-overlapping time intervals (of any length) are independent.

There are some situations in which this assumption is unlikely to hold. Relaxations of the assumption are discussed in chapter 9.

Feller (1957) showed that any stationary discrete variable satisfying the modified second assumption must be a member of the class of compound Poisson distributions. This family includes :

* Poisson, for single item orders.
* Clumped Poisson (Ritchie and Kingsman (1985)), for multiple item orders for the same SKU of a fixed 'clump size'.
* Negative binomial (see, for example, Sherbrooke (1968)), for order sizes following a logarithmic distribution.

A summary of the analogies (shown by arrows) between demand distributions based on discrete and continuous processes is shown in Table 8.3:

TABLE 8.3
Analogies between discrete and continuous independent demand processes

| | **Bernoulli Process (Discrete)** | | **Poisson Process (Continuous)** |
|---|---|---|---|
| **Single Item Purchases** | BINOMIAL | ⟷ | POISSON |
| **Multiple Item Purchases** | COMPOUND BINOMIAL | ⟷ | COMPOUND POISSON |
| | including : | | including : |
| | LOG-ZERO POISSON | | NEGATIVE BINOMIAL |

There are further properties of the compound Poisson distribution which strengthen its theoretical justification. These properties will be reviewed in the next section.

### 8.5.2 Representation of demand by the compound Poisson distribution

All compound Poisson distributions have a variance greater than or equal to the mean. This limits the flexibility of the distribution, as noted by Kwan (1991) who examined some SKUs with the property that the sample variance of demand is less than the sample mean.

Kwan found the Poisson to perform poorly as a demand distribution, fitting only 40% of a sample of 85 SKUs. Sherbrooke (1992) commented that although the American Air Force uses Poisson demand in many stock models, the distribution is based on unrealistic assumptions and tends to mis-allocate inventory investment.

In a case study, Ritchie and Kingsman (1985) observed that some items follow a Poisson pattern but demand occurs in 'clumps'. Demand incidence, rather than demand itself, was found to be Poisson distributed. The clump sizes may be determined by pack-sizes or because demand naturally occurs in clumps. The authors give an example of pistons for four cylinder cars, which would usually be sold in fours, since all pistons are likely to be replaced in an engine overhaul.

Ritchie and Kingsman admit that the demand distribution is complicated by orders for single items of goods normally sold in clumps. This situation may require more complex Poisson models than the clumped Poisson with fixed clump-size.

## 8.6 The compound Poisson distribution for non-stationary demand

In the previous section, two alternative models (mixing over time and compounding) have been found to lead to a negative binomial distribution of demand over time. Their joint effect may be evaluated by using a result proved by Bühlmann and Buzzi (1971), motivated by research in the field of insurance mathematics. The result is given below:

*Bühlmann and Buzzi's Theorem* A mixed compound Poisson process may be represented as a (non-mixed) compound Poisson variable if, and only if, the mixing function is infinitely divisible.

There are four assumptions required for Bühlmann and Buzzi's result to hold. Firstly, a random (Poisson) process is assumed. Herniter (1971) suggested that there is likely to be a 'dead' period immediately following a consumer's purchase. In this case, demand in one period is not necessarily independent of demand in previous periods and inter-arrival times are more regular than might be predicted by the negative exponential distribution. Herniter suggested the Erlang-2 as an alternative. This distribution will be reviewed in the next chapter.

The second assumption is that the distribution of mean demands over time follows an infinitely divisible distribution. Although it is not certain that this assumption will be valid, it is not very restrictive. For example, the following distributions have been shown to be infinitely divisible : gamma and generalised gamma (gamma$^{\alpha}$, for $\alpha \geq 1$) (Bondesson (1978)), lognormal (Thorin (1977a)) and Pareto (Thorin (1977b)).

The third assumption is that the mean is finite. This was needed for Bühlmann and Buzzi's proof of the necessity of infinite divisibility for the mixing distribution. However, deviations from this assumption are of no practical concern.

The final assumption is that the order-size distribution is independent of mean incidence of demand. This was questioned by Paull (1978) who claimed that "*most data do not support this assumption of independence*". However, Paull produced only one example of an SKU, facial tissue, to justify this claim. In chapter 12, the assumption of independence will be shown to be critical when analysing the relationship between the variance and mean of demand over a range of SKUs.

In this brief review of assumptions for the Bühlmann and Buzzi theorem, it has been shown that the second and third assumptions are not problematic. Empirical evidence from one study has cast doubt on the fourth assumption; hence, this assumption should be tested in practice. Empirical evidence from a number of studies have given support to the Erlang-2 inter-purchase distribution. This evidence will be reviewed in more detail in chapter 9 and the implications of the Erlang-2 inter-purchase distribution on variance laws will be assessed in chapter 12.

## 8.7 Conclusions

In this chapter, it has been proposed that demand distributions should be assessed against three criteria: *a priori* justification; flexibility to model different types of demand; and empirical evidence. The underlying mechanisms behind three distributions, which have received empirical support, have been investigated. This analysis has enabled the inter-relationships between the distributions, and the families to which they belong, to be brought into focus.

A theoretical model has been suggested for the log-zero Poisson distribution, based on a Bernoulli process of order-lines and a logarithmic-Poisson distribution of the number of items per order-line. However, no justification has been found for the use of the logarithmic-Poisson distribution in this situation.

Two derivations of the negative binomial distribution (NBD) as a demand model over a fixed lead-time were reviewed: the mixing model and the compounding model. Although both models are plausible, it was noted that no research has been published in the inventory field on their joint effect. It was shown, using a result from insurance mathematics, that the joint effect of mixing and compounding, under certain conditions, is a compound Poisson distribution. Since there is empirical evidence (Sherbrooke (1992)) to support the 'mixing' of the mean demand over time, this result strengthens the case for the 'compound Poisson' distribution. The assumption of the random arrival of orders will be investigated in the next chapter.

The gamma distribution has the backing of empirical evidence (Johnston and Milne (1972), Johnston (1980)) but no theoretical justification has been advanced for it. If, however, it is recalled that the NBD is the discrete analogue of the gamma, then it may be observed that the gamma is related to a distribution with theoretical support.

In the theoretical model for the log-zero-Poisson distribution, if a general order-size distribution replaces the logarithmic-Poisson, then a 'compound binomial' distribution is obtained. The 'compound binomial' distribution treats time as a discrete variable. If this assumption is replaced by the more natural treatment of time as a continuous variable, then a 'compound Poisson' distribution is obtained. Specific examples of members of the compound Poisson family of distributions include the clumped Poisson and the negative binomial. Empirical support for these distributions has been provided by Ritchie and Kingsman (1985) and Kwan (1991).

# CHAPTER 9

## *Inter-Purchase Time Distribution Models*

## 9.1 Introduction

In the previous chapter, an *a priori* case was developed for the compound Poisson demand distribution. The justification depends on the assumption that demands in any two non-overlapping time intervals are independent and, consequently, inter-demand times may be modelled by a negative exponential distribution. In this chapter, three alternative distributions will be reviewed, namely the geometric, the Erlang-2 and the inverse Gaussian. As in the previous chapter, the distributions will be assessed according to their theoretical justification as well as the empirical evidence to support their use. The implications of Erlang-2 inter-demand on distributions of demand incidence and demand will be examined.

## 9.2 Criticisms of the assumption of 'random' demand from individual consumers

### 9.2.1 Purchase habit effect

Kahn and Morrison (1989) argue that, even if a consumer's need for a product occurs at negative exponential intervals, the purchase may be delayed to the next regularly scheduled shopping trip. In these circumstances, times between 'need' are continuous but times between 'purchase' are discrete. The authors suggest using the geometric distribution - the discrete analogue to the negative exponential - in this situation.

### 9.2.2 Consumption effect

Herniter (1971) noted that an individual's inter-demand pattern is random only if the probability of purchasing in any time interval of given length, is constant. Herniter argued that, for many products, as the time since the last purchase increases the probability of needing to repurchase the product will increase. This is particularly relevant to 'consumable' products. Wheat and Morrison (1990) made the same point: *"A person who has just bought a one-pound can of coffee is unlikely to buy another tomorrow"*. Of course, if the product is replaced on failure (eg light bulb, car tyre), then the random assumption is more natural. To overcome the problem of non-randomness of inter-demand times for consumable products, Herniter proposed using the Erlang-2 distribution. Chatfield and Goodhardt (1973) supported Herniter's case, providing empirical evidence of an Erlang shape parameter of two.

### 9.2.3 Variability in shape of inter-demand distributions

Banerjee and Bhattacharyya (1976) commented on the wide variability in the shape parameter around the average of two observed by Chatfield and Goodhardt. This led the authors to propose the two parameter inverse Gaussian distribution to represent inter-demand times, since it can accommodate a rich variety of shapes.

## 9.3 The geometric inter-purchase distribution

### 9.3.1 *A priori* case for the geometric distribution

Kahn and Morrison (1989) justify the geometric distribution for modelling the inter-purchase time for individual consumers. The distribution arises if, on each shopping occasion, a given product is bought with a constant probability p, regardless of the number of past occasions since the product was last bought. Although Kahn and Morrison considered individual consumers, their argument also applies to retailers or wholesalers, who have a regular day of the week when stock orders are placed on a regional or national distribution centre.

The geometric distribution assumes a 'memory-less' process: the probability of a purchase occurring is independent of the time since the last purchase. Suppose that, in addition, it is assumed that the purchase occasions of individual consumers are mutually independent. If there are N consumers with a particular regular shopping day, then the probability of any consumer making a purchase on the day in question is given by :

$$P(\text{purchase occurrence}) = 1 - \prod_{j=1}^{N} (1 - p_j)$$

where $p_j$ is the probability of a purchase by the *j*th consumer ($j = 1,\ldots,N$).

In the case of individual consumers, it is unlikely that the above probability will be the same for each shopping day. The number of consumers with Saturday as their regular shopping day may exceed the number with Thursday as a regular day, for example. In an industrial setting, however, it is possible that a more even distribution of stock-ordering days may be achieved, in agreement with the retailers or wholesalers being served.

### 9.3.2 Empirical evidence on the geometric distribution

There is little evidence available on the goodness-of-fit of the geometric distribution to empirical inter-purchase time distributions. Kwan (1991) found the geometric to fit only 20% of her sample of SKUs, using a chi-square test at a significance level of 5%. This compared unfavourably to the negative exponential distribution, which fitted 42% of the SKUs.

Since little empirical work has been done in this area, an examination of the goodness-of-fit of the geometric distribution will be presented in chapter 11, using a sample of SKUs from an engineering supplies company.

## 9.4 The Erlang-2 inter-purchase distribution

### 9.4.1 *A priori* case for the Erlang-n distribution

Chatfield and Goodhardt (1973) comment that a simple theoretical model gives rise to the Erlang-n distribution. Consider a 'censored' Poisson process in which every *n*th event is 'marked'. Then the distribution of the time between the 'marked' events is Erlang with shape parameter n. Thus, if a consumer buys n items of a particular product on each shopping occasion, and if the time to requiring the next item is Erlang-1 distributed (negative exponential), then the time to the next demand on a stockist will be Erlang-n distributed.

It should be noted that the above argument depends on the same quantity being purchased on each shopping occasion and, hence, a regular 'marking' process occurring. If the 'marking' is random, then the time between demands is no longer Erlang-n. The 'colouring theorem' of marked Poisson processes states that :

*If the points of a Poisson process (mean $\mu$) are coloured randomly with k colours, the probability that a point receiving the ith colour being $p_i$ and the colours of different points being independent (of one another and of the positions of the points), then the set of points with the ith colour are independent Poisson processes with means $\mu_i = p_i\,\mu$ $(i=1,2,...,k)$.*

A discussion and proof of the colouring theorem is given by Kingman (1993).

Thus, if a Poisson process is marked randomly, the marked Poisson processes are also random and the inter-demand times are negative exponential. This discussion shows that the assumption of consistent purchase quantities (leading to a regular 'marking') is a critical assumption in the derivation of an Erlang-n inter-purchase distribution.

Gupta (1991) examined the effect of five covariates on inter-purchase times in stochastic models based on negative exponential and Erlang-2 inter-purchase times. The five covariates are listed below :

* household weekly inventory
* regular price
* promotional price cut
* feature and display (zero-one variable)
* feature or display (zero-one variable).

It was found that 'household weekly inventory' had the greatest effect on inter-purchase time. This empirical finding lends support to the a priori case for the Erlang-n distribution, since the conditions for a 'marking' process were found to be present.

The arguments summarised above support the use of the Erlang-2 as an inter-purchase distribution for *individual* consumers or households. No arguments have been presented to support the Erlang-2 for inter-demand times on a stockist (ie when *total* consumer demand is considered).

9.4.2 Empirical evidence based on consumer data

Herniter (1971) presented histograms of time between purchases for three product types: facial tissues, aluminium foil and laundry detergent. They each show unimodal distributions with non-zero modal values. Herniter reports that such shapes are typical for frequently purchased consumer products. Assuming an Erlang distribution, an analysis of the coefficient of variation of inter-purchase times for aluminium foil gave an estimate of 1.9 for the shape parameter, rounded to 2 since an integer value is required for the Erlang. Coefficients of variation were not reported for the other two product types.

Chatfield and Goodhardt (1973) reported an analysis of consumer data showing the average coefficient of variation of inter-purchase times of individual consumers to be between 0.7 and 0.8. This result is consistent with an Erlang-2 distribution which has a coefficient of variation of approximately 0.71 ($1/\sqrt{2}$ exactly).

Wheat (1987) summarised the statistical limitations of the coefficient of variation approach. Firstly, estimates of the mean and variance of inter-purchase times for each consumer are required. This requires a large number of inter-purchase times for each consumer. Thus, researchers often limit their analyses to consumers who have made at least ten purchases in the time period examined, biasing the results in favour of heavier users. Secondly, long purchase histories are often used to ensure an adequate base of consumers with at least ten purchases. However, the assumption that consumers' purchasing habits remain stationary becomes more questionable as the time period lengthens.

Jeuland, Bass and Wright (1980) examined a sample of French consumer data for cooking oil over the entire two year period 1971-72 using a maximum likelihood approach. They compared the fit of Erlang-1 (negative exponential), Erlang-2, Erlang-3, Erlang-4 and Erlang-5 distributions to inter-purchase times for 200 customers. The values of the likelihood function for each of the distributions were compared. The authors report that the Erlang-2 yielded a higher likelihood function than the negative exponential for 71.8% of customers. It also outperformed the Erlang-3, Erlang-4 and Erlang-5 (for 66.2 %, 72.3% and 74.4% of customers respectively). In this study, no attempt was made to find how many customers had shape parameters significantly different from unity.

Dunn, Reader and Wrigley (1983) examined data from a consumer panel in Cardiff for a 24-week period in 1982. The products covered were baked beans, toilet tissue and instant coffee. Dunn, Reader and Wrigley reported sample sizes of 168 customers for baked beans and 177 customers for toilet tissue; the sample size for instant coffee was not given. For each customer-product combination, assuming a gamma distribution of inter-purchase times, the shape parameter was estimated by maximum likelihood. Then the null hypothesis that the shape parameter is equal to one (negative exponential inter-purchase times) was tested for each customer-product combination. It was found that for almost all non-heavy buyers (ie consumers with fewer than 12 purchases of a product in 24 weeks), the null hypothesis was not rejected. However, for the heavy buyers, 57% of baked beans consumers and 62% of toilet tissue consumers had shape parameters significantly greater than one (significant at the 5% level). Detailed figures were not given for instant coffee, though the authors note that *"very similar results were obtained"*. The evidence from the three products indicates that the negative exponential cannot necessarily be used as a universal inter-purchase model.

Dunn, Reader and Wrigley discuss the major limitation of maximum likelihood estimation for this problem : the method introduces some bias in parameter estimation for the small samples observed for non-heavy buyers. However, they comment that the bias does not seem sufficiently large to change their conclusions for the lighter buyers. The samples are large enough for the heavy buyers to be hardly affected by bias at all.

Gupta (1988) analysed the purchase patterns of coffee of 100 households over the two-year period 1980-82 at Pittsfield, Massachusetts. The maximum likelihood estimate of the shape parameter was 2.03. Also, the Erlang-2 null hypothesis was not rejected for 92 out of 100 households, using the Kolmogorov-Smirnov test. Unfortunately, the significance level of this test was not given.

### 9.4.3 Empirical evidence based on inter-demand times on a stockist

Little empirical evidence is available, but Kwan (1991) found the Erlang-2 inter-demand distribution did not fit the demand histories of any of the 85 slow-moving spare parts. An empirical analysis on the demand histories of engineering supplies SKUs, including slow and fast-moving items, will be presented in chapter 11.

### 9.4.4 Condensed-Poisson demand incidence distribution

The distribution of the number of purchases arising from Erlang-2 inter-purchase times may be obtained by the following argument, summarised from Chatfield and Goodhardt (1973). The Erlang-2 distribution may be derived by considering a 'censored' Poisson process in which only every second event is recorded. If counting begins at a time independent of the process being counted, an odd number of events in the Poisson process, say 2r+1, is equally likely to give r or r+1 events in the censored process. Similarly, an even number of events in the Poisson process, say 2r, is certain to give r events in the censored process.

Hence, the probabilities are given by :

$$\text{Prob(no purchases)} = P_p(0) + \tfrac{1}{2} P_p(1)$$

$$\text{Prob( k purchases)} = \tfrac{1}{2} P_p(2k-1) + P_p(2k) + \tfrac{1}{2} P_p(2k+1)$$

where the second equation applies for $k = 1,2,3,...$ and $P_p$ refers to the probability of p events occurring in the uncensored process.

If the mean of the censored process is $\lambda$, then the mean of the uncensored process is $2\lambda$, and using the above relationships, the condensed Poisson distribution is given by:

$$p_0(\lambda t) = e^{-2\lambda t} + \lambda t\, e^{-2\lambda t}$$

$$p_k(\lambda t) = \frac{1}{2} e^{-2\lambda t} \frac{(2\lambda t)^{2k-1}}{(2k-1)!} + \frac{e^{-2\lambda t}(2\lambda t)^{2k}}{(2k)!} + \frac{1}{2} e^{-2\lambda t} \frac{(2\lambda t)^{2k+1}}{(2k+1)!} \quad \text{for } k \geq 1.$$

where t is the time period of concern.

Chatfield and Goodhardt (1973) named the above distribution 'condensed Poisson' since its variance is always less than its mean.

The compound Poisson distribution, discussed in the previous chapter, has a variance to mean ratio greater than or equal to one. Compound Poisson demand is based on Poisson demand incidence, hence resulting in the minimum variance to mean value of unity. The condensed-Poisson demand incidence distribution extends the potential variance to mean ratios to the range 0.5 to 1. For the condensed-Poisson, the ratio is :

$$\frac{\text{Variance}}{\text{Mean}} = \frac{1}{2} + \frac{e^{-2\lambda t}\sinh(2\lambda t)}{4\lambda}$$

It may be observed that as $\lambda$ tends to zero, the variance to mean ratio tends to one and, as $\lambda$ tends to infinity, the ratio tends to a half. This indicates that for small values of $\lambda$, the purchase distributions will be similar to Poisson, but for larger values of $\lambda$, there is a greater difference between the Poisson and condensed Poisson distributions, which should be discernible from real data. Such a difference was distinguished for 'heavy' buyers by Dunn, Reader and Wrigley (1983), whose empirical work was reviewed in the previous section.

### 9.4.5 Condensed negative binomial demand distribution

If the mean rate of demand incidence is not constant, but varies according to a gamma distribution, then demand incidence will follow a condensed negative binomial distribution (CNBD). Chatfield and Goodhardt showed that this distribution is given by the following recursive formulae :

$$\text{Prob(no purchases)} = P_N(0) + \tfrac{1}{2} P_N(1)$$

$$\text{Prob( k purchases)} = \tfrac{1}{2} P_N(2k-1) + P_N(2k) + \tfrac{1}{2} P_N(2k+1)$$

where the second equation applies for $k = 1,2,3,...$ and $P_N$ refers to the probability of p events occurring in an uncensored process following a negative binomial distribution (NBD). Although this distribution has been derived for demand incidence, it may also be used for demand as an alternative to the NBD, on the grounds of 'flexibility' since it has a wider range of variance to mean ratios than the NBD. Although the condensed negative binomial extends the lower end of the range of the variance to mean ratio to 0.5, it should be noted that it is not as flexible as the log-zero-Poisson which permits any ratio greater than zero.

## 9.5 The inverse Gaussian inter-purchase distribution

### 9.5.1 *A priori* case for the inverse Gaussian distribution

Banerjee and Bhattacharyya (1976) proposed a model based on inverse Gaussian inter-purchase times for individual households. The density of the inverse Gaussian (IG) distribution, with parameters $\psi$ and $\lambda$ is given by :

$$f(x \mid \psi,\lambda) \;=\; (\lambda/2\pi)^{1/2}\, x^{-3/2}\, \exp\{\, -\, \lambda x\, (\psi - 1/x)^2 / 2 \,\} \qquad x > 0,\ \lambda > 0,\ \psi > 0.$$

The mean and variance of the distribution are $\psi^{-1}$ and $\psi^{-3}\,\lambda^{-1}$ respectively. Hence, if the time between successive occurrences of an event follows an inverse Gaussian distribution, then the parameter $\psi$ may be interpreted as the average number of occurrences of the event per unit time.

Banerjee and Bhattacharyya justified the use of the IG distribution by recalling its application to problems of Brownian motion. Schrödinger (1915) showed that the time taken for a particle in Brownian motion to reach a barrier may be modelled by the inverse Gaussian distribution. (Schrödinger's original article is in German, but a summary of his paper may be found in English in the first chapter of Seshadri (1993) which also has reviews of some alternative derivations of the inverse Gaussian model for Brownian motion problems.) Banerjee and Bhattacharyya argue that, *"This interpretation of the IG as a first passage time distribution suggests its potential usefulness in modelling life times or usage times as an alternative to other conventional models such as the Weibull, gamma and lognormal".*

Although the above argument is attractive, it contains a significant weakness. It would be natural to assume that the 'passage time' is the time taken for stocks to fall below a critical level, or to reach zero, thus prompting another purchase. However, one would expect the 'drift' in stocks, for an individual household, to be always downwards. This contradicts the inverse Gaussian model, in which 'drifts' in Brownian motion may be away from a barrier as well as towards it.

### 9.5.2 Empirical evidence of household purchases

Banerjee and Bhattacharyya (1976) examined the inter-purchase patterns of toothpaste by just under 300 households in Chicago, USA, over a period from January 1958 to April 1963. The original sample was exactly 300 but 11 households were excluded since they reported "*unreasonably extreme values*", possibly due to lack of care in recording the data. The authors compared two distributions of the number of purchases per household :

1. **The 'compound' inverse Gaussian distribution.** This distribution is obtained by assuming that inter-purchase times are inverse Gaussian and the parameters $\psi$ and $\lambda$ vary across the households according to a bivariate density function.

   The term 'compound' is used by Banerjee and Bhattacharyya as the term 'mixed' was used in the previous chapter (although the authors are 'mixing' over customers, rather than 'mixing' over time). In this sub-section, 'compound' will be used in Banerjee and Bhattacharyya's sense, but quotation marks will be retained to highlight its different usage. In future sections, quotation marks will be dropped to indicate that the original meaning (orders for possibly more than one item of the same SKU) is intended.

2. **The negative binomial distribution.** This distribution is obtained by assuming that inter-purchase times are negative exponential, and the mean purchases are gamma distributed across households. This follows Ehrenberg's (1959) model of household purchases.

Banerjee and Bhattacharyya concluded that the 'compound' IG distribution was preferable to the NBD for two reasons, summarised below.

Firstly, the 'compound' IG provides a better fit to the data than the negative binomial distribution .

Secondly, the NBD is restricted in its scope of application. Ehrenberg recommended the use of the method of 'means and zeros' to estimate the shape parameter, k. The shape parameter is found by solving the following equation for k :

$$p_0(t) = \left(1 + \frac{m}{k}\right)^{-k}$$

where m is the observed mean number of purchases per household, and

$p_0(t)$ is the observed proportion of households making no purchases in time period t.

Banerjee and Bhattacharyya note that the above equation is solvable if and only if $m > -\log p_o(t)$. They comment that if $m < -\log p_0(t)$ then *"the NBD model may not be appropriate"*.

This point is a criticism of the method of 'means and zeros' rather than of the NBD distribution itself. If the method of maximum likelihood is used, Anscombe (1950) showed that estimates may be calculated provided :

$$(N - 1)\, s^2 > N m$$

where N is the number of observations in the sample
m is the sample mean
$s^2$ is the sample variance.

Thus, Banerjee and Bhattacharyya's argument that *"should $m < -\log p_0(t)$ happen, it indicates that a negative binomial distribution may not be appropriate"* is imprecise. There may be occasions when Anscombe's condition is met, although there is no solution to Ehrenberg's 'means and zeros' equation. In the analysis which follows, it is shown that such an occasion arises in the analysis of household toothpaste purchases analysed by Banerjee and Bhattacharyya themselves.

In order to check Banerjee and Bhattacharyya's claim that the 'compound' IG distribution gives a better fit than the NBD to toothpaste purchases across households, their analysis has been replicated. In the replication, the NBD distribution is estimated using two methods: 'means and zeros' and maximum likelihood. These results are then compared with the fit obtained by Banerjee and Bhattacharyya for the 'compound' IG distribution. The 'compound' IG distribution was estimated by the authors in two stages :

1. The IG parameters $\psi$ and $\lambda$ were estimated for each of the 289 sampled households using maximum likelihood.

2. The 'compounding' distribution was fitted to the bivariate frequency distribution obtained from the first stage of the analysis. The 'compounding' distribution has the following probability density function (pdf) :

$$(\beta/\alpha)^{1/2} (\gamma\alpha/2)^{\gamma/2} [\, H_\nu(\xi)\, B(\nu/2,1/2)\, \Gamma(\gamma/2)\, ]^{-1} \exp\{ - \gamma\alpha/2\, [\, 1 + \beta/\alpha\, (\psi - 1/\beta)^2\, ]\, \lambda \}\, \lambda^{(\gamma/2)-1}$$

for $\psi > 0, \lambda > 0$

where $\alpha > 0, \beta > 0, \gamma > 1$ are the parameters of the distribution,
B and $\Gamma$ are the complete beta and gamma functions,
$\nu = \gamma - 1$, $\xi = (\alpha\beta/\nu)^{-1/2}$ and $H_\nu(.)$ is the cumulative distribution function of Student's t-distribution with $\nu$ degrees of freedom.

Banerjee and Bhattacharyya note that : *"Due to complexity of the pdf, the method of maximum likelihood for parameter estimation would be extremely difficult to apply. We have employed the method of moments to estimate the parameters $\alpha$, $\beta$ and $\gamma$ of the compounding distribution".*

The parameters have been estimated by the methods outlined above and the goodness-of-fit of each combination of distribution and estimation method compared using chi-square statistics. The results are summarised in Table 9.1:

TABLE 9.1
Chi-square 'goodness of fit' statistics for toothpaste purchase data

| | *Negative Binomial Distribution* | | | | *'Compound' IG* | |
|---|---|---|---|---|---|---|
| *Time Period* | *df* | *Means & Zeros* | *df* | *Max Likelihood* | *df* | *Hybrid* |
| 1 | N/A | N/A | N/A | N/A | N/A | 11.186 |
| 2 | 3 | N/A | 3-5 | 2.538 | 2 | 3.596 |
| 3 | 5 | 8.398 | 5-7 | 8.380 | 4 | 5.781 |
| 4 | 6 | 25.468 | 6-8 | 13.762 | 5 | 11.510 |
| 5 | 7 | 29.586 | 7-9 | 13.059 | 6 | 17.284 |

where df - degrees of freedom
N/A - method is not applicable.

The degrees of freedom for the NBD ('means and zeros') are always one greater than for the 'compound' IG since the latter distribution has one more parameter (three, instead of two). The degrees of freedom for the NBD (maximum likelihood) are expressed in terms of a range of possible values since the sample chi-square statistic is bounded between a $\chi^2_{n-k-1}$ and a $\chi^2_{n-1}$ variable (Kendall and Stuart (1979)), where n is the number of classes and k is the number of parameters to be estimated.

The degrees of freedom for the 'compound' IG distribution for the time period of length one is not given, since grouping of data resulted in such a small number of classes that no degrees of freedom remained after taking account of the number of parameters to be estimated. The chi-square statistic is recorded to show that estimation was feasible, although it is not meaningful to test the statistic for significance.

Each model was tested for acceptance ('A') or rejection ('R') using a 5% significance level, for all lengths of time periods except the first. In the case of the NBD (maximum likelihood), tests were performed at each end of the range of possible degrees of freedom, namely n-1 and n-3, where n is the number of classes after grouping. The acceptance / rejection results of the tests are summarised in Table 9.2, with the corresponding significance levels in brackets :

TABLE 9.2
Acceptance or rejection of models for toothpaste purchase data

| *Time Period* | *Negative Binomial Distribution* | | | *'Compound' Inverse Gaussian* |
|---|---|---|---|---|
| | *Means & Zeros* | *Maximum Likelihood* Degrees of Freedom : *n-3* | *n-1* | |
| 2 | N/A | A (47%) | A (77%) | A (17%) |
| 3 | A (14%) | A (14%) | A (30%) | A (18%) |
| 4 | R (< 0.1%) | R (4%) | A (9%) | R (4%) |
| 5 | R (< 0.1%) | A (7%) | A (17%) | R (1%) |

Three points from Tables 9.1 and 9.2 are highlighted below :

1. For the time period of length one, the NBD model is inappropriate. (Neither Anscombe's criterion nor the requirements to solve Ehrenberg's equation are met.) This supports Banerjee and Bhattacharyya's criticism of the NBD's lack of comprehensiveness.

2. For the time period of length two, Ehrenberg's equation cannot be solved; thus, according to Banerjee and Bhattacharyya's reasoning, the NBD model is inappropriate. However, Anscombe's condition is satisfied and maximum likelihood (ML) estimates may be derived. Furthermore, Table 9.2 shows that the NBD model, based on ML estimates, not being rejected at the 5% level.

3. For time periods of length four and five, the NBD ('means and zeros') model was rejected decisively. The 'compound' IG model was also rejected,

although performing somewhat better. The NBD (maximum likelihood) method clearly outperforms both other methods for these time periods. This demonstrates that Banerjee and Bhattacharyya's claim to have provided *"some improvement over the NBD model"* depends on the use of the poorer estimation method of 'means and zeros'.

In conclusion, the case for the 'compound' inverse Gaussian is quite weak, since the empirical evidence has been reviewed and shown to favour the NBD if maximum likelihood estimation is used. Moreover, the *a priori* justification for the IG distribution has been shown to be have a significant weakness in that consumers' stocks drift downwards but not upwards between purchases.

## 9.6 Conclusions

### 9.6.1 Negative exponential inter-purchase distribution

It is well known that there are strong theoretical grounds to support the negative exponential distribution, with its 'memory-less' property, as a model for inter-purchase times. However, theoretical and empirical evidence has been presented in the literature which challenges the assumption of such negative exponential inter-purchase times for individual consumers.

Although there is some empirical evidence which does not support the negative exponential for 'heavy' buyers, evidence is available to support its use for 'light' and 'medium' buyers. Moreover, in a study of total customer demand for two sets of SKUs, Kwan (1991) found that the negative exponential fitted the inter-demand times of more SKUs than the Erlang-2 and the geometric distributions.

### 9.6.2 Geometric inter-purchase distribution

Kahn and Morrison (1989) argued that the geometric distribution is more appropriate than the negative exponential to model those situations when purchases are delayed to the next shopping occasion. It was argued that a similar argument may be applied to retailers or wholesalers who have a designated day of the week on which they submit stock-orders to a regional or national distribution centre.

There is little empirical evidence to support the geometric as an inter-purchase time distribution. Kwan found only 20% of her sample of SKUs to be well-fitted by the geometric distribution.

### 9.6.3 Erlang inter-purchase distribution

It has been argued that the theoretical case for the Erlang inter-purchase distribution, for an individual consumer, is coherent but dependent on a strong assumption. A consumer who purchases n items of the same SKU on each shopping occasion, and has negative exponential times between the requirement for a new item, will have Erlang-n inter-purchase times. Using the 'colouring theorem', it can be shown that random purchase quantities lead to negative exponential inter-purchase times. Furthermore, Erlang inter-purchases at individual consumer level do not necessarily lead to Erlang purchases at a total consumer level, unlike the negative exponential.

A number of studies have provided empirical evidence to support the use of the Erlang-2 distribution to model inter-purchases at individual household level. There is also empirical evidence available to support the 'household inventory effect' which was hypothesised as part of the theoretical justification for the use of the Erlang distribution. However, there have been hardly any studies of the Erlang-2 inter-demand model for total customer demand. A study by Kwan (1991) provided no support for the Erlang-2 distribution at total customer level. Further empirical work on this question will be presented in chapter 11.

### 9.6.4 Inverse Gaussian inter-purchase distribution

Banerjee and Bhattacharyya (1976) justified the use of the inverse Gaussian for modelling inter-purchase times on the basis of its application as a 'first passage time' distribution. However, it was argued that the weakness of this justification is that consumers' stocks drift downwards between purchases. It makes no sense to consider them drifting upwards, as required by the inverse Gaussian distribution.

Banerjee and Bhattacharyya offered some direct empirical evidence to support the use of the inverse Gaussian. Twelve households' inter-purchase times for toothpaste fitted the IG distribution, but no comparison was made with other distributions such as the negative exponential or Erlang-2. The authors also provided some indirect empirical evidence by comparing the goodness of fit of the 'compound' IG (based on IG inter-purchases) with the negative binomial distribution (NBD, based on negative exponential inter-purchases) to the observed distribution of purchases across households. It was shown that the improvement obtained by using the 'compound' IG instead of the NBD depended on the use of a poor estimation method. When the results were replicated using maximum likelihood, it was shown that the 'compound' IG has inferior goodness of fit to the NBD (also fitted by maximum likelihood).

These findings highlight the importance of choosing an appropriate estimation method for comparing inter-demand distributions. In the next chapter, estimation methods for inter-purchase time, demand incidence and demand distributions will be discussed.

# CHAPTER 10

## *Efficiency and Bias of Estimation Methods*

### 10.1 Introduction

In this chapter, the literature on the efficiency and bias of various estimation methods will be reviewed for four demand distributions: gamma, negative binomial (NBD), condensed negative binomial and log-zero Poisson (lzP). The properties of estimators for the condensed Poisson demand incidence distribution will also be examined. The compound Poisson and compound binomial will not be considered directly, since they are families of distributions; unless the compounding distributions are specified, it is not possible to estimate the parameters. Of course, the NBD is an example of a compound Poisson distribution and the lzP is a member of the compound binomial family. The compound inverse Gaussian will not be considered since, as argued in the previous chapter, the theoretical and empirical evidence supporting it is weak.

It is not the aim of this chapter to extend the statistical theory of efficiency and bias. This would digress from the purpose of the thesis. However, gaps in theory will be noted, as appropriate, and a maximum likelihood estimator for the condensed Poisson distribution is derived.

Analysis of the statistical properties of estimation methods is not merely an academic exercise. In chapter 9, it was shown that the NBD performed better as a distribution of purchases across households when the parameters were estimated using the method of maximum likelihood, than when the 'means and zeros' method was used.

A comprehensive analysis of distributions using various estimation methods is required, in order to compare their goodness of fit to empirical data. Such an analysis will be presented in chapter 11. In this chapter, the conditions under which each estimation method may be applied will be identified and the efficiency and bias of each method will be examined.

## 10.2 Inter-purchase time and demand incidence distributions

### 10.2.1 Inter-purchase time distributions

The moments estimators and maximum likelihood estimators coincide for each of the three distributions to be considered in the next chapter : geometric, negative exponential and Erlang-2. Therefore, these estimators will be used in the empirical analysis. Problems of calculating mean inter-purchase times for fast moving SKUs will be discussed in section 11.4.

### 10.2.2 Poisson and binomial demand incidence distributions

The moments and maximum likelihood estimator coincides for the Poisson distribution. For the binomial distribution, the maximum likelihood estimator r/n will be used for the 'probability of a purchase occurrence', where r is the number of working days on which a purchase occurs and n is the total number of working days. The difficulties in using this estimator when there is more than one demand occurrence in a working day will be discussed in sub-section 11.5.3.

### 10.2.3 Condensed Poisson demand incidence distribution

For this distribution, the moments and maximum likelihood estimators do not coincide. The moments estimator is simply $\Sigma_{j=1}^{n} x_j / n$, where $x_j$ (j=1,...,n) are the recorded inter-purchase times. The maximum likelihood estimate is a root of the following equation in $\lambda$ :

$$2n - \frac{m}{1+\lambda} = \frac{2}{\lambda} \sum_{j=m+1}^{n} x_j + \sum_{j=m+1}^{n} \frac{1/(2x_j+1) - x_j/2\lambda^2}{\lambda/(2x_j+1) + 1 + x_j/2\lambda}$$

where there are n observations, m of which are zero.

The derivation of the maximum likelihood equation is given in Appendix 10.1. The properties of the equation are discussed in the appendix. In brief, it is easy to show that the equation must have at least one positive finite root but it has not been proven that there is only one such root.

## 10.3 Gamma demand distribution

### 10.3.1 Method of moments

Burgin and Wild (1967), in their exposition of the gamma distribution for stock control, recommend that the distribution be fitted by the method of moments. The estimate of the shape parameter is then $k = D_1^2 / D_2$, where $D_1$ and $D_2$ are the sample estimates of the mean and variance of demand. It is immediately apparent that the estimation method may be used for any data with non-zero sample variance.

It may be shown (see, for example, Bury (1975)) that the asymptotic efficiencies of the method of moments for the two-parameter Gamma distribution are as follows :

$$\text{Efficiency (k estimator)} = \frac{1}{2\,[\,k\,\psi'(k) - 1\,]\,(\,1 + k\,)}$$

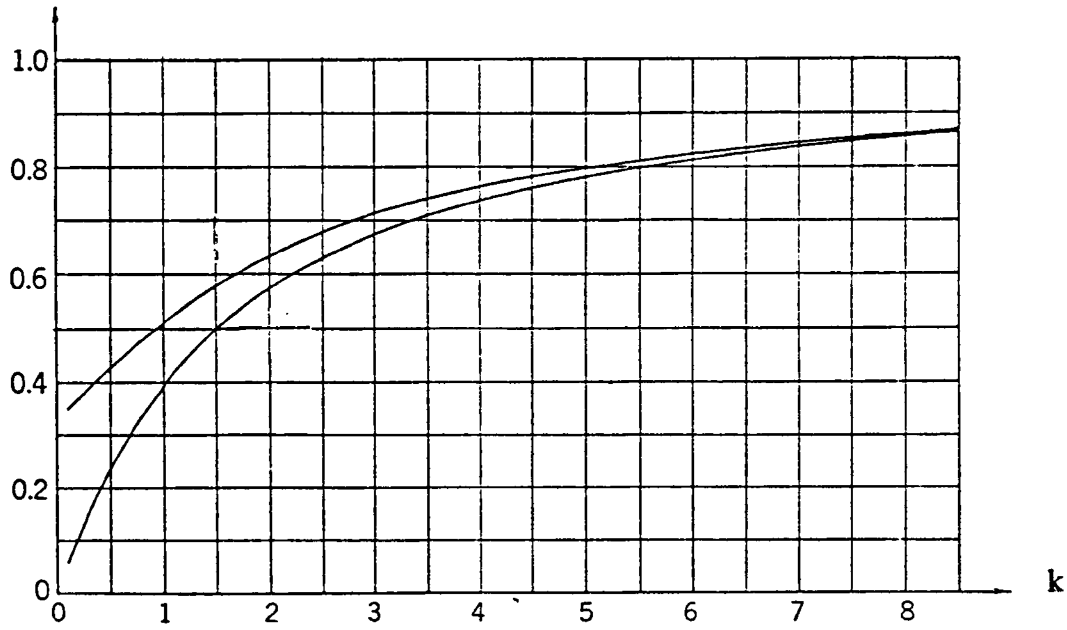
$$\text{Efficiency }(\alpha\text{ estimator}) = \frac{1}{[\,k\,\psi'(k) - 1\,]\,(\,2k + 3\,)}$$

where k and $\alpha$ are the shape and scale parameters

$\psi'(k)$ is the trigamma function : $\dfrac{d^2 \ln \Gamma(k)}{dk^2}$

Bury (1975) plotted the efficiencies of the two estimators against the shape parameter, k, as shown in Figure 10.1. The graph shows that the moments estimators may be inefficient for low values of k and particularly so if k is less than two.

Figure 10.1
Efficiency of moments estimators for the gamma distribution




*Upper line* Efficiency of scale parameter
*Lower line* Efficiency of shape parameter

### 10.3.2 Maximum likelihood estimation

Since the method of moments may be inefficient, the additional complexity of maximum likelihood can be justified by potential efficiency gains. Maximum likelihood estimates are obtained by solving the following two equations :

$$\alpha = \frac{k}{m}$$

$$\ln(k) - \psi(k) = \ln(m) - \ln(G)$$

where m and G are the arithmetic and geometric means of the observed data

and $\psi(k) = \frac{d}{dk} \ln \Gamma(k)$ is the digamma function.

Although efficiency gains may be made by using maximum likelihood, there are a number of difficulties associated with using the method.

The first difficulty associated with maximum likelihood (ML) estimation, in the context of demand distributions, is apparent from the second likelihood equation : the geometric mean, G, must be greater than zero for ln(G) to be well-defined. When the gamma distribution is used to represent times between events, all inter-event times are positive and it follows that G is also positive. However, if the gamma distribution is used to represent demand, then zero values, and a zero geometric mean, are possible. In these cases, which may be common in inventory management, the maximum likelihood estimate is not well-defined.

The second difficulty is that ML estimates are biased for small samples. Shenton and Bowman (1977) found approximate functions for the bias of the ML estimate of the gamma shape parameter. Using Shenton and Bowman's approximation, the following sample sizes, shown in Table 10.1, are required to ensure that the bias is less than 10% of the population shape parameter :

TABLE 10.1
Sample sizes required for ML bias of less than 10% for Gamma shape parameter

| | | | | | |
|---|---|---|---|---|---|
| *Population shape parameter* | 0.1 | 0.5 | 1.0 | 5.0 | 25.0 |
| *Sample size required* | 20 | 25 | 28 | 32 | 33 |

The final difficulty is that when maximum likelihood estimation is possible, it may be computationally expensive. However, approximation methods have been devised for ML estimates for the gamma distribution. Two of these methods will now be reviewed. The estimation of the shape parameter is the focus of the discussion. The ML estimate of the scale parameter, $\alpha$, is related to the estimate of the shape parameter, k, by the equation $\alpha = k / m$, where m is the sample mean.

Burgin (1975) provided a neat approximation to the solution of the likelihood equations for the shape parameter k, given below :

$$k \approx \frac{1}{4\lambda}\left\{ 1 + \left( 1 + \frac{4\lambda}{3} \right)^{1/2} \right\}$$

where $\lambda = \ln(m) - \ln(G)$
and m is the sample arithmetic mean
G is the sample geometric mean.

However, no error bounds to this approximation were given.

Greenwood and Durand (1960) gave the following approximation :

$$k \approx \frac{1}{y}\,[\,0.5000876 + 0.1648852\,y - 0.0544274\,y^2) \qquad \text{for } 0 \le y \le 0.5772$$

$$k \approx \frac{8.898919 + 9.059950\,y + 0.9775373\,y^2}{y\,[\,17.79728 + 11.968477\,y + y^2\,]} \qquad \text{for } 0.5772 \le y \le 17.0$$

where $y = \ln(m / G)$.

Greenwood and Durand estimated the maximum error in the first range ($0 \le y \le 0.5772$) as 0.0088% and the maximum error in the second range ($0.5772 \le y \le 17.0$) as 0.0054%. Therefore, the approximation will give an error no greater than 0.01 % in the overall range ($0 \le y \le 17.0$), an accuracy which is acceptable in practice. Since the Greenwood and Durand approximation is simple to use and tight error bounds have been established, this approximate estimation method will be used in the empirical analysis of chapter 11.

## 10.4 Negative binomial demand distribution

### 10.4.1 Estimation methods for the NBD

Anscombe (1950) reviewed five alternative estimation methods to maximum likelihood for the negative binomial distribution (NBD). Two methods were discounted as they are always outperformed, in terms of efficiency, by one of the remaining three methods. These methods are as follows :

1. Means and Zeros
2. Method of Moments
3. Transformation Method.

The first method was discussed in chapter 9. It may be used whenever there are some zero observations. If there are no zero observations, then the 'means and zeros' equation, from which the NBD exponent is derived, has no solution.

The second method was also discussed in chapter 9. The method may be used whenever the sample variance is strictly greater than the sample mean.

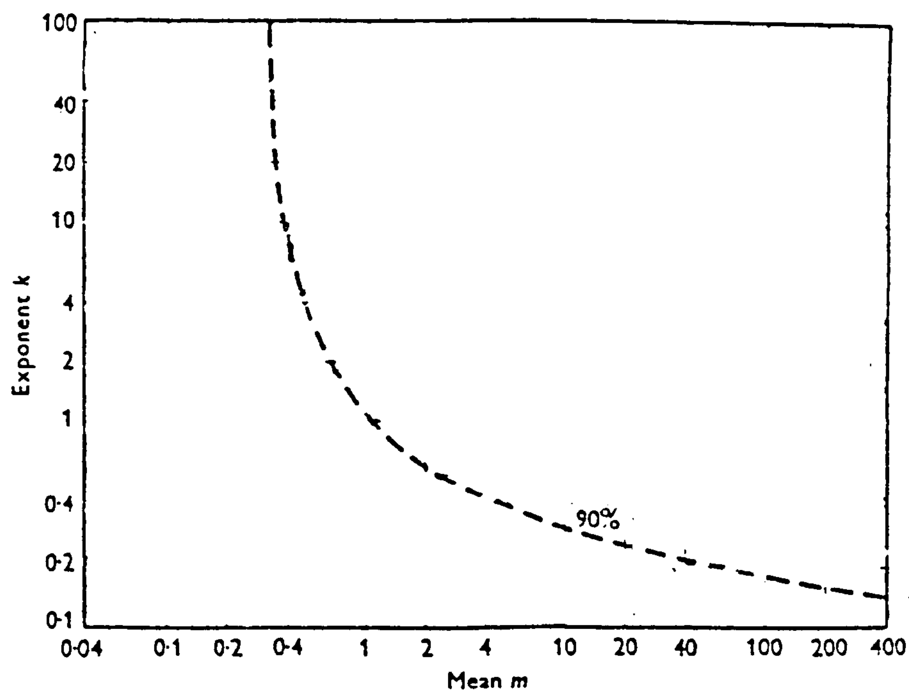
The 'transformation method' was introduced by Anscombe (1948) and is based on a logarithmic or inverse hyperbolic sine transformation, depending on the values of the exponent, k, and the sample mean, m. Anscombe (1950) found that the transformation method had high efficiency (over 90%) in only a small region of the parameter space not already covered by high efficiency of one of the first two methods. Moreover, Pieters et al (1977) found the method to give a significant upward bias in the estimation of the NBD exponent. Therefore, the transformation method will not be analysed further.

10.4.2 Efficiency of the 'means and zeros' method

Ehrenberg (1959) claimed that this method is "*at least 90% efficient for our kind of data, and often a good deal more so*". In this quotation, 'our kind of data' means market research data obtained from consumer panels. It is important to note that Ehrenberg and his associates used the NBD as a model of purchases across households; they did not model total consumer purchases over time.

Anscombe (1950) investigated the efficiency of the 'means and zeros' method. His analysis showed that the values of the exponent, k, and the sample mean, m, for which the method is at least 90% asymptotically efficient are in the region below the curve shown in Figure 10.2 :

Figure 10.2
Large sample efficiencies for 'means and zeros'




This graph shows that if the exponent is below 0.2 and the mean value is below 40, then the method of 'means and zeros' is always at least 90% efficient. Many studies of consumer purchases across households show values of k and m in this region; early examples include Ehrenberg (1959), Chatfield et al (1966), Grahn (1969) and Chatfield and Goodhardt (1973). In each of these studies, the method of means and zeros was used and gave good efficiency. However, it may be recalled from the previous chapter that Banerjee and Bhattacharyya's (1976) estimation of the NBD was much improved for two time periods by using maximum likelihood instead of means and zeros. The reason for this poor performance may now be explained. The values of the mean (sample mean) and exponent (maximum likelihood estimate) for these time periods are summarised in Table 10.2 :

TABLE 10.2
Parameter values of NBD in Banerjee & Bhattacharyya's study

| *Length of Time Period* | *Exponent (k)* | *Mean (m)* |
|---|---|---|
| 4 | 4.758 | 2.564 |
| 5 | 4.582 | 3.173 |

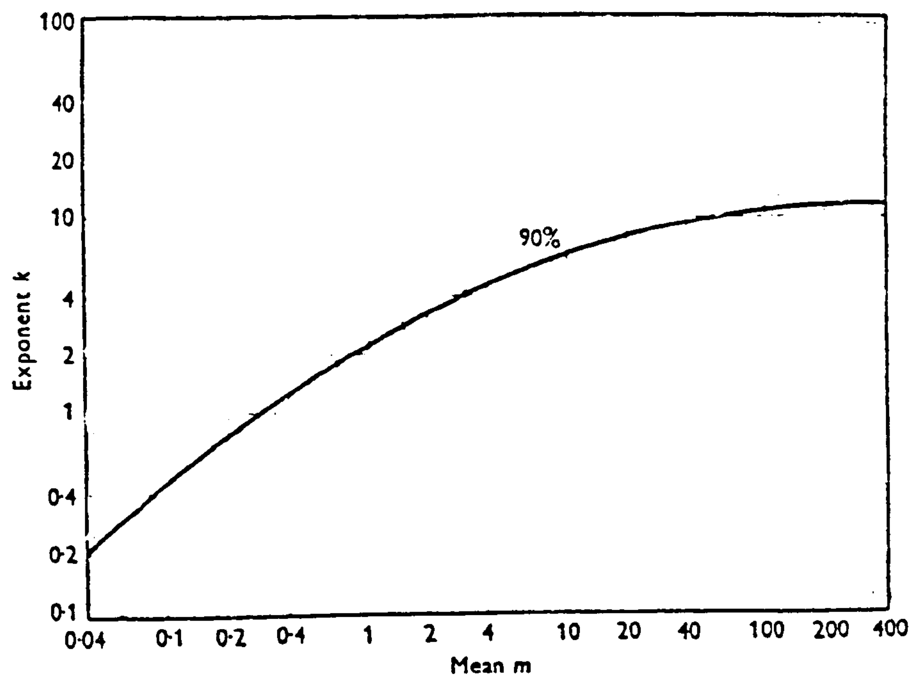
Referring back to Figure 10.2, it is clear that these values lie above the curve and, therefore, outside the region in which 90% efficiency is attained. In fact, the efficiency at both points is approximately 50%, thus explaining the poor performance of the means and zeros method in this case.

### 10.4.3 Efficiency of the method of moments

Kwan (1991) estimated the parameters of the NBD distribution of demand over time using the method of moments. She commented on the NBD that "*both the parameters* *and the probabilities are easy to calculate. Therefore, it is highly recommended".*

Anscombe (1950) identified the region of the parameter space in which at least 90% asymptotic efficiency will be attained; it is the region above the curve shown in Figure 10.3 :

Figure 10.3
Large sample efficiencies for the method of moments




The SKUs investigated by Kwan (1991) have been analysed to find the efficiency expected by the method of moments. The results are shown in Table 10.3. The results in the body of the table show the number of SKUs (and the percentage of each category) which satisfy the efficiency criteria on the left hand side of the table.

TABLE 10.3
Efficiency of moment estimator in Kwan's study

| | *Vehicle Spares* | *Steel Manufacturing Spares* | *Total* |
|---|---|---|---|
| 90% Efficient | 10 (47.6%) | 9 (17.6%) | 19 (26.4%) |
| 75% Efficient | 18 (85.7%) | 17 (33.3%) | 35 (48.6%) |
| 50% Efficient | 20 (95.3%) | 41 (80.4%) | 61 (84.7%) |

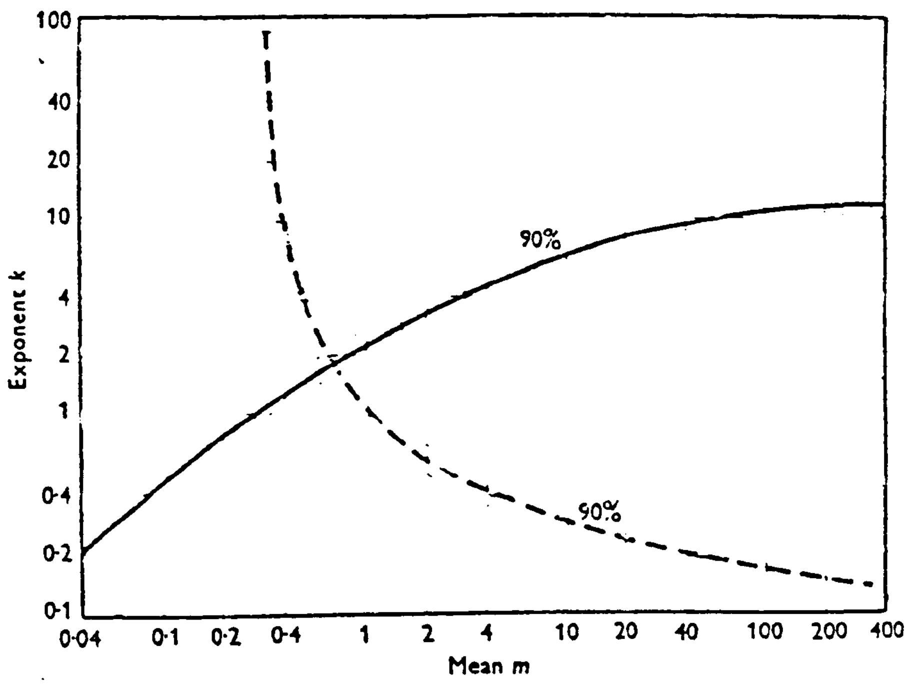
Note: The table excludes those SKUs with sample variance less than the sample mean.

Table 10.3 shows that the method of moments attains ninety per cent efficiency for a minority of SKUs. Seventy five per cent efficiency is achieved for most vehicle spares, but for only 33% of the steel manufacturing spares. Therefore, for this latter category of spares, the method of moments must be used with some caution. Fifty per cent efficiency is achieved for over 80% of spares, but for the remaining proportion the method of moments is not to be recommended.

10.4.4 Maximum likelihood estimation

If the two previous graphs are superimposed, then the following graph is obtained, shown in Figure 10.4 :

Figure 10.4
Large sample efficiencies for moments and 'means and zeros'




The graph shows that the methods reviewed in this section, moments and 'means and zeros' are complementary since they perform well in different regions of the parameter space, with only a relatively small intersection on the left side of the graph. However, the two methods do not cover all of the parameter space. The method of maximum likelihood is required for the remaining area on the right side of the graph.

Anscombe (1950) showed that the maximum likelihood estimator of the exponent k is obtained by solving the following equation for x :

$$N \ln (1 + m / x) = \Sigma_{j=1}^{\infty} n_j \; [ 1/x + 1/(x+1) + \ldots + 1/(x+j-1) ]$$

where N is the sample size
m is the sample mean
$n_j$ is the number of observations equal to j.

The solution to the above equation is not prohibitively expensive to obtain in computational terms. Since it may provide significant efficiency gains in the area of the parameter space not covered by moments or 'means and zeros', maximum likelihood will be used in the empirical analysis of the next chapter, with Anscombe's equation used to derive the ML estimates.

10.4.5 Bias of method of moments, 'means and zeros' and maximum likelihood

Pieters et al (1977) conducted a simulation experiment to determine the bias of the above methods in estimating the exponent and scale parameters of the NBD. They found the 'means and zeros' estimator to be inferior to the moments and maximum likelihood estimators in terms of bias. They concluded that there was little to choose between moments and maximum likelihood in terms of the frequency or magnitude of bias. However, there is some evidence that maximum likelihood is less biased than moments when the parameter p (p = mean / exponent) is large relative to the exponent, k.

## 10.5 Condensed negative binomial demand distribution

### 10.5.1 Method of moments

The moments of the condensed negative binomial distribution were derived by Chatfield and Goodhardt (1973). Equating the sample moments to the population moments,

$$m = a\,k$$

$$s^2 = m\,a + \frac{m}{2} + \frac{[\,1 - (1+4a)^{-m/a}\,]}{8}$$

The method may be used whenever the sample variance is strictly greater than half the sample mean.

No studies have been found in the literature on the efficiency and bias of the moments estimator for the condensed negative binomial distribution (CNBD), nor on estimation by maximum likelihood. A detailed investigation into the efficiency of estimation methods would be very welcome, along similar lines to the PhD thesis by Rao (1969) on the properties of estimators for the log-zero Poisson distribution. However, such an investigation is beyond the scope of this thesis.

Although little is known about the efficiency and bias of moments estimation for the CNBD, some empirical evidence is available which compares moments estimators to 'means and zeros' estimators. Morrison and Schmittlein (1981) analysed the distribution of purchases across customers for peanut butter, garbage bags, food bags and lawn bags. The data was taken from a panel of over ten thousand households who recorded their purchases for the calendar year of 1977. Schmittlein and Morrison's comparisons between estimates of the exponent, k, and scale parameter, $\alpha$, showed the following results :

TABLE 10.4
Schmittlein and Morrison's comparison of CNBD exponent estimators

| | *Peanut Butter* | *Garbage Bags* | *Food Bags* | *Lawn Bags* |
|---|---|---|---|---|
| Moments | 0.6702 | 0.5092 | 0.5613 | 0.1338 |
| Means and Zeros | 0.4964 | 0.2815 | 0.2385 | 0.0503 |

TABLE 10.5
Schmittlein and Morrison's comparison of CNBD scale parameter estimators

| | *Peanut Butter* | *Garbage Bags* | *Food Bags* | *Lawn Bags* |
|---|---|---|---|---|
| Moments | 0.1890 | 0.2281 | 0.6159 | 1.0290 |
| Means and Zeros | 0.1400 | 0.1261 | 0.2617 | 0.3872 |

It is apparent that the two estimation methods give quite different results and that the moments estimators are consistently higher than the 'means and zeros' estimates. Hence, a more detailed review of the efficiency and bias of the method of 'means and zeros' is required. This is an area for future research.

### 10.5.2 'Means and zeros'

In the previous sub-section, some empirical evidence was reviewed which demonstrated that widely different parameter estimates may be obtained by using the two estimation methods of moments and 'means and zeros'. Schmittlein and Morrison (1981) showed that better predictions of the average number of purchases (by consumers who had consumed a given amount in a previous period) were obtained by the method of moments than 'means and zeros' when using the CNBD model of purchases across customers. For two product groups, food bags and lawn bags, the improvement in predictive accuracy was dramatic.

Schmittlein and Morrison's evidence is not of immediate relevance to estimation of parameters for the distribution of total purchases (ie purchases from all consumers). However, the analysis does demonstrate that wide differences in both the estimates themselves and their predictive accuracy are possible. For this reason, it is necessary to use both estimation methods in order to assess the goodness of fit of the condensed negative binomial distribution.

## 10.6 Log-zero-Poisson demand distribution

### 10.6.1 Efficiency of the method of factorial moments

Kwan (1991) fitted the log-zero-Poisson (lzP) distribution to the demand histories of 85 SKUs using the method of factorial moments. The reason for this choice was the method's tractability. The estimators are given below :

$$\theta = \frac{f_0}{n}$$

$$\tau = \frac{2\,\mu_{(2)}}{\mu_{(1)}} - \frac{\mu_{(3)}}{\mu_{(2)}}$$

$$p = \frac{\mu_{(3)}/\mu_{(2)} - \mu_{(2)}/\mu_{(1)}}{\tau}$$

where $\theta$, $\tau$ and p are the parameters of the lzP distribution, as described in chapter 8, n is the sample size, $f_0$ is the number of zero observations and $\mu_{(1)}$, $\mu_{(2)}$ and $\mu_{(3)}$ are the first, second and third sample factorial moments.

Rao (1969) proved that the estimator $f_0/n$ is the unique minimum variance unbiased estimator of the parameter $\theta$.

Rao examined the 'relative efficiency' of 27 alternative estimators to maximum likelihood. In each method, the statistic $f_0/n$ was used as the estimator for $\theta$; alternative estimators for the other two parameters were explored. The ratio of the generalised variance of each method to that of the best method (of the 27) was calculated to give the 'relative efficiency' of each method. Rao concluded that the method of factorial moments performed very poorly : "*The relative efficiency is very* *low. Hence, the use of this method should be restricted to preliminary experimentation where one is interested in a quick simple estimator for the sole purpose of guiding towards more careful experimentation*".

Kwan's results on the goodness of fit of the log-zero-Poisson distribution may require re-assessment in the light of Rao's findings. It is possible that alternative estimation methods may provide better fit to the data, thereby strengthening the case for the lzP as a demand distribution.

10.6.2 Maximum likelihood estimation

Katti and Rao (1970) showed that the likelihood equations for the lzP distribution are as follows :

$$\theta = \frac{f_0}{n}$$

$$\frac{\tau\, p\,(1-\theta)}{\log\,[1 + p - p\, e^{-\tau}\,]} = m$$

$$\sum_{i=1}^{\infty} (i+1)\,\frac{f_i}{n}\,\frac{P(i+1)}{P(i)} + P(1) = m$$

where m is the sample mean, $f_i$ denotes the frequency of observation *i*, and P(i) denotes the probability of *i* according to the lzP distribution.

A number of studies were conducted by Katti's PhD students - Rao (1969), Huque (1973) and Khedr (1980) - but none included an analysis of the biases in likelihood estimates of $\tau$ and p. Consequently, the small sample properties of ML estimates are not well understood. This is an issue which requires further research.

## 10.7 Conclusions

In chapter 9, it was argued that taking care over the choice of estimation method for demand distributions is not merely an academic exercise. Significant improvements in estimation may be made simply by avoiding inappropriate methods. Moreover, a number of authors have used such inappropriate methods in their analysis of demand, demonstrating the need for this issue to be highlighted.

In this chapter, a literature review was conducted on the statistical properties of a range of estimators for the gamma, NBD, CNBD and lzP distributions. The properties of estimators for the condensed Poisson distribution were also examined.

It was shown that the moments and maximum likelihood (ML) estimators for the condensed Poisson distribution do not coincide. The equation for ML estimators was derived. It was shown that this equation must have at least one positive finite root, but it remains to be proven that there is only one such root.

It was noted that the moment estimators of the Gamma shape parameter may be inefficient, particularly if the population shape parameter is low. Maximum likelihood estimates suffer from problem of non-computability if there are any zero observations and bias for small samples. However, an excellent approximation to the solution of the likelihood equations has been provided by Greenwood and Durand (1960), making the estimation method computationally inexpensive.

Anscombe (1950) showed that the negative binomial distribution (NBD) may be estimated efficiently by the method of 'means and zeros', but only in a restricted

# CHAPTER 11

## *Empirical Analysis of Demand Distributions*

### 11.1 Introduction

In this chapter, a sample of stock-keeping units from Pillar Engineering Supplies (PES) Ltd, Leicester, will be analysed. The company supplies tools such as drills and hammers to customers in an area concentrating on the central counties of England.

There are a number of reasons for conducting this empirical analysis. Firstly, it will enable the assumption of Poisson demand incidence to be tested. The compound Poisson family of demand distributions forms the foundation upon which the variance law theory of chapter 12 is built.

The second reason for conducting this analysis is connected with the discussion in the previous chapter. Very little evidence is available on the effect of using different estimation methods on the parameter estimates and the goodness-of-fit of the corresponding distributions in the context of inventory management.

Thirdly, the analysis will replicate some previous studies on the goodness of fit of the demand distributions discussed in chapters 8 and 9 - log-zero-Poisson, negative binomial, condensed negative binomial and gamma - using some new data. Although this work is not essential in the later development of theory, the results are presented here since the literature on this subject is surprisingly meagre and, therefore, a replication study is of interest in its own right.

## 11.2 Stock management at Pillar Engineering Supplies Ltd

### 11.2.1 Policy objectives

According to PES internal documentation (Pillar Engineering Supplies Ltd (1993)), the company's aims are to optimise :

* Financial investment in stock
* Service levels to customers
* A rationalised supplier base
* Stock turn

and to tailor the PES product base to business and market requirements.

### 11.2.2 SKU categories at Pillar Engineering Supplies

One measure taken by PES to help achieve the above objectives was to introduce a product categorisation system which reflected the number of customers demanding a product. The introduction of this system was intended to reduce inventories of products required by few customers. The classification system is as follows :

*A* Products sold to 16 or more customers company wide during the last year

*B* Products sold to 4 to 15 customers company wide during the last year

*C* Products within active promotion profile

*D* All products that do not fall into other categories - including all 'specials' - items without a PES code

*E* Obsolete products.

automatically for A and B classes, with manual intervention for the C class. The D class constitutes 'customer specific stock'; demand for these products has not expired but, as there are so few customers, the stock is purchased on demand. Class E products are not re-ordered if any items are ever sold.

The empirical analysis of this chapter and chapter 13 will concentrate on the A and B stock categories. The reasons for not including the other categories in the analysis are summarised below.

It would be difficult to examine the *underlying* distribution and variance of demand for class C items, since demand patterns may change dramatically at the beginning and end of a period of promotional activity. Consequently, these items are excluded.

For PES, class D items would also be unaffected by centralisation, since recommended stock levels are zero and stock is purchased on demand. However, other companies may include such items in their stocked range with very low recommended stock levels. The effect of centralisation of such items may be assessed using the methods discussed in chapter 6; the variance law theory to be developed in part IV of the thesis is not required for these SKUs. Therefore, class D items are also not considered. Class E items are excluded since centralisation of obsolete stock affects neither the inventory requirement (zero) nor the inventory holding. Of course, obsolete stock may be centralised for the purpose of sell-off or scrap, but this is a matter beyond the scope of this investigation.

### 11.2.3 Ordering policies

For A and B category stock, the company uses a minimum / maximum system, with periodic review at least once per week. The minima and maxima are set at multiples of weeks demand. If the stock falls below the minimum level, then an order is triggered. Sani and Kingsman (1995) show that if the parameters of such an operational system are set appropriately, then it can perform similarly to more sophisticated alternatives in terms of stock-holding and service level.

## 11.3 Data for analysis

### 11.3.1 Sample of SKUs

Since the SKUs in categories in categories A and B number over ten thousand, it was necessary to extract a sample for analysis. The objectives in choosing the sample were as follows :

* To ensure that, within each category, the widest possible range of movements are covered, from the slowest to the fastest moving SKUs.
* To sample higher proportions of stock from groupings with higher levels of demand. It will be required to test the sensitivity of the compound Poisson distribution and its resultant variance law against a sufficient sample of fast-moving SKUs. Since, following the Pareto effect, the number of such SKUs declines as demand increases, higher sampling proportions are needed.
* To make the sampling method as simple as possible to turn into computer code for PES staff, who were under great work pressure at the time the sample was taken.

To meet the second objective, it was necessary to stratify the A and B categories into a number of segments. Although it would be natural to segment by number of movements or volume of sales, the easiest segmentation to implement was 'number of customers' since this was already stored as a key field on the stock movement files. Since sales volume and movements are correlated with number of customers, it was considered that this would be an acceptable approximation.

Within each segment, a proportion of SKUs were sampled using a 'systematic sampling' method, selecting every *n*th SKU in the segment. Since the SKUs were sorted in descending order of number of customers, 'systematic sampling' could have introduced some bias in the fastest moving SKUs, according to the choice of the first SKU to be sampled. However, this problem did not arise, as a 100% sample of the fastest moving SKUs was taken. The details of the sampling proportions are given in Table 11.1 :

TABLE 11.1
Proportions of SKUs sampled

| *Stock Category* | *Number of Customers* | *Sampling Proportion* |
|---|---|---|
| A | ≥ 500 | 100% |
| A | 200 - 499 | 20% |
| A | 100 - 199 | 10% |
| A | 50 - 99 | 5% |
| A | 15 - 49 | 2% |
| B | ≥ 10 | 2% |
| B | < 10 | 1% |

This sampling method yielded a sample of 230 SKUs. This was considered to be an adequate sample size to permit comparison of demand distributions and analysis of variance laws.

### 11.3.2 Length of data history

In any inventory problem, operational or strategic, it is necessary to predict the distribution and the parameters which will pertain over the forthcoming lead-time. In practice, distributions must be tested against historical data and it is worth noting that the distribution or its parameters, or both, may have changed over the historical period. This principle underpins all forecasting systems where data is discounted as it becomes older.

The findings of Kwan (1991) indicated strong empirical support for the log-zero-Poisson but these results were based on data histories as long as ten years. However, this does not necessarily imply that the log-zero-Poisson was a good distribution of demand at any particular time, unless a stationary mean is assumed. This assumption is unlikely to be sustainable for all of the 85 SKUs examined over such long time periods.

Empirical evidence to be presented in chapter 13 shows that for Pillar Engineering Supplies, the order-size distribution may be assumed to be stationary but the same assumption does not hold for the demand and demand incidence distributions.

For the reasons outlined above, longer data histories are not necessarily required to ascertain current demand distributions. However, very short data histories are not sufficient to permit accurate estimation of the parameters and testing of the goodness of fit of the distributions.

For consistency with PES ordering policies, which express minimum and maximum stock levels in terms of number of weeks demand, it was decided to measure weekly demand and demand incidence. This means that at least six months data would be required for reasonably accurate estimates of the variance of demand and demand incidence. Six months data would also suffice to test the goodness of fit of demand,

demand incidence and inter-purchase distributions for faster moving SKUs. However, no tests could be performed on the slowest moving SKUs, with just four movements per annum, if only six months data were available. At least twelve months data is required to allow a Kolmogorov-Smirnov test on the goodness-of-fit of the inter-purchase distribution. Hence, it was decided to analyse a period of twelve months demand data for the sampled SKUs; this covered the whole calendar year of 1994.

### 11.3.3 Data fields extracted

The following fields were extracted from the stock movement files :

1. PES product code
2. Movement date
3. Account code
4. Movement quantity
5. Return to stock flag
6. Resultant stock quantity.

The analyses in this chapter require the first, second and fourth fields. The 'account code' is used to exclude receipts from suppliers. The 'return to stock flag' is used to exclude inter-company transfers and customer returns. The 'resultant stock quantity' allows a check to be made on the comprehensiveness of the data supplied; if a record were missing, this would be revealed by the current stock quantity being inconsistent with the previous stock quantity and subsequent sales and receipts. Checks on a sub-sample of SKUs revealed no omissions or inconsistencies in the records of sales and receipts.

### 11.3.4 Calculation of inter-purchase times, weekly demand and weekly demand incidence

Inter-purchase times have been calculated as the number of working days since the last purchase. Thus, Saturdays and Sundays are excluded from consideration, as are bank holidays (except for Good Friday when orders were received and goods were sold by the company). Since inter-purchase time calculations were complicated by the 1994 Christmas period, it was decided to exclude the last weeks of December 1994 from the analysis.

The demand and demand incidence values were totalled for each week, with no adjustment for those weeks containing a bank holiday. It was considered that this four-day working week effect was relatively unimportant and may be compensated by additional buying on the day after the bank holiday, if trade customers' stocks had been further depleted over the bank holiday weekend.

### 11.3.5 Seasonality of demand

PES does not adjust demand forecasts for seasonality, since it is has been a consistent finding that its products exhibit no strong seasonal patterns. Therefore, no seasonal adjustments to demand has been made for the analysis in this chapter and chapter 13.

## 11.4 Estimation methods and goodness-of-fit tests

### 11.4.1 Estimation methods

In chapter 10, a number of estimation methods were examined for inter-purchase, demand incidence and demand distributions. A summary of the methods discussed for each distribution is presented in Table 11.2 :

TABLE 11.2
Estimation methods

| | | | |
|---|---|---|---|
| **Inter-purchase distributions** | *Exponential* | *Geometric* | *Erlang-2* |
| | Moments (=ML) | Moments (=ML) | Moments (=ML) |
| **Demand incidence distributions** | *Poisson* | *Binomial* | *Condensed Poisson* |
| | Moments (=ML) | ML | Moments, ML |
| **Demand distributions** | *NBD* | *lzP* | *CNBD* |
| | Moments, ML means and zeros | Factorial moments ML | Moments, means and zeros |
| | *Gamma* | | |
| | Moments, ML | | |

Notes :

| | |
|---|---|
| NBD | negative binomial distribution |
| CNBD | condensed negative binomial distribution |
| lzP | log-zero-Poisson distribution |
| ML | maximum likelihood. |

Each of the methods shown above will be used for the appropriate distributions in this chapter. This analysis will allow a comparison of distributions, for the same estimation method, and a comparison of estimation methods, for the same distribution.

## 11.4.2 Goodness-of-fit tests

Two methods of testing goodness-of-fit will be used in this analysis :

* Chi-square test, for demand and demand incidence distributions.
* Kolmogorov-Smirnov test, for inter-purchase distributions.

Since demand and demand incidence are discrete variables, represented by discrete distributions (with the exception of the gamma distribution), the chi-square test is a natural choice. It has been found that approximately 200 of the 230 sampled SKUs are testable by chi-square; the exact figures vary by distribution. The remaining SKUs do not have a sufficient number of 'cells' with an expected frequency of five or more to allow a chi-square test to proceed. However, the inter-purchase time distributions of these SKUs may be tested using the Kolmogorov-Smirnov (KS) test. Unlike the chi-square, the KS test may be applied to such small samples.

The KS test applies to continuous data, but for PES the date of purchase is the finest level of detail available, which yields discrete data. This is not a serious problem for slow-moving items since the effect of 'integerising' (rounding to a whole number) the inter-purchase time will be small. However, Johnston (1975) showed how the integerisation effect can give very misleading results. Using a real example, he demonstrated that a faster-moving SKU, with a good fit to a Poisson demand incidence distribution, did not have a good fit to an exponential inter-purchase distribution because of the integerisation effect.

From this discussion, it may be observed that the two tests are complementary. The chi-square test cannot be used for the demand and demand incidence distributions of the slowest moving SKUs, but the KS test can be applied to their inter-purchase distributions. For faster moving SKUs, the KS test may encounter problems of 'integerisation' of inter-purchase times, but the chi-square test can be used for demand and demand incidence distributions.

## 11.5 Goodness-of-fit of demand incidence and inter-purchase time distributions

### 11.5.1 Presentation of results

Each demand incidence distribution will be tested using a chi-square test at a 5% significance level and each inter-purchase distribution using a Kolmogorov-Smirnov test at a 5% significance level. The results will be presented for all SKUs and sub-divided into the following categories, as shown in Table 11.3 :

TABLE 11.3
Categorisation of SKUs by speed of movement

| | *Range of Movements* | *Number of SKUs* |
|---|---|---|
| Class I | ≥ 500 | 59 |
| Class II | 150 - 499 | 54 |
| Class III | 50 - 149 | 62 |
| Class IV | < 50 | 55 |

The movement categories have been chosen somewhat arbitrarily and are not related to any categorisation used by PES. Class IV corresponds to SKUs with less than one movement per week, class III to less than three and class II to less than ten.

Presentation of results by speed of movement will allow each distribution to be assessed for its particular suitability to classes of movement as well as its general suitability across all classes.

### 11.5.2 Poisson demand incidence and exponential inter-purchase times

Since the sample mean is identical to the maximum likelihood estimate, the goodness-of-fit test must be performed for both n-1 and n-2 degrees of freedom, where n is the number of classes with expected values greater than or equal to five. The results for both values of the degrees of freedom are summarised in Table 11.4 :

TABLE 11.4
Percentage of SKUs with demand incidence fitted by the Poisson distribution

| *Degrees of freedom* | *Class I* | *Class II* | *Class III* | *Class IV* | *Total* |
|---|---|---|---|---|---|
| n-2 | 84.7 | 85.2 | 82.3 | 76.9 | 83.1 |
| n-1 | 89.8 | 92.6 | 88.7 | 94.2 | 91.2 |

All of the SKUs in classes I, II and III are included in the above calculations. For class IV, 26 SKUs are included in the 'n-2 degrees of freedom' analysis and 52 SKUs in the 'n-1 degrees of freedom' analysis.

The above results show very high percentages of SKUs with demand incidence well-fitted by the Poisson distribution. It is also noteworthy that the percentages are very consistent between classes I, II and III. The percentage range of class IV (76.9% to 94.2%) is wider than the others. However, the result is consistent with 87.6% of SKUs in this class having inter-purchase times well-fitted by the exponential distribution. The other results for the exponential distribution are not recorded here. They give much lower percentages, as a result of the 'integerisation' effect discussed in sub-section 11.4.2.

### 11.5.3 Binomial demand incidence and geometric inter-purchase times

Most class I and class II SKUs have had purchases on every working day, making it impossible to test a non-trivial binomial demand incidence model. Since daily information is the finest level of detail available from the PES files, the 'probability of success' of a purchase being made in a day must be estimated at unity for these items. The use of moments estimates was considered but was infeasible, since the sample variance of demand incidence exceeds the sample mean for almost all of the sampled SKUs.

Results were obtained for the goodness-of-fit of the binomial for class III and IV SKUs using the ratio of the number of movements, over the whole year, to the total number of working days as the estimator for the 'probability of success'. This ratio coincides with the maximum likelihood estimator if there are no working days with more than more than one incidence of demand. The results of this analysis are summarised in Table 11.5 :

TABLE 11.5
Percentage of SKUs with demand incidence fitted by the binomial distribution

| *Degrees of freedom* | *Class III* | *Class IV* | *Total* |
|---|---|---|---|
| n-2 | 30.6<br>(19 / 62) | 64.0<br>(16 / 25) | 40.2<br>(35 / 87) |
| n-1 | 48.4<br>(30 / 62) | 96.2<br>(50 / 52) | 70.2<br>(80 / 114) |

The bracketed figures in Table 11.5 show the number of SKUs fitted by the binomial as a ratio of the number of SKUs with sufficient 'cells' to allow a chi-square goodness-of-fit test to proceed. The results for class IV show a high percentage of SKUs fitted by the binomial, a finding which is consistent with the 90.7% of these

SKUs (49 / 54) with inter-purchase times fitted by the geometric distribution. The results for class III show a lower percentage of SKUs with a good fit to the binomial, but this is inconsistent with the 85.2% of SKUs (46 / 54) with inter-purchase times fitted by the geometric. The inconsistency is most noticeable for the faster-moving SKUs within class III. For these SKUs, the occurrence of more than one demand incidence in some working days leads to a greater variability of demand than that predicted by the binomial distribution.

### 11.5.4 Condensed-Poisson demand incidence and Erlang-2 inter-purchase times

As discussed in chapter 10, the moments and maximum likelihood estimators do not coincide for the condensed-Poisson distribution. The goodness-of-fit results for both methods are summarised in Table 11.6 :

TABLE 11.6
Percentage of SKUs with demand incidence fitted by the condensed-Poisson distribution

| *Estimation Method* | *Class I* | *Class II* | *Class III* | *Class IV* | *Total* |
|---|---|---|---|---|---|
| Moments | 23.7 | 24.2 | 24.2 | 14.3 | 23.0 |
| ML (n-2 df) | 23.7 | 25.9 | 16.1 | 0.0 | 18.7 |
| ML (n-1 df) | 33.9 | 29.6 | 25.8 | 16.4 | 26.5 |

As for the Poisson distribution, all of the SKUs in classes I, II and III are included in the above calculations. For class IV, 21 SKUs are included in the moments analysis, 28 in the ML calculations with n-2 degrees of freedom and all SKUs in the analysis with n-1 degrees of freedom.

The results show a low percentage of SKUs fitted by the condensed-Poisson distribution and this finding is consistent across all movement classes. This result is also obtained for Erlang-2 inter-purchase times, with just 21.1% of SKUs being well-fitted. From the discussion in chapter 9, it may be expected that the condensed-Poisson would perform better for the faster moving SKUs, particularly in class I. However, it is clear that the Poisson distribution provides a good fit to more SKUs in all classes.

Maximum likelihood does not represent a significant improvement on moments estimation for the condensed-Poisson distribution. It yields a small improvement for classes I and II, but does not fit any SKUs in class IV. The reason for this is not clear, although it is possible that there are bias problems for the slowest moving SKUs. This is a matter which requires further research.

### 11.5.5 Comparison of demand incidence distributions

From the results in this section, clear empirical support has been obtained for the Poisson demand incidence model for all movement categories. The binomial distribution suffers from estimation problems for faster-moving items, although good fits are confirmed for the slowest movers. No support for the condensed-Poisson has been found.

## 11.6 Goodness-of-fit of demand distributions

### 11.6.1 Negative binomial distribution

Three methods of estimation for the negative binomial distribution (NBD) were discussed in chapter 10: moments, 'means and zeros' and maximum likelihood. This discussion was motivated by the improved goodness-of-fit of the NBD by using maximum likelihood in the replication of the empirical analysis of Banerjee and Bhattacharyya (1976).

The results of the goodness-of-fit analysis using each of the three estimation methods is summarised in Table 11.7 :

TABLE 11.7
Percentage of SKUs with demand fitted by the negative binomial distribution

| *Estimation Method* | *Class I* | *Class II* | *Class III* | *Class IV* | *Total* |
|---|---|---|---|---|---|
| Moments | 76.3<br>(45 / 59) | 72.2<br>(39 / 54) | 61.0<br>(36 / 59) | 48.0<br>(12 / 25) | 67.0<br>(132 / 197) |
| Means & zeros | 12.1<br>(4 / 33) | 51.0<br>(26 / 51) | 71.0<br>(44 / 62) | 65.2<br>(15 / 23) | 52.7<br>(89 / 169) |
| ML (n-3 df) | 84.7<br>(50 / 59) | 85.2<br>(36 / 54) | 67.7<br>(42 / 62) | 64.0<br>(16 / 25) | 77.0<br>(154 / 200) |
| ML (n-1 df) | 86.4<br>(51 / 59) | 90.7<br>(49 / 54) | 80.6<br>(50 / 62) | 94.2<br>(49 / 52) | 87.7<br>(199 / 227) |

The maximum likelihood results show a high percentage of SKUs well-fitted by the NBD. This finding is consistent with Kwan (1991) who also found the NBD to fit demand patterns well. The high percentages are particularly marked for classes I and

II, but good results are also obtained for classes III and IV.

As anticipated in chapter 10, the method of moments performs poorly for slow-moving SKUs and the 'means and zeros' method performs poorly for faster-moving SKUs. Maximum likelihood is better than both of these methods, particularly for class I and II.

### 11.6.2 Gamma distribution

It was noted in chapter 8 that the NBD is the discrete analogue of the gamma, so similar results may be anticipated. Two estimation methods were reviewed in chapter 10: moments and maximum likelihood. Results from the two methods are summarised in Table 11.8 :

TABLE 11.8
Percentage of SKUs with demand fitted by the gamma distribution

| *Estimation Method* | *Class I* | *Class II* | *Class III* | *Class IV* | *Total* |
|---|---|---|---|---|---|
| Moments | 78.0<br>(46 / 59) | 70.4<br>(38 / 54) | 54.2<br>(32 / 59) | 45.5<br>(10 / 22) | 64.9<br>(126 / 194) |
| ML<br>(n-3 df) | 81.4<br>(48 / 59) | 75.9<br>(41 / 54) | 10.3<br>(6 / 58) | 0.0<br>(0 / 13) | 51.6<br>(95 / 184) |
| ML<br>(n-1 df) | 89.8<br>(53 / 59) | 85.2<br>(46 / 54) | 14.5<br>(9 / 62) | 25.0<br>(12 / 48) | 53.8<br>(120 / 227) |

As expected, the maximum likelihood results for classes I and II are similar to the results for the NBD. ML gives a small but noticeable improvement in goodness-of-fit over the method of moments for these top two movement classes.

Maximum likelihood estimation is markedly inferior to the method of moments for SKUs in classes III and IV. This is not unexpected, as it was noted in chapter 10 that ML estimates are biased for small samples. The moments results for the gamma are comparable to moments results for the NBD but are not as good as ML results for the NBD.

### 11.6.3 Condensed negative binomial distribution

Maximum likelihood estimates have not been derived for the condensed negative binomial distribution (CNBD). The methods of moments and 'means and zeros' were reviewed in chapter 10. The goodness-of-fit results for these two estimation methods are given in Table 11.9 :

TABLE 11.9
Percentage of SKUs with demand fitted
by the condensed negative binomial distribution

| *Estimation Method* | *Class I* | *Class II* | *Class III* | *Class IV* | *Total* |
|---|---|---|---|---|---|
| Moments | 78.0<br>(46 / 59) | 68.5<br>(37 / 54) | 57.6<br>(34 / 59) | 44.0<br>(11 / 25) | 65.0<br>(128 / 197) |
| Means & zeros | 6.1<br>(2 / 33) | 45.1<br>(23 / 51) | 67.7<br>(42 / 62) | 48.1<br>(13 / 27) | 46.2<br>(80 / 173) |

These figures are very similar to the corresponding results from the NBD. This may seem surprising, since the Poisson demand incidence model out-performed the condensed-Poisson. However, Chatfield and Goodhardt (1973) obtained similar empirical results for consumer purchases. They concluded that the difference between the two fitted distributions, NBD and CNBD, is small.

### 11.6.4 Log-zero-Poisson distribution

Two estimation methods were discussed in chapter 10: factorial moments and maximum likelihood. Both methods have been used for the lzP distribution. The goodness-of-fit results are summarised in Table 11.10 :

TABLE 11.10
Percentage of SKUs with demand fitted by the log-zero-Poisson distribution

| *Estimation Method* | *Class I* | *Class II* | *Class III* | *Class IV* | *Total* |
|---|---|---|---|---|---|
| Factorial Moments | 0.0 (0 / 55) | 4.0 (2 / 50) | 8.6 (5 / 58) | 0.0 (0 / 4) | 4.2 (7 / 167) |
| ML (n-4 df) | 28.6 (2 / 7) | 61.9 (13 / 21) | 75.0 (18 / 24) | 100.0 (1 / 1) | 64.2 (34 / 53) |
| ML (n-1 df) | 57.1 (4 / 7) | 71.4 (15 / 21) | 91.7 (22 / 24) | 100.0 (23 / 23) | 85.3 (64 / 75) |

The factorial moments results are very poor for classes I, II and III. There are an insufficient number of SKUs in class IV which have sufficient degrees of freedom, after allowing for grouping and the three parameters to be estimated, to give results of any significance.

Kwan (1991) found the lzP to perform well for slow-moving SKUs. Seventy nine of the SKUs in her sample had average inter-purchase times of more than a week, corresponding to class IV items. Sixty four of these SKUs (86.1%) were well-fitted by the lzP distribution. However, only six SKUs in Kwan's sample had average inter-purchase times of less than a week. This sample is too small to give significant results.

The poor performance of factorial moments was discussed in chapter 10. Rao (1969) showed that the relative efficiency of the method is greater than 30% only if both the tau and P parameters are less than 1.5. None of the SKUs in classes I, II and III have estimated values of tau and P in this part of the parameter space.

Maximum likelihood estimates gave good results for those SKUs where such estimates were obtained. However, ML estimates were found for less than a quarter of the total sample. A grid search was conducted on the parameter tau from 0.1 to 5.1 times the value of the factorial moments estimator, with grid-points at every 0.2 times the estimator. The P parameter was treated identically and all combinations of tau and P were examined. Most SKUs had maxima at an edge of this grid, thereby providing no assurance of optimality. The results presented in Table 11.10 are for those SKUs for which a local maximum of the likelihood function was found.

It may be concluded, then, that improvements on the grid-search method are needed for the lzP to be applied to all SKUs. One possible enhancement is a more appropriate 'initial solution' than the factorial moments estimator. Khedr (1980) found that the use of factorial moments as an initial estimate in searching for the maximum likelihood estimate to be less computationally efficient than certain ad hoc estimators. A second enhancement would be a search routine based on steepest descent methods, rather than grid-search, to speed up the computations.

For those SKUs for which an ML estimator was obtained, the overall results are similar to the NBD. However, since it is not known whether the goodness-of-fit is representative of those SKUs whose ML estimators are yet to be found, these findings must be treated with some caution.

### 11.6.5 Comparison of demand distributions

The results in this section show that there are greater differences between methods of estimation (for the same distribution) than between distributions (for the same method of estimation). In this sub-section, the differences between estimation methods are quantified for some illustrative examples. Whereas in the previous sub-sections the differences in percentages of SKUs well-fitted by different estimation methods were shown, in this sub-section examples of the magnitude of improvement in goodness of fit, and the effect on parameters, are given. Although the examples are not untypical of the differences which may occur, it should be emphasised that for many SKUs there is little discernible difference between estimation methods. In this sub-section, attention is focussed on those SKUs which do show a difference. These SKUs will benefit most from the operational recommendations in section 11.7.

The NBD and gamma distributions both have easily obtainable ML estimates, unlike the CNBD and lzP. Although the negative binomial and gamma distributions, using ML estimation, give similar goodness-of-fit for faster-moving SKUs, the NBD provides a better fit for slow-moving items. The lzP distribution performs similarly to the NBD for those items for which ML estimates were obtained. However, the problem of finding the estimates without prohibitive computation requires further research.

Examining the NBD in more detail, it is apparent that 'means and zeros' estimation is inappropriate for class I SKUs. For example, for SKU 472514, a zero estimate is returned for the exponent, since no zero values were observed. Almost half of the SKUs in class I suffered from the same problem. For those items with some zero observations in class I, 'means and zeros' performed poorly. For example, for SKU 410007, the following results were obtained :

TABLE 11.11
Example of poor performance of 'means and zeros' in NBD estimation for fast moving SKUs

| | *Exponent* | *Chi-square (degrees of freedom)* |
|---|---|---|
| Means and zeros | 0.77 | 19.00 (6) |
| Moments | 1.50 | 14.00 (6) |
| Max Likelihood | 0.76 | 9.77 (6) |

Maximum likelihood estimation outperforms moments for slower classes too. However, for such items, there is often little difference between 'means and zeros' and maximum likelihood. For example, for SKU 417016, a class III item, the following results were obtained :

TABLE 11.12
Example of good performance of 'means and zeros' in NBD estimation for slower moving SKUs

| | *Exponent* | *Chi-square (degrees of freedom)* |
|---|---|---|
| Means and zeros | 0.74 | 10.31 (5) |
| Moments | 0.63 | 15.24 (6) |
| Max Likelihood | 1.76 | 10.29 (5) |

Turning to the gamma distribution, it is apparent that the method of maximum likelihood performs poorly for classes III and IV. For classes I and II, however, some improvement in goodness of fit is possible by using maximum likelihood. Three examples are given below which show the potential improvement and the difference in the estimates of the shape parameter :

TABLE 11.13
Example of good performance of maximum likelihood in Gamma estimation for fast moving SKUs

| *SKU* | *Class* | *Shape Moments* | *Shape ML* | *Chi-square Moments* | *Chi-square ML* |
|---|---|---|---|---|---|
| 472514 | I | 4.415 | 3.682 | 15.077 (8) | 10.536 (8) |
| 702053 | I | 1.453 | 2.262 | 18.886 (8) | 6.890 (8) |
| 1212100 | II | 0.572 | 1.222 | 9.943 (5) | 1.853 (5) |

where the bracketed figures indicate 'degrees of freedom'. In each example, and especially for the second and third SKUs, there is a wide difference in parameter estimates and goodness of fit, demonstrating the benefit of consideration of using maximum likelihood for class I and II items.

For the condensed negative binomial distribution, it is apparent that 'means and zeros' estimation is inappropriate for class I SKUs. Just as for the negative binomial, for SKU 472514, a zero estimate is returned for the exponent, since no zero values were observed. As noted earlier, a substantial number of the SKUs in class I suffered from this problem. For class II items, the method of moments outperforms 'means and zeros'. For example, for SKU 702174, the following results were obtained :

TABLE 11.14
Example of poor performance of 'means and zeros' in CNBD estimation for faster moving SKUs

| | *Exponent* | *Chi-square (degrees of freedom)* |
|---|---|---|
| Means and zeros | 0.80 | 14.83 (5) |
| Moments | 1.15 | 5.20 (5) |

Thus, a similar finding is obtained as for the NBD: the use of 'means and zeros' should be restricted to slower moving classes of inventory.

As far as the log-zero-Poisson distribution is concerned, it is clear that more work on estimation properties is required, particularly for inventory data. The estimators of Huque (1973) and Khedr (1980) require records for the number of occasions when two and three items were ordered within a defined time period. Such methods are not immediately applicable to many SKUs and further work on ad hoc estimators is needed since maximum likelihood estimators are computationally expensive.

## 11.7 Conclusions

In this chapter, the assumption of Poisson demand incidence has been tested on a sample of 230 SKUs from Pillar Engineering Supplies (PES) Ltd. The range of speed of movement of the SKUs in the sample extends from one every ten weeks to fifty per week. The Poisson distribution was found to perform consistently well over all movement categories.

The discrete analogue of the Poisson, the binomial distribution, also performed well for those SKUs for which it was possible to obtain an estimate of the 'probability of success' of under one. However, the condensed Poisson performed much worse than the Poisson, even for the fastest moving items.

Strong support has been provided for the negative binomial as a demand distribution. It provided a good fit to a high percentage of SKUs in all movement categories. The results also showed consistently better goodness-of-fit by using maximum likelihood estimation rather than the method of moments.

The gamma distribution, the continuous analogue of the NBD, showed good results, although not as many SKUs were well-fitted by the gamma as the NBD, particularly for slow-moving SKUs. The condensed negative binomial distribution also showed similar results to the NBD for corresponding methods of estimation. However, maximum likelihood estimates have not been derived for this distribution.

The log-zero-Poisson distribution performed very poorly using factorial moments estimates, showing that Kwan's good results using this method for slow-moving SKUs do not carry over to faster-moving SKUs. Substantial improvements in goodness-of-fit were obtained by using maximum likelihood estimation, but estimates were obtained for less than a quarter of the sampled SKUs. Further research is needed to find better computational methods.

For a practitioner wishing to decide on an estimation method for an operational system, the following summary of results may be useful :

* For NBD estimation, 'means and zeros' is inappropriate for class I items. For class II items, maximum likelihood outperforms 'means and zeros' in terms of goodness of fit. For classes III and IV, little difference is apparent.

* For the gamma distribution, a small but noticeable improvement in goodness of fit over the method of moments is afforded by maximum likelihood for classes I and II. For classes III and IV, maximum likelihood is not appropriate and the method of moments is preferable.

* For the condensed negative binomial distribution (CNBD), 'means and zeros' is not suitable for class I items. The method of moments is superior to 'means and zeros for class II. There is some evidence that 'means and zeros' is the better method for classes III and IV.

* For the log-zero Poisson distribution, factorial moments may be used for class IV items only. Stable estimation methods have not been found which are suitable for the whole range of inventory items within faster moving classes.

# PART IV

# DEMAND VARIANCE LAWS

## Summary of Part IV

In this part of the thesis, issues of operationalising inventory amalgamation models are examined, with particular reference to the estimation of demand variance. In chapter 12, the theory of variance laws is discussed. Models in the literature are found to be limited by their assumption of constant order-sizes. A more general formulation is proposed, based on compound Poisson demand, which takes into account non-constant order sizes and non-constant mean demand incidence. It is shown that, under certain conditions, there is a quadratic relationship between the variance and mean of demand. Two models are proposed. If mean order-size is independent of mean demand, then the relationship is exactly quadratic. If mean order-size depends on a power of mean demand, then another term must be added to the quadratic variance law to give a 'general variance law'.

In chapter 13, the conditions for a quadratic variance are tested using the sample of data from Pillar Engineering Supplies (PES) which was used in chapter 11 to test demand incidence and demand distributions. Each of the three assumptions required for both quadratic and general variance laws are found to hold strongly. The mean order-size is found to be weakly correlated with a power of the mean demand. Although the correlation is weak, it is found to be statistically significant. The general variance law is not supported by the data because of the weak correlation. Hence, the most appropriate model for PES is found to be a quadratic variance law.

In chapter 14, the relative merits of the quadratic and general variance laws are discussed in the context of centralisation modelling. The general variance law cannot carry over to a centralised setting. There is no reason why higher mean demands at the centralised depot, resulting from receipt of orders from more customers than at a decentralised depot, should lead to higher mean demands from customers. Therefore, a quadratic variance law is more appropriate. The parameters of the variance law may change as a result of centralisation. Methods for assessing this change are discussed, to enable the variance law approach to be useful in practice.

# CHAPTER 12

## *Variance Laws for Inventory Management*

### 12.1 Introduction

In chapter 2, it was shown that if the variance of demand is unequal at some decentralised depots, then estimates of demand variances are required in order to use the depot centralisation models proposed in the literature. However, the estimation of demand variance can be problematic. In the following situations, it may be difficult to obtain accurate demand variance estimates:

1. The SKU is near the beginning of its product life cycle and insufficient data is available to allow a reasonable estimate of demand variance to be made.

2. There has been a step-change in the mean demand level. In this case, the variance of demand expected over future periods is not reflected by past variance, which is inflated by the change in mean level.

3. The SKU is a slow-moving product but is not yet obsolete. There are few data points as the product moves so slowly and variance estimation will be subject to large errors.

In each of the above cases, a reasonably reliable mean estimate is easier to find than

---

*Note : The first four sections of this chapter are based on the article by Boylan and Johnston (1996), entitled 'Variance Laws for Inventory Management' (Int J Production Economics, Vol 45, pp 343-352).*

a variance estimate. If a variance-mean relationship is known, then the mean estimate may be readily translated into a variance estimate.

Jacobs and Wagner (1989) analysed the performance of a 'power law' variance-mean relationship, using regression-based estimators, compared with a sample variance estimator. They simulated a wide range of scenarios, monitoring the inventory systems costs (set-up costs, holding costs and penalty costs) of both methods in each scenario. They concluded, *"The results showed that when a family of items has the same variance-mean relation, a regression estimator can be far superior to an individual item sample variance estimator"*. Therefore, estimation benefits are not necessarily limited to the three situations outlined above.

In this chapter, a theory of functional relationships is developed for compound-Poisson demand. The robustness of functional relationships to departures from the assumptions underpinning the compound-Poisson demand model, discussed in chapters 8 and 9, are analysed.

## 12.2 Variance laws

### 12.2.1 Definition of a variance law

A variance law is a relationship linking the population mean ($\mu_i$) and the population variance ($\sigma_i^2$) of demand for the *i*th SKU of the form:

$$\sigma_i^2 = f(\mu_i) \qquad \text{for all i} \quad \text{(ie for all SKUs).}$$

The simplest relationship is of the form :

$$\sigma_i^2 = A\,\mu_i.$$

where A is a constant.

This would hold if demand were well represented by a Poisson distribution. If this were the case, safety stocks would be a small fraction of those found to be necessary in practice, since demand is almost always found to be more variable than suggested by the above equation.

The power law of the form :

$$\sigma_i^2 = A\,\mu_i^B$$

was suggested by Brown (1963).

A quadratic law :

$$\sigma_i^2 = A\,\mu_i + B\,\mu_i^2,$$

was proposed by Burgin and Wild (1967).

Sherbrooke (1992) suggested a model of the form :

$$\sigma_i^2 = \mu_i + A\,\mu_i^B.$$

Empirical evidence has been collected by researchers to support the power law (eg

Johnston et al (1988a) and Sherbrooke's law (Sherbrooke (1992)). However, no theoretical reasoning has been advanced to support either of these laws. The quadratic law, on the other hand, does have some theoretical support. This is reviewed in the next sub-section.

### 12.2.2 Stevens' joint demand model

Stevens (1974) concluded that a quadratic law is preferable to a power law since it may be derived from a 'joint demand model'. The joint demand model proposed by Stevens is of the form:

$$y_{it} = \mu_i + \varepsilon_{it} + \eta_{it}$$

| | | |
|---|---|---|
| where | $y_{it}$ | is the demand for the *i*th SKU at time t |
| | $\mu_i$ | is the mean demand for the *i*th SKU |
| | $\varepsilon_{it}$ | is the random component of demand attributable to the *i*th SKU alone |
| and | $\eta_{it}$ | $= p_i \eta_t$ |
| | $\eta_t$ | is the random component of total demand for the family of SKUs as a whole |
| | $p_i$ | is the proportion of $f_t$ attributable to the *i*th SKU. |

Thus, the $\eta_{it}$ component is the 'share' of the family effect $\eta_t$ which is allocated to the *i*th SKU. The difficulties associated with such a concept are discussed below, after an overview of the assumptions underpinning Stevens' model.

Stevens made a number of assumptions in his model :

1. The individual random element, $\varepsilon_{it}$, and the family random element, $f_t$, are independent of each other.

2. The two random elements, $\varepsilon_{it}$ and $f_t$, are independent of earlier values.

3. The expected values of $\varepsilon_{it}$ and $f_t$ are zero.

4. The variances of $\varepsilon_{it}$ and $f_t$ do not vary over time.

5. The individual random element, $\varepsilon_{it}$, follows a Poisson distribution.

6. The family effect, $f_t$, is attributable to each SKU in direct proportion to the SKU's mean demand. The distribution of $f_t$ is not specified.

The fifth assumption requires a correction. For non-zero $\varepsilon_{it}$ values, zero expectation (assumption 3) and a Poisson distribution (assumption 5) are incompatible. The assumption should be that $\mu_i + \varepsilon_{it}$ follows a Poisson distribution, thus allowing for deviations below the mean as well as above the mean.

It should be noted that $y_{it}$, according to Stevens' model, may be non-integer since the $p_i f_t$ term may be non-integer. This is a weakness in the formulation, particularly for slow-moving SKUs for which non-integer demand values may be a poor approximation.

On the basis of the joint demand model and its assumptions, given above, Stevens found a relationship of the following form :

$$\sigma_i^2 = \mathrm{Var}(y_{it}) = A\,\mu_i + B\,\mu_i^2$$

where A is the unit of sale (assumed constant)

B is the squared coefficient of variation of the common component of total demand.

If there is no joint effect, then the second term vanishes, leaving the standard linear variance-mean relationship for a Poisson distribution.

Stevens' joint demand model marked an advance by providing a derivation of a variance law from first principles. However, in addition to the weaknesses outlined above, the model is also limited by a constant unit of sale being assumed, whereas non-constant values are more commonly observed in practice. A re-formulation is needed to overcome these difficulties.

### 12.2.3 A re-formulation of the joint demand model

A re-formulation of the joint demand model is proposed which differs from Stevens' model in two respects. Firstly, disturbance factors are assumed to affect mean values, for consistency with the 'mixing model'. Secondly, the unit of sale, A, is introduced as a first step towards a 'compounding model'.

The model is split into two parts. The first equation relates the demand for the *i*th SKU, $y_{it}$, to the current mean level of demand incidence, $\lambda_{it}$. The second equation relates the current mean level of demand incidence to an underlying mean level. The re-formulation is as follows :

*Equation 1* $$y_{it} = A\,(\lambda_{it} + \varepsilon_{it})$$

*Equation 2* $$\lambda_{it} = \lambda_i + \eta_{it}$$

where $\lambda_{it}$ is the mean demand incidence, at time t, for the *i*th SKU

$\lambda_i$ is the underlying mean demand incidence for the *i*th SKU.

It is clear that the linear coefficient, A, in the quadratic law follows directly from 'Equation 1'. If A is integer, then the demand, $y_{it}$, must also be integer, thereby overcoming a drawback of Stevens' formulation. Moreover, the introduction of 'Equation 2' facilitates a comparison with other work on 'mixing models' of compound Poisson demand.

## 12.3 Quadratic variance law for compound Poisson demand

### 12.3.1 The stationary mean model : a generalisation of the re-formulated model

The re-formulated model of the previous section may be generalised to allow for non-constant order-sizes. The equations introduced in the previous section are generalised to :

*Equation 1* $\qquad y_{it} = \Sigma_{j=1}^{n_{it}} A_{ijt}(\lambda_{it})$

*Equation 2* $\qquad n_{it} = \lambda_{it} + \varepsilon_{it}$

*Equation 3* $\qquad \lambda_{it} = \lambda_i + \eta_{it}$

where

| | | |
|---|---|---|
| $A_{ijt}$ | is the size of the *j*th order in time period t for the *i*th SKU | |
| $n_{it}$ | is the number of orders in time period t for the *i*th SKU | |
| $\varepsilon_{it}$ ) | are stochastic | ( $\lambda_i + \varepsilon_{it}$ is Poisson distributed |
| $\eta_{it}$ ) | disturbance terms | ( $\eta_{it}$ distribution not specified. |

Whilst the non-constant order-sizes are expressed as generally as possible, the variation in mean level is based on the assumption of an underlying stationary mean. This model will therefore be referred to as a 'stationary mean' model. A further generalisation, based on a random walk of the mean level and known as a 'steady state' model, is explored in sub-sections 12.5.6 and 12.5.7. The assumption of negative exponential inter-demand times is retained. The sensitivity to some deviations from this assumption is examined in section 12.4.

## 12.3.2 Variance law and the effect of dependence of mean order-size on mean order-incidence

Suppose that the assumptions of the previous sub-section are maintained, namely : compound Poisson demand and Poisson demand incidence with a time varying mean subject to a general probability distribution. Then, if the mean demand incidence is treated as a continuous variable over time, rather than as a variable subject to discrete shocks at given points in time, the expressions for the mean and variance of demand over time t, are as follows :

$$c_1 = \int b_1(\lambda)\, \lambda t \, dS(\lambda)$$

$$c_2 = \int [\, b_2(\lambda) + b_1(\lambda)^2 \,]\, \lambda t \, dS(\lambda) \;+\; \mathrm{Var}\,[\, b_1(\lambda)\, \lambda t \,]$$

where $\lambda$ is mean demand incidence (varying over time)
S is the cumulative distribution function of $\lambda$
$a_1$, $a_2$ are the mean and variance of mean demand incidence
$b_1$, $b_2$ are the mean and variance of order-sizes
$c_1$, $c_2$ are the mean and variance of demand.

The result given above is expressed quite generally. An important special case is when the mean order-size is independent of mean demand incidence. Since there is no *a priori* reason why these two variables should be related, the assumption is a reasonable working hypothesis. In the next chapter, the hypothesis will be tested using the data from PES described in chapter 11. If mean order-size and mean demand incidence are independent, then $b_1(\lambda) = b_1$ and the equation for the mean demand simplifies to :

$$c_1 = b_1 \int \lambda t \, dS(\lambda) \;=\; b_1\, a_1\, t$$

and the equation for the variance simplifies to :

$$c_2 = [\, b_2 + b_1^2 \,] \int \lambda t \, dS(\lambda) + b_1^2 \; \mathrm{Var}[\, \lambda t \,]$$

$$= [\, b_2 + b_1^2 \,] \, a_1 \, t + b_1^2 \, a_2 \, t^2$$

After a little algebraic substitution, the following relationship is obtained :

$$c_2 = \left( \frac{b_2}{b_1^2} + 1 \right) b_1 \, c_1 + \frac{a_2}{a_1^2} c_1^2$$

Thus, a relationship between the variance and mean of demand has been derived from first principles. The relationship may be described as a 'quadratic law', in the sense of section 12.2, if three conditions are satisfied :

1. Mean order-size ($b_1$) is independent of mean demand ($c_1$)

2. Order-size variance to mean-squared ratio ($b_2/b_1^2$) is independent of mean demand

3. Distribution of mean order-incidence over time, using the variance to mean squared measure ($a_2/a_1^2$), is independent of mean demand.

Two special cases of the quadratic variance law illustrate its application. If demand incidence is Poisson and there is no variation in the mean level ($a_2 = 0$), then demand is compound Poisson distributed and the variance is given by :

$$c_2 = c_1 \left( \frac{b_2}{b_1} + b_1 \right) .$$

If demand incidence is Poisson and mean demand is gamma distributed ($a_2/a_1^2 = 1/k$), then the variance is given by:

$$c_2 = c_1 \left( \frac{b_2}{b_1} + b_1 \right) + \frac{c_1^2}{k} .$$

### 12.3.3 Analogous results from insurance mathematics

The results given in the previous sub-section are new to inventory modelling but are not new to applied mathematics. Indeed, the results were motivated by work in the field of insurance mathematics.

In insurance mathematics, it is often assumed that individual policyholders have accident claims which form, for each individual, a Poisson process, with an individual mean claim rate. A function, U(x), called the 'risk distribution', is used to denote the cumulative distribution of mean numbers of claims ('risks') across a group of policyholders. Each claim has a value and the distribution function, V(y), gives the probability that the value of the claim will not exceed y. This is known as the 'collective risk' model.

In demand modelling, it may be assumed that each customer arrives randomly, each with their own individual arrival rate. Customers' mean arrival rates may be described by the (cumulative) distribution function U(x). At each customer arrival, an order is placed and the distribution function V(y) gives the probability that the size of order will not exceed y. This analogy is summarised in Table 12.1 :

TABLE 12.1
Analogy between insurance claims and demand for inventory

| *Insurance Claims* | *Inventory Demand* |
|---|---|
| Policyholder | Customer |
| Claim made | Order placed |
| Size of claim | Size of order |

Although the analogy is strong, it is important to note that both models are across customers, not over time. However, it is possible to re-interpret the U distribution. If U represents the distribution function of mean claim-rates over time (rather than across customers), then there is a *direct* analogy between insurance mathematics and inventory modelling. Moreover, since the form of the results is not changed by the interpretation, insurance mathematics results on the distribution of risks across customers can be carried over to the modelling of inventory demand over time.

Lundberg (1964) analysed the 'collective risk' model and found the relationships between the moments of the total claims and the moments of the claim arrival rates and claim sizes. Using the same notation as in the previous two sub-sections (but denoting moments about the origin by a dashed superscript), Lundberg showed that:

$$c'_1 = t\, a'_1\, b'_1$$

$$c'_2 = t\, a'_1\, b'_2 + t^2\, a'_2\, b'^2_1$$

From these results, a relationship between the *moments about the mean* (same notation as above but without the dashes) may be obtained. The relationship is identical to that derived in the previous section, namely :

$$c_2 = \left( \frac{b_2}{b_1^2} + 1 \right) b_1\, c_1 + \frac{a_2}{a_1^2}\, c_1^2$$

Thus, the same relationship applies to both insurance mathematics and inventory modelling, since the assumptions are of identical mathematical form and allow analogous interpretations in the two fields of study.

## 12.4 Condensed-Poisson demand incidence

### 12.4.1 General relationship between variance and mean of demand

In chapter 9, the Erlang-2 distribution was examined as an alternative to the negative exponential. In this section, the 'stationary mean' model is examined under the assumption of Erlang-2 inter demand times (or, equivalently, condensed Poisson demand incidence). The sensitivity to this deviation from a Poisson process is investigated. No assumptions are made about the distribution of mean demand level.

A condensed-Poisson process is obtained by considering every second event in a Poisson process. If the mean rate of demand incidence in the original process is $2\lambda$, then the mean rate of demand incidence in the censored process is $\lambda$. The expressions for the mean and variance of demand over time t are given below :

$$c_1 = \int b_1(\lambda)\, \lambda t \, dS(\lambda)$$

$$c_2 = \int \left( b_2(\lambda) + \frac{b_1(\lambda)^2}{2} \right) \lambda t \, dS(\lambda) + \frac{1}{4} \int b_1(\lambda)^2 \, e^{-2\lambda t} \sinh(2\lambda t) \, dS(\lambda)$$

$$+ \operatorname{Var}[\, b_1(\lambda)\, 2\lambda t \,]$$

where the notation is unchanged from section 12.3.

The proofs of these results may be derived using the methods given in Appendix 12.1. The results simplify if the mean order-size is independent of the mean demand incidence, a situation which is examined in the next sub-section.

## 12.4.2 Variance law if mean order-size independent of mean demand incidence

If mean order-size and mean demand incidence are independent, then $b_1(\lambda) = b_1$ and the mean and variance equations simplify to :

$$c_1 = b_1 \int \lambda t \, dS(\lambda)$$

$$= b_1 a_1 t$$

$$c_2 = \left( b_2 + \frac{b_1^2}{2} \right) \int \lambda t \, dS(\lambda) + \frac{b_1^2}{4} \int e^{-2\lambda t} \sinh(2\lambda t) \, dS(\lambda)$$

$$+ b_1^2 \operatorname{Var}[\lambda t]$$

$$= \left( b_2 + \frac{b_1^2}{2} \right) a_1 t + \frac{b_1^2}{4} \int e^{-2\lambda t} \sinh(2\lambda t) \, dS(\lambda) + b_1^2 a_2 t^2$$

After some algebraic re-arrangements, the following relationship is obtained :

$$c_2 = \left( \frac{b_2}{b_1^2} + \frac{1}{2} \right) b_1 c_1 + \frac{b_1^2}{4} \int e^{-2\lambda t} \sinh(2\lambda t) \, dS(\lambda) + \frac{a_2}{a_1^2} c_1^2$$

It is immediately apparent that the middle term in the above expression for variance is likely to preclude the quadratic variance law for most distributions of mean demands. In the special case of the Gamma distribution, the results are as follows :

$$c_1 = b_1 \frac{\gamma}{\alpha} t$$

$$c_2 = \left( b_2 + \frac{b_1^2}{2} \right) \frac{k}{\alpha} t + \frac{b_1^2}{8} \left\{ 1 - \left( 1 + \frac{4t}{\alpha} \right)^{-k} \right\} + b_1^2 \frac{k}{\alpha^2} t^2$$

where $k$ and $\alpha$ are the shape and scale parameters of the gamma distribution.

Proofs are given in Appendix 12.1.

### 12.4.3 Comparison with Chatfield and Goodhardt's formulation

For the special case of unit orders with no variability of order-sizes ($b_1 = 1$, $b_2 = 0$), a little re-arrangement of the variance expression gives :

$$\text{Var} = \frac{c_1}{2} + \frac{1}{8}\left\{1 - \left(1 + \frac{4c_1}{k}\right)^{-k}\right\} + \frac{c_1^2}{k}$$

This is a similar, but not identical, formulation to Chatfield and Goodhardt (1973), who found the expression for variance of demand across customers and expressed the final term as $c_1/\alpha$. The above formulation is preferable for both variance across customers and variance over time since, as the mean demand ($c_1$) increases, it is likely to be more strongly linked with increases in the scale parameter ($\alpha$) than the shape parameter (k). In practice, some analysis of empirical data is required to check this assertion. At once, it may be seen that an approximate quadratic variance-to-mean relationship holds if two conditions apply :

1. The shape parameter (k) is independent of the mean ($c_1$).
2. The middle term in the above expression is negligible.

It is apparent that, as the mean demand values increase, so the middle term becomes less important.

## 12.5 General variance laws

### 12.5.1 Order-size assumptions

In section 12.3, a quadratic variance law was derived based on mean order-size being independent of mean demand. A model based on demand-dependent order-sizes may be obtained under the following set of assumptions :

1. Mean order-size is independent of mean demand incidence (but not of mean demand).
2. Poisson stream of demand incidence.
3. Mean demand incidence is stationary, fluctuating about a constant underlying mean.

Under the third assumption, a 'stationary mean' model applies. Relaxation to a 'steady state' model is examined later in this section.

Earlier in this chapter, the following variance law was derived :

$$c_2 = \left( \frac{b_2}{b_1^2} + 1 \right) b_1 \, c_1 + \frac{a_2}{a_1^2} c_1^2$$

If order-size is not assumed independent of demand, then an understanding of two relationships are required in order to establish the variance of demand :

1. The relationship between mean order-size, $b_1$, and mean demand, $c_1$.
2. The relationship between the variance to mean squared ratio of order-size, $b_2/b_1^2$, and the mean demand.

These relationships are examined in the next two sub-sections.

### 12.5.2 Relationship between mean order-size and mean demand

Just as an organisation may use some form of Economic Order Quantity (EOQ) formula to determine their order-sizes on suppliers, so also their customers may use an EOQ approach in determining the size of their orders. This leads to higher order-sizes for those SKUs with higher demand. In the case of a single customer, the order-size is proportional to the square root of demand. Since customers must have a minimum order-size of unity, this leads to the following relationship :

$$b_1 = 1 + \alpha c_1^{1/2} .$$

This relationship no longer holds when an SKU has more than one customer (as is the case for all Pillar Engineering's A and B category stock). However, consideration of this functional form has motivated the following postulated relationship :

$$b_1 = 1 + \alpha c_1^{\beta} .$$

The parameters $\alpha$ and $\beta$ may be estimated empirically, using historical data on mean demands and order-sizes from a sample of SKUs.

### 12.5.3 Relationship between variance and mean of order-size

Although it is conceivable that those SKUs with higher mean order-sizes may have similar order-size variances to SKUs with lower mean order-sizes, it is not likely to be observed in practice. If an SKU is ordered singly and also has some order-sizes of a hundred, then it will almost certainly have a higher order-size variance than another SKU which has a maximum order-size of ten.

It is difficult to specify a functional relationship between the variance and mean order-size which will apply for all SKUs. However, the following distribution of order-sizes is very commonly observed:

* A large number of small orders (singles or small pack-sizes).
* A relatively small number of large orders (hundreds, thousands, etc).

Burgin (1975) showed how the gamma distribution may be used as a robust distribution for such skewed data. For a gamma distribution of shape parameter k :

$$\frac{\text{Variance}}{\text{Mean}^2} = \frac{1}{k}$$

Since a similar degree of skewness of order-sizes is often observed for many SKUs in the range, a stronger assumption may be made that the same shape parameter (k) applies to the order-sizes of all SKUs. If this assumption is made, then :

$$\frac{b_2}{b_1^2} = \gamma \text{ (constant).}$$

Naturally, this assumption must be tested in practice, as it may be untenable for some sets of SKUs.

## 12.5.4 General variance law based on 'stationary mean' model

The relationships postulated in the previous two sub-sections are taken as the basis for the derivation of a variance law in this sub-section. In section 12.3, the following result was derived under the assumption of a stationary mean :

$$c_2 = \left( \frac{b_2}{b_1^2} + 1 \right) b_1 c_1 + \frac{a_2}{a_1^2} c_1^2$$

Substituting the order-size relationship postulated in sub-section 12.5.2 :

$$b_1 = 1 + \alpha c_1^{\beta}$$

and the variance to mean squared order-size relationship postulated in sub-section 12.5.3:

$$\frac{b_2}{b_1^2} = \gamma \text{ (constant).}$$

gives the following result :

$$c_2 = (1 + \gamma) c_1 + \alpha (1 + \gamma) c_1^{1+\beta} + \delta c_1^2$$

where $\delta = a_2 / a_1^2$ is assumed independent of $c_1$ (final condition of section 12.3).

This result may also be written in the following form :

$$c_2 = A c_1 + B c_1^C + D c_1^2$$

This variance law is termed 'general' as it contains the quadratic law, the power law and Sherbrooke's law as special cases.

### 12.5.5 Comparison of the general variance law with Sherbrooke's law

Sherbrooke's variance law is of the form :

$$c_2 = c_1 + B\, c_1^{C}$$

A significant difference between Sherbrooke's law and the general variance law is in the method of derivation. Sherbrooke (1992) gave empirical evidence from SKUs held by the US Air Force to support his proposed variance relationship. He states that "*another empirical analysis ... provided more evidence that demand over short periods of time does follow a Poisson process. The larger variance-to-mean ratios observed over longer periods of time arise because the mean changes*". Sherbrooke gave two examples of these 'larger variance-to-mean ratios': in one study, the power C was found to be 1.58 and in another it was found to be 1.50.

The general variance law, on the other hand, has been derived from first principles. The derivation clarifies a number of issues :

1. A relationship such as Sherbrooke's may be derived for *stationary* data if there is a functional relationship between the mean order-size and mean demand of the type postulated.

2. The additional term due to a varying mean level is of the form : $(a_2 / a_1^2)\, c_1^2$. The addition of a term of the form $c_1^{C}$ for a varying mean will not occur in general but only if there is a relationship of the form : $(a_2 / a_1^2) = A\, c_1^{C-2}$.

12.5.6 Relaxation of stationary mean assumption - the steady state model

The final assumption of section 12.3 is that the mean demand may vary about a constant underlying mean level with the measure $a_2 / a_1^2$ independent of mean demand. This assumption is relaxed by using a steady state-model which permits a random walk for the mean demand incidence. The formulation is given below :

*Equation 1* $$y_{it} = \Sigma_{j=1}^{n_{it}} A_{ijt}$$

*Equation 2* $$n_{it} = \lambda_{it} + \varepsilon_{it}$$

*Equation 3* $$\lambda_{it} = \lambda_{it-1} + \eta_{it}$$

The implications of the steady state model, proposed by Harrison (1967), for the properties of demand variance over different lead-times has been investigated by Johnston and Harrison (1986).

12.5.7 Variance law based on 'steady state' model

Johnston and Harrison analysed the variance of demand for a steady state model of the form :

*Equation 1* $$y_{it} = \mu_{it} + v_{it}$$

*Equation 2* $$\mu_{it} = \mu_{it-1} + w_{it}$$

The variance of demand over t periods of time from such a model is given by :

$$c_2 = t\,V + \frac{t\,(t+1)\,(2t+1)}{6}\,W$$

where $V = \mathrm{var}(v_{it})$, $W = \mathrm{var}(w_{it})$.

In Johnston and Harrison's original equation, there was an additional term for the estimation error of the mean. Since this analysis is concerned with relationships between population parameters, the additional term is ignored.

The second term in the above variance formula replaces the expression $t^2 W$ which would apply for a stationary mean model. The additional complexity arises from the auto-covariance between demand in periods 1 ... t in the steady state model. (If a stationary mean model holds, then there is no auto-covariance). However, for one period of time (t = 1), the same result applies for both the steady state and stationary mean models, namely :

$$c_2 = V + W$$

From previous analyses, for compound Poisson demand :

$$V = \left( \frac{b_2}{b_1^2} + 1 \right) b_1 c_1$$

and :

$$W = b_1^2 a_2$$

Hence, the same variance law applies for one period of time :

$$c_2 = \left( \frac{b_2}{b_1^2} + 1 \right) b_1 c_1 + \frac{a_2}{a_1^2} c_1^2$$

Analysis of the more general case of the variance of compound Poisson demand over t periods of time, with auto-covariance, is a subject for future research.

## 12.6 Conclusions

The starting point for this chapter was Stevens' joint demand model, which provided a basis upon which a quadratic variance law may be derived. However, it was demonstrated that the joint demand model has a number of limitations. It was shown how a quadratic variance law may be derived under the more general conditions of compound Poisson demand with time-varying mean and non-constant order-sizes. The derivation is qualified by three assumptions: firstly, that mean order-size is independent of mean demand; secondly, that the order-size variance to mean-squared ratio is independent of mean demand; finally, that the expected demand variance to mean-squared ratio is also independent of mean demand.

The assumption of a Poisson demand stream was relaxed and the effect of an Erlang-2 process, giving condensed Poisson demand incidence, was investigated. Exact relationships between the variance and mean of demand were found. These exact relationships are not of quadratic form. However, if it is assumed that the shape parameter is independent of the mean, the relationships are *approximately* quadratic. The approximation improves as the mean demand increases.

The two assumptions of demand-independent order-sizes were also examined. Under stationary demand, it was shown that a general variance law results of the form : $c_2 = A\,c_1 + B\,c_1^C + D\,c_1^2$, where $c_1$ and $c_2$ represent the mean and variance of demand and A, B, C and D are parameters to be estimated. This expression was compared with the formula of Sherbrooke (1992). Although Sherbrooke's formula had strong empirical support, it was shown that the term $B\,c_1^C$ is not necessarily dependent on a time-varying mean, as Sherbrooke supposed.

Finally, the assumption that the mean varies about a constant underlying value was replaced by the assumption of a steady state model. Under these conditions, it was shown that the same variance expression applies for one period of time, but a more complex expression is required for the variance of demand over more than one period of time. This is a topic for future research.

# CHAPTER 13

## *Empirical Analysis of Variance Laws*

### 13.1 Introduction

In chapter 12, it was shown how a stationary mean model and a one-step ahead steady state model lead to a quadratic variance law of the form:

$$\sigma^2 = A\mu + B\mu^2$$

assuming that mean order-size is independent of mean demand. The linear term is due to an underlying Poisson process of demand incidence. The quadratic term is due to variation in the mean demand level over time.

If mean order-size ($b_1$) is related to mean demand ($\mu$) : $b_1 = 1 + \alpha\,\mu^\beta$, then the law changes to a more general form :

$$\sigma^2 = A\mu + B\mu^C + D\mu^2 .$$

Three conditions are necessary for both the quadratic and general variance laws :

* The ratio of the variance of mean demand incidence to the square of the expected mean demand incidence is independent of mean demand.
* Mean order-size is independent of mean demand incidence.
* The ratio of the variance of order-size to the square of mean order-size is independent of mean demand,

In this chapter, these three conditions are tested empirically using the data from Pillar

Engineering Supplies (PES) described in chapter 11. The second and third conditions may be tested straightforwardly; the results are reported in sections 13.6 and 13.7. The first condition is difficult to test directly, since estimates of the variance of mean demand incidence are required. However, if the first condition holds, then it follows that the following relationship for demand incidence applies :

$$\sigma^2 = \mu + A\mu^2$$

where $\sigma^2$ is the variance of demand incidence, rather than the variance of mean demand incidence. This relationship is tested in section 13.5.

The relationship between mean order-size and mean demand, which discriminates between the quadratic and general laws, is tested in section 13.8.

Finally, results are reported for the overall relationship between the mean and variance of demand in section 13.9.

## 13.2 Mean and variance estimation methods

In order to estimate the parameters in the general variance laws and the supporting relationships, estimates of the mean and variance of demand incidence, order-size and demand are required for each SKU.

### 13.2.1 Estimation of mean demand incidence

The underlying demand incidence model may be expressed in two forms :

*Stationary Mean Model*

$$n_t = \lambda_t + \varepsilon_t$$
$$\lambda_t = \lambda + \eta_t$$

*Steady State Model*

$$n_t = \lambda_t + \varepsilon_t$$
$$\lambda_t = \lambda_{t-1} + \eta_t$$

using the same notation as in chapter 12.

For the stationary mean model, older data and newer data should be given equal weight in estimating the underlying mean demand incidence. Therefore, the sample mean, over a period in which the stationarity assumption is not violated, will be used for this model.

Harrison (1967) showed that the optimal predictor for the steady state model is the exponentially weighted moving average (EWMA) of past observations, with a smoothing parameter, $\alpha$, given by :

$$\alpha = \frac{[(1 + 4R)^{1/2} - 1]}{2R}$$

where $$R = Var(\varepsilon_t) / Var(\eta_t) \quad .$$

It is assumed that all error terms have zero expected values, are not serially correlated and are not correlated with each other. No distributional assumptions are made about the errors.

### 13.2.2 Estimation of the standard deviation of demand incidence

An appropriate estimate of standard deviation for the stationary mean model, in the absence of extreme outliers, is the unbiased sample estimate.

Two estimators have been proposed for the steady state model. Brown (1959, 1963) suggested the use of the smoothed mean absolute deviation, scaled by a factor of 1.25 (assuming a normal distribution) as a suitable robust estimator of the standard deviation. Bretschneider (1986) showed that this estimate may be out-performed by the smoothed mean square error (MSE), even in the presence of some outliers.

In the Pillar Engineering Supplies data, demand incidence outliers are quite infrequent and not extremely severe. Therefore, Bretschneider's suggestion of using the smoothed MSE will be followed.

### 13.2.3 Estimation of mean order-size

In the previous chapter, a model for order-sizes was not specified. However, a stationary mean model was implicit, since no allowance was made for the auto-covariance between order-sizes which would result from a steady state model. The assumption of a stationary mean is a natural one, since it may be expected that the mean order-size would change very slowly over time. It should be checked in practice.

If the mean is stationary, then a sample mean should be used as an estimate, for the reasons given in sub-section 13.2.1.

### 13.2.4 Estimation of the standard deviation of order-size

The distribution of order-sizes is subject to some extremely high observations. This makes the estimation of standard deviation problematic, since the estimate may be dominated by just one data point. To overcome this difficulty, the mean absolute deviation is used and converted to a standard deviation estimate. The estimate will be approximate, since the conversion factor of 1.25 is based on a normal distribution of order-sizes, whereas highly skewed distributions are observed in practice. Since a stationary mean model is assumed, the calculation of the MAD is not smoothed, but based on equal weightings of all observations.

### 13.2.5 Estimation of mean demand

The same considerations apply as for mean demand incidence.

### 13.2.6 Estimation of standard deviation of demand

Since demand variation is affected by the high variability in order-sizes, a MAD measure is to be preferred to a measure based on squared errors. Therefore, if demand incidence follows a stationary mean model, then the overall (equally weighted) MAD should be used as the basis for the estimate of the standard deviation of demand. If, on the other hand, demand incidence follows a steady state model, then the MAD measure should be smoothed over time to reflect the changing levels in mean demand incidence.

## 13.3 Identification of the most accurate estimation methods

### 13.3.1 Indication of underlying model

In the previous section, the most appropriate mean and variance estimation methods for 'stationary mean' and 'steady state' models were identified for demand incidence (henceforth called order-incidence in this chapter), order size and demand itself. In practice, however, the underlying model must first be specified. Since a mean average estimate is more appropriate for a 'stationary mean', and an EWMA is more appropriate for a 'steady state', it is possible to gain an indication of which model is better by the relative performance of the two measures. When the model is thus specified, the most suitable method of variance estimation may be identified using the arguments of the previous section.

### 13.3.2 Testing of one-step ahead forecast accuracy

The method of EWMA was tested for each of the three statistics (order-incidence, order-size and demand). In each case, four smoothing constants were used: 0.05, 0.1, 0.15 and 0.2. The squared error was calculated for each SKU and for each week. A comparison of the total squared errors for each of the statistics is given in Table 13.1:

TABLE 13.1
Total squared errors (EWMA comparison)

| | *Smoothing constant* | | | |
|---|---|---|---|---|
| | *0.05* | *0.10* | *0.15* | *0.20* |
| Order-incidence | 97503 | 94417 | 94075 | 94953 |
| Order-size | 96340770 | 97756260 | 99168740 | 100714900 |
| Demand | 122929200 | 119798300 | 120228600 | 121769500 |

The above table shows that the best smoothing constant of the four analysed for order-incidence is 0.15, indicating some drift in the mean level over time. The best smoothing constant for order-size, on the other hand, is 0.05. This indicates a much slower movement of the mean level. The overall demand is influenced by both order-incidence and order-size and its best smoothing constant is between the other two, at 0.10, indicating some drift in the mean level.

If a stationary mean model is appropriate, then it may be expected that a moving average (MA) calculation will give a more accurate estimate of the mean. The previous analysis of EWMA has been repeated for 8, 12 and 16 period moving averages. The results are given in Table 13.2 :

TABLE 13.2
Total squared errors (MA comparison)

| | *Number of periods* | | |
|---|---|---|---|
| | *8* | *12* | *16* |
| Order-incidence | 95625 | 96739 | 97771 |
| Order-size | 99180010 | 95657240 | 93626700 |
| Demand | 127188000 | 123265800 | 121744200 |

The above table shows that the moving average forecasts do not perform as well as the best EWMA forecasts for order-incidence and demand, confirming that a 'steady state' model may be more suitable for these two statistics. Order-size is best forecast by a 16-point moving average, confirming that a 'stationary mean' model is more appropriate for this statistic.

### 13.3.3 Identification of estimation methods

Based on the findings of the previous sub-section that a 'stationary mean' model is suitable for order-sizes and a 'steady state' model for demand and demand incidence, and the arguments of section 13.2, the following estimation methods are recommended, as shown in Table 13.3 :

TABLE 13.3
Recommended estimation methods

| | *Mean* | *Variance* |
|---|---|---|
| Order-incidence | EWMA | Smoothed MSE |
| Order-size | Sample mean | MAD |
| Demand | EWMA | Smoothed MAD |

For each SKU and each statistic, a one-step ahead forecast will be made using the methods identified above.

## 13.4 Heteroscedasticity in variance law estimation

Stevens (1974) illustrated the problems of heteroscedasticity in the estimation of a quadratic variance relationship. He showed that if heteroscedasticity is not taken into account, then inaccurate and even absurd estimates may result. Stevens recommended weighting each variance observation by the reciprocal of the squared predictor to overcome the problem of heteroscedasticity. On introducing this weighting, the parameters must be estimated iteratively but Stevens found the estimates to converge rapidly.

Stevens specifies the model for fitting as :

$$s_i^2 = A\,\mu_i + B\,\mu_i^2 + \psi_i$$

where $s_i^2$ is the variance estimate
$\psi_i$ is the error term.

Since it is assumed that the variance estimate is unbiased, it follows that the expected value of the error term is zero. However, Stevens does not specify the distribution of the error term; nor does he specify the method of calculation of the variance estimate. Nevertheless, he claims that $\mathrm{Var}(s^2) / E(s^2)^2$, the coefficient of variation of the estimates, takes a constant value.

If n values are used in the standard sample variance estimate, then it is known (see Kendall and Stuart (1979)) that :

$$n \frac{\mathrm{Var}(s^2)}{E(s^2)^2} = \frac{\mu_4}{\mu_2^2} - 1$$

where $\mu_2$ is the population second moment (about the mean)
$\mu_4$ is the population fourth moment (about the mean)
$\mu_4/\mu_2^2$ is the population kurtosis.

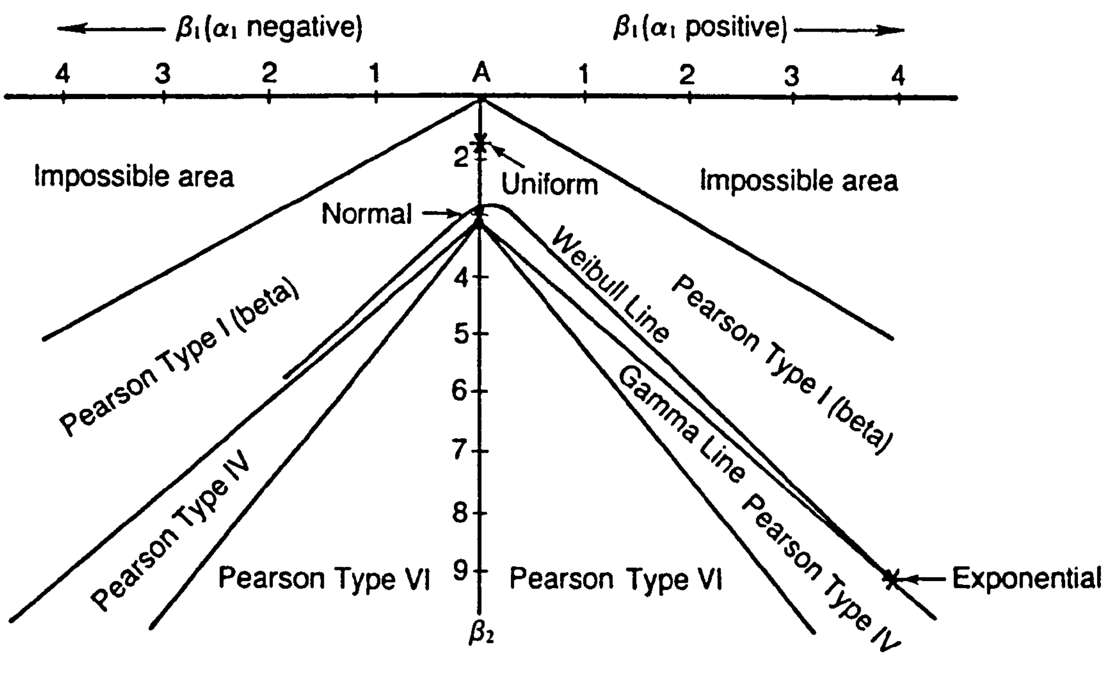
Since, for the normal distribution, kurtosis takes a constant value of three,

$$\frac{Var(s^2)}{E(s^2)^2} = \frac{2}{n}$$

Hence, if it is assumed that the sample size, n, is the same for all estimates, then the coefficient of variation of the estimates is constant.

However, as Figure 13.1, taken form Lau (1989), shows, some distributions do not have constant kurtosis values :

Figure 13.1
Kurtosis ($\beta_2$) - skewness ($\beta_1$) diagram



This argument shows that Stevens' result is not as general as was claimed. However, if a zero expected error is assumed, then a normal distribution of errors is a natural choice. For the empirical analysis of variance laws, it will be assumed that the coefficient of variation of all variance estimates, by all methods of estimation, are constant and Stevens' approach will be used.

The above discussion is necessary in order to demonstrate that the supposition of normal errors is, indeed, an *assumption*. Further research is required to examine the effect of relaxing the assumption to allow non-constant coefficients of variation on the parameter estimates in the variance laws under examination.

## 13.5 Estimation of the relationship between variance and mean order-incidence

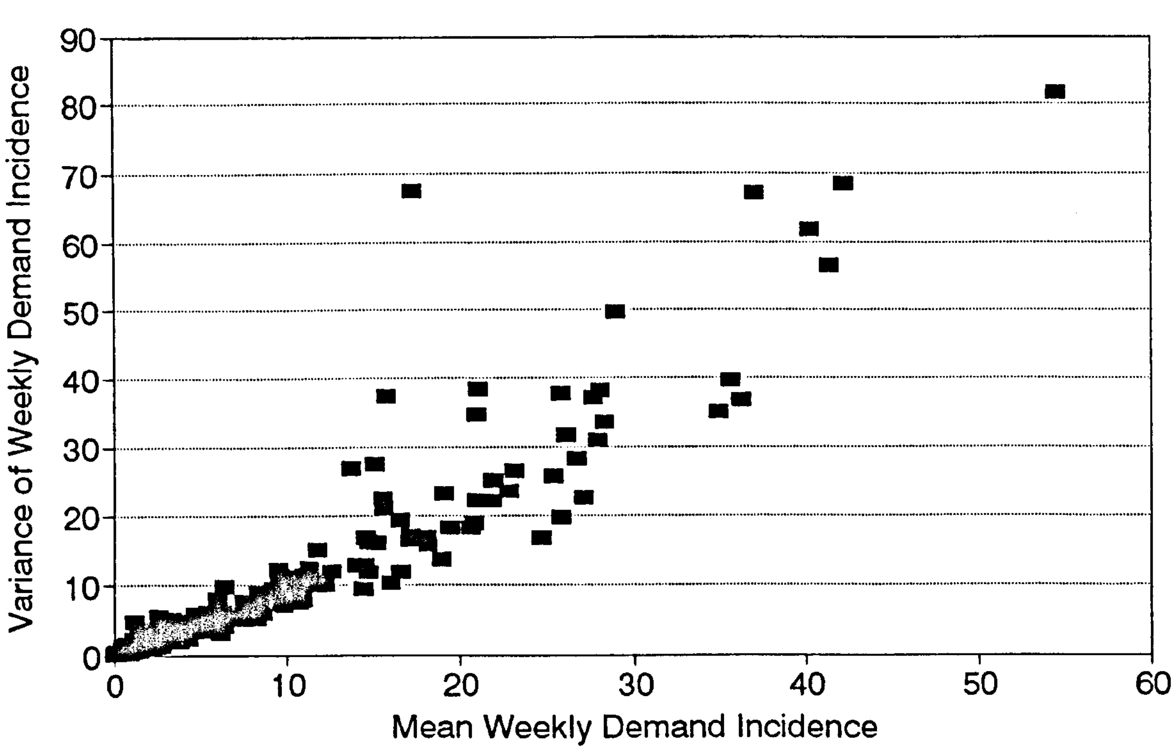
### 13.5.1 Ordinary least squares estimation

According to the analysis of chapter 12, if all orders are of a constant size of unity, then negative exponential inter-demand time leads to the following relationship for both the 'stationary mean' and one-step ahead 'steady state' model :

$$\text{Variance of Incidence} = \text{Mean Incidence} + B \ (\text{Mean Incidence})^2$$

A scatter graph of Pillar Engineering data is shown in Figure 13.2 :

Figure 13.2
Mean and variance of order-incidence for PES data



In this case, it is difficult to distinguish between a linear and a quadratic relationship. A goodness-of-fit analysis was performed for each of the following relationships :

| | |
|---|---|
| *Linear* | Variance = A * Mean |
| *Pure Quadratic* | Variance = B * $(Mean)^2$ |
| *Mixed Quadratic* | Variance = A * Mean + B * $(Mean)^2$ |
| *Mixed Quadratic (A=1)* | Variance = Mean + B * $(Mean)^2$ |

The $R^2$ values for the four relationships are given in Table 13.4 :

TABLE 13.4
Order-incidence variance - mean $R^2$ measure

| | *R - squared* |
|---|---|
| Linear relationship | 85.8 % |
| Pure quadratic relationship | 75.8 % |
| Mixed quadratic relationship | 87.6 % |
| Mixed quadratic (A = 1) | 87.5 % |

The table indicates that most of the variability in the variance-mean relationship may be explained by the linear model. The squared term in the 'mixed quadratic relationship' adds a little extra explanatory power; this is barely affected by forcing the linear parameter to unity, as required by the theory developed in chapter 12.

A more formal argument demonstrates that the contribution to improvement in $R^2$ from the quadratic term, although small, is statistically significant. The ratio of the reduction in squared errors (by adding a quadratic term) to the remaining squared errors (after the quadratic term is added), both divided by their relevant degrees of freedom, is the following F-statistic:

$$F = \frac{(0.876 - 0.858) / 1}{(1 - 0.876) / 227}$$

$$= 32.95$$

This statistic is highly significant for the relevant degrees of freedom, thereby demonstrating the significance of the contribution of the quadratic term.

The parameters and their standard errors, for each of the four postulated relationships are shown in Table 13.5 :

TABLE 13.5
Order-incidence variance - mean parameter estimates

| | *Linear Term* | | *Quadratic Term* | |
|---|---|---|---|---|
| | Parameter Estimate | Standard Error | Parameter Estimate | Standard Error |
| Linear | 1.2406 | 0.0272 | | |
| Pure Quadratic | | | 0.0384 | 0.0011 |
| Mixed Quadratic | 0.9166 | 0.0622 | 0.0115 | 0.0020 |
| Mixed Quadratic (A=1) | 1.0000 | | 0.0090 | 0.0008 |

Table 13.5 shows that :

1. The quadratic parameter in both quadratic models is significantly higher than zero, confirming that the quadratic term is significant.

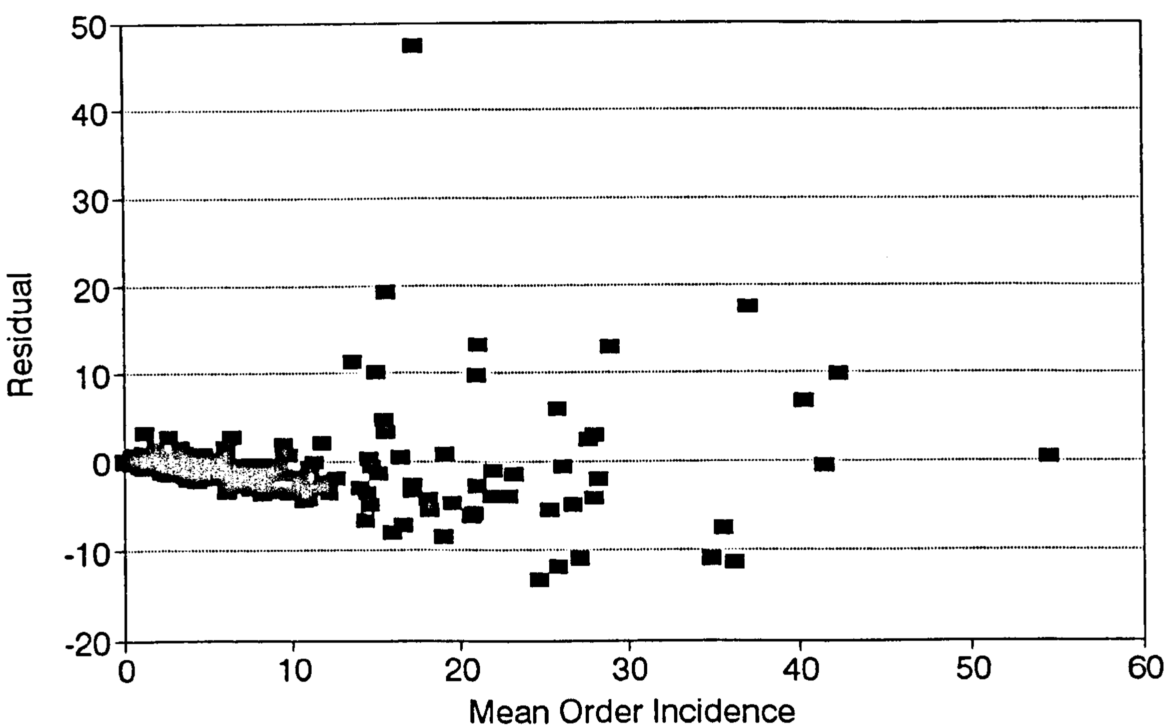
2. The coefficient of A = 1.2406 has no natural interpretation in the linear model. For demand incidence, a value of A = 1 would be expected for negative-exponential inter-arrivals and a value of A < 1 would be expected for more regular arrival patterns.

3. When the linear term is not forced to unity in the mixed quadratic relationship, the estimate is not significantly different from unity. This gives some support to the model postulated in chapter 12.

13.5.2 Weighted least squares estimation

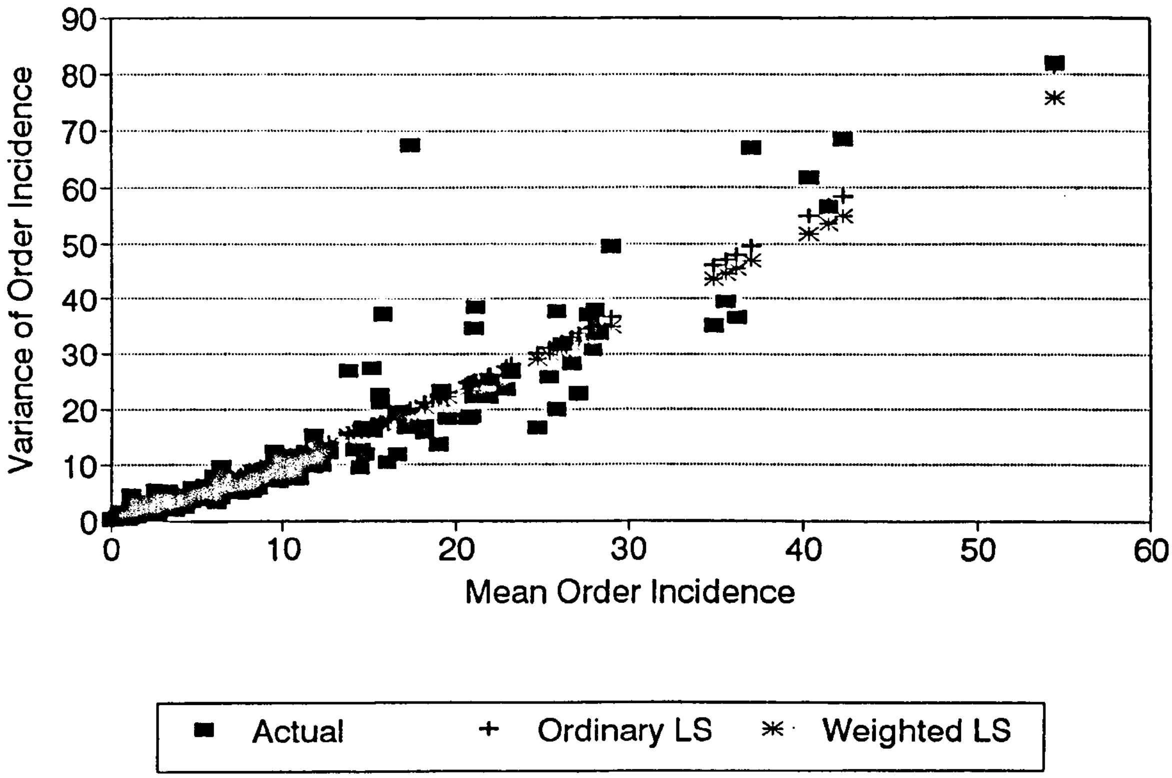
The residuals from the mixed quadratic relationship with the linear parameter forced to unity are shown in Figure 13.3 :

Figure 13.3
Residuals from order-incidence variance - mean relationship (OLS)



The graph shows evidence of heteroscedasticity. The relationship was re-estimated using weighted least squares, using the squared predictor as the weighting factor, as suggested by Stevens (1974). The effect of weighting is to reduce the quadratic parameter from 0.0090 to 0.0071. The change in the fit is shown in Figure 13.4 :

Figure 13.4
Order-incidence variance - mean relationship (WLS)




The weighted least squares (WLS) fit is a little less influenced by the SKUs with highest variance, as would be expected. The demand incidence variance-mean relationship derived from WLS is shown below :

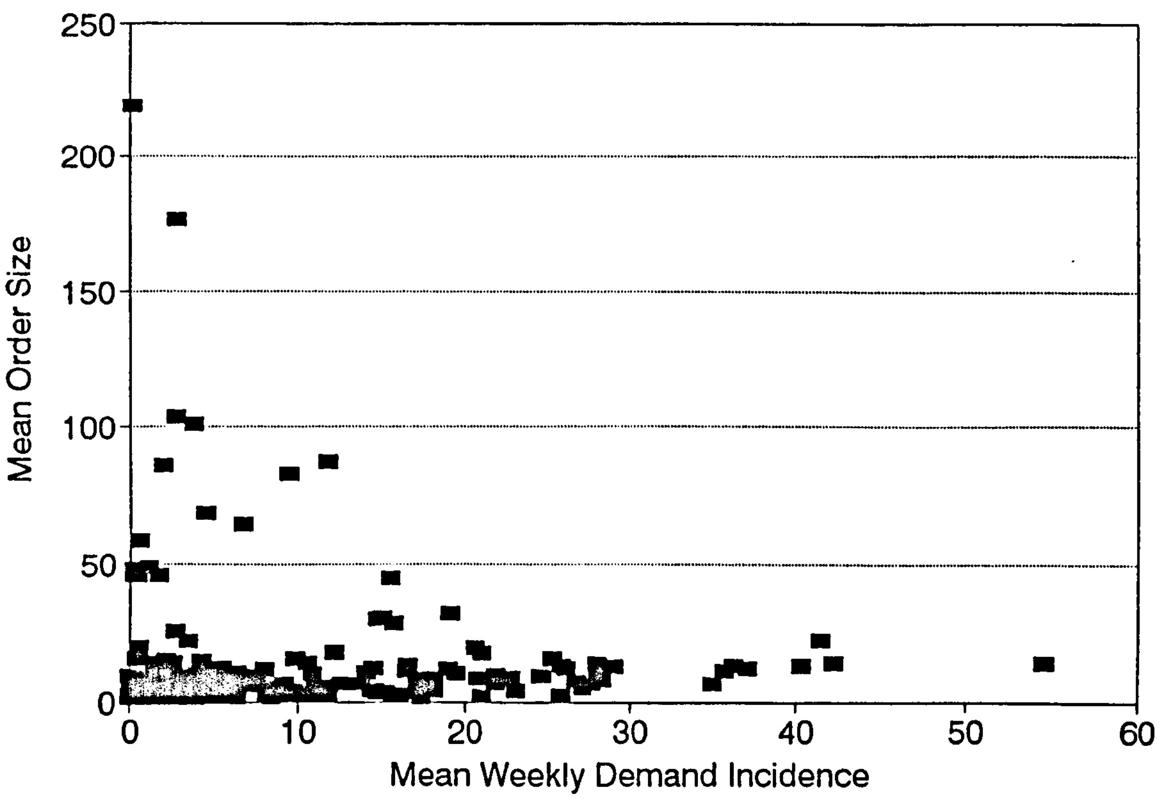
$$\text{Variance} = \text{Mean} + 0.0071 * (\text{Mean})^2$$

This relationship is consistent with the theoretical development of the previous chapter and takes into account the heteroscedasticity identified in this data.

## 13.6 Testing the assumption of independent mean order-size and mean order-incidence

In chapter 12, it was necessary to assume that mean order-size and mean order-incidence are independent in order to establish the 'variance law' quadratic and general functional relationships. A scatter-graph of these two variables for the Pillar Engineering SKUs is shown in Figure 13.5 :

Figure 13.5
Mean order-size and mean order-incidence for PES data




It seems that there is little correlation between the two variables. An analysis of the goodness of fit for the linear relationship :

$$\text{Mean order-size} = A + B\,(\text{Mean order-incidence})$$

and the inverse relationship :

$$\text{Mean order-size} = A + B \frac{1}{\text{Mean order-incidence}}$$

gives the following results, summarised in Table 13.6 :

TABLE 13.6
Independence of mean order-size and mean order-incidence

| | *R - squared* |
|---|---|
| Linear relationship | 0.017 % |
| Inverse relationship | 0.171 % |

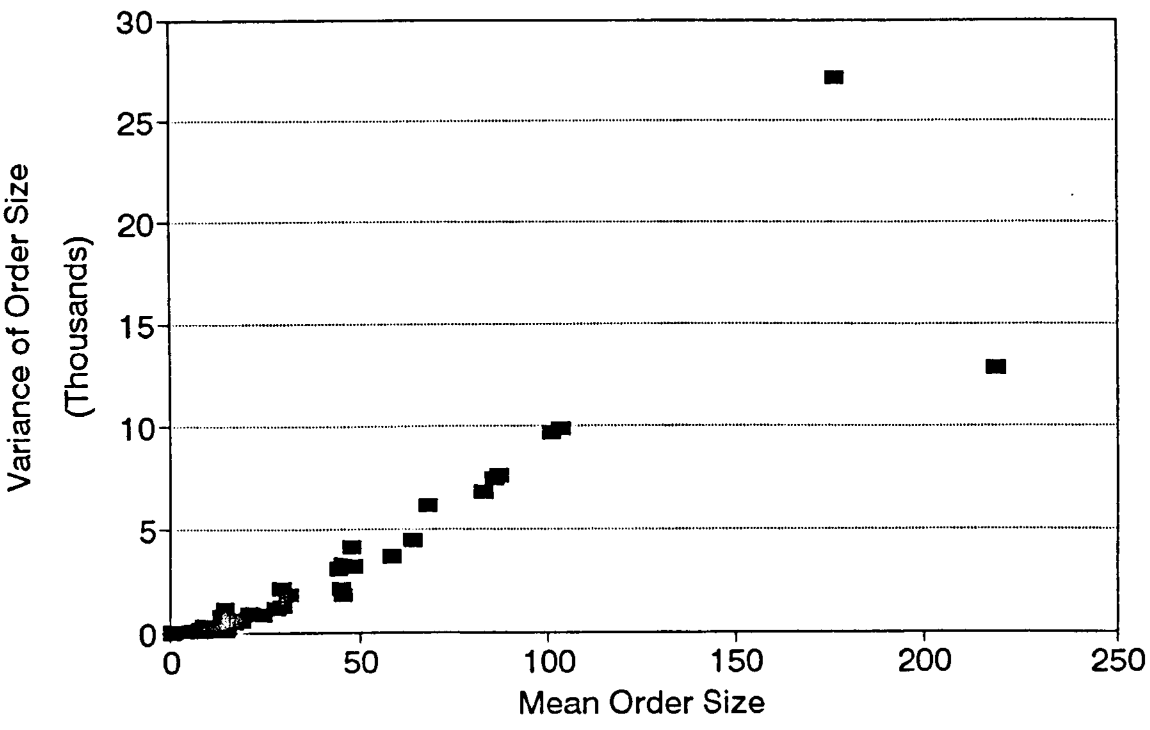
In both cases, there is no correlation between order-size and order-incidence and it may be concluded that the assumption of independence is supported by the evidence. Since the fit is so poor, the A and B parameters are not reported; for both models, the B parameter was not significantly different from zero (5% significance level).

## 13.7 Estimation of the relationship between variance and mean order-size

### 13.7.1 Ordinary least squares estimation

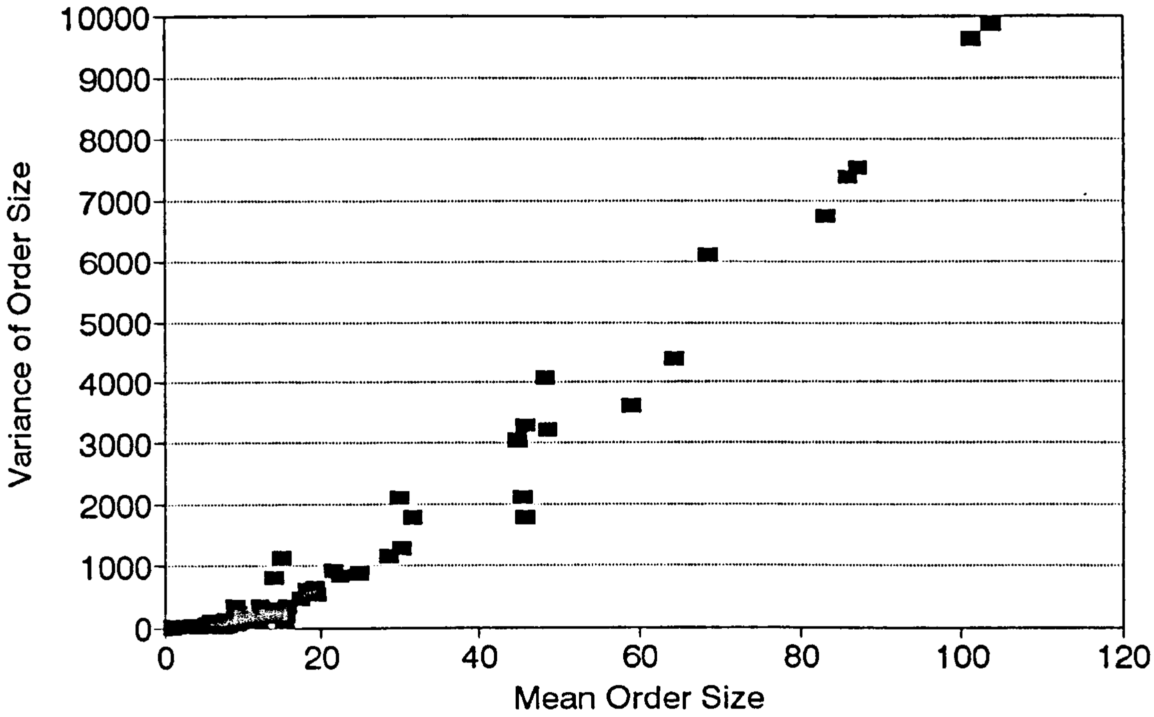
In the previous chapter, it was assumed that the variance to mean-squared ratio of order-sizes is constant. The scatter-chart shown in Figure 13.6 indicates the nature of the observed relationship between the variance and mean of order-size.

Figure 13.6
Mean and variance of order-size for PES data




The relationship is difficult to assess by eye because of the two observations with high mean order-sizes. If these two observations are removed, then the picture is somewhat clearer, as shown in Figure 13.7 :

Figure 13.7
Mean and variance of order-size for PES data
(two outlying observations removed)




The scatter-chart seems to indicate a non-linear relationship between the two variables. To examine this matter further, two relationships were compared, namely a linear relationship :

Variance of order-size = A * Mean order-size and a quadratic relationship of the form :

$$\text{Variance of order-size} = A * (\text{Mean order-size})^2$$

A summary analysis of the $R^2$ measure is shown in Table 13.7 :

TABLE 13.7
Order-size variance - mean $R^2$ measure

| | *All data* | *With two highest mean order-sizes excluded* |
|---|---|---|
| Linear relationship | 81.0 % | 87.5 % |
| Quadratic relationship | 70.2 % | 96.9 % |

The table shows the strong influence exerted by the two observations with the highest mean order-sizes. This illustrates the difficulty in using ordinary least squares when the data is heteroscedastic. With the two high observations excluded, it is clear that the quadratic relationship is preferable to the linear. The influence of the two high observations may also be seen from Table 13.8, which shows the effect on estimation of the parameter, A :

TABLE 13.8
Order-size variance - mean parameter estimates

| | *All data* | | *With two highest mean order-sizes excluded* | |
|---|---|---|---|---|
| | Parameter Estimate | Standard Error | Parameter Estimate | Standard Error |
| Linear Relationship | 83.724 | 2.585 | 72.114 | 1.721 |
| Quadratic relationship | 0.517 | 0.021 | 1.012 | 0.011 |

Table 13.8 shows a marked reduction in the standard errors when SKUs with the two highest mean values are excluded, since the excluded observations have outlying variances, as shown in Figure 13.6.

## 13.7.2 Weighted least squares estimation

In the previous sub-section, linear and quadratic relationships between the variance and mean of order-size were estimated using ordinary least squares. However, the residuals (differences between variance and fitted variance) may be expected to increase as mean order-size increases. This expectation is confirmed for both linear and quadratic models, as shown in figures 13.8 and 13.9. The two SKUs with highest mean order-sizes are excluded from these graphs, for clarity of exposition.

Figure 13.8
Residuals from order-size variance - mean linear relationship

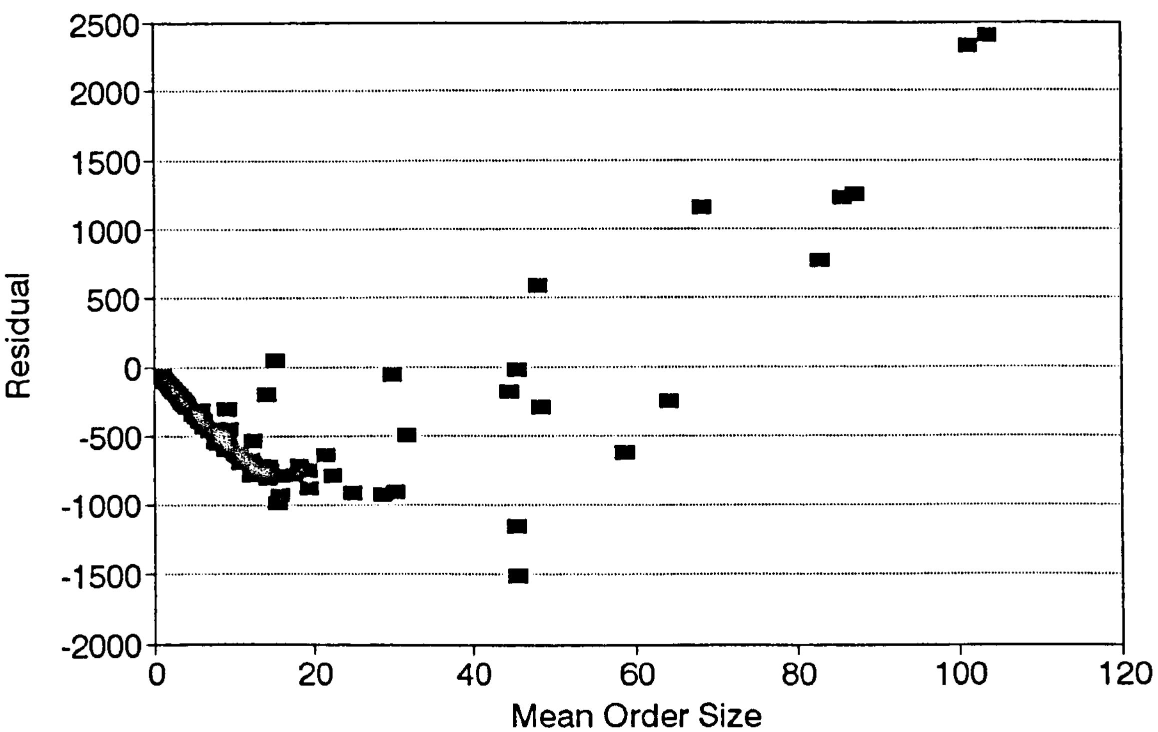
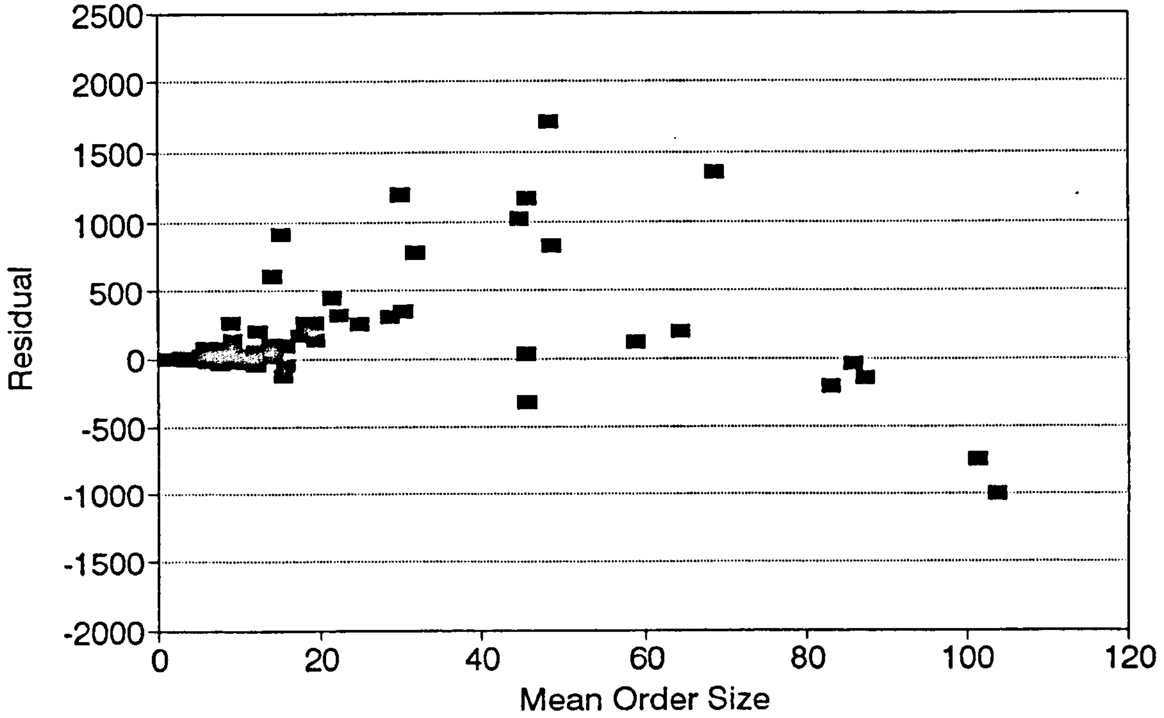


Figure 13.9
Residuals from order-size variance - mean quadratic relationship




Both of the figures reveal heteroscedasticity. This requires a similar adjustment to that proposed by Stevens for the demand variance - mean relationship. If it is assumed that the errors in the fitting are normally distributed, then the reciprocal of the square of the fitted variance may be used as a suitable weighting factor.

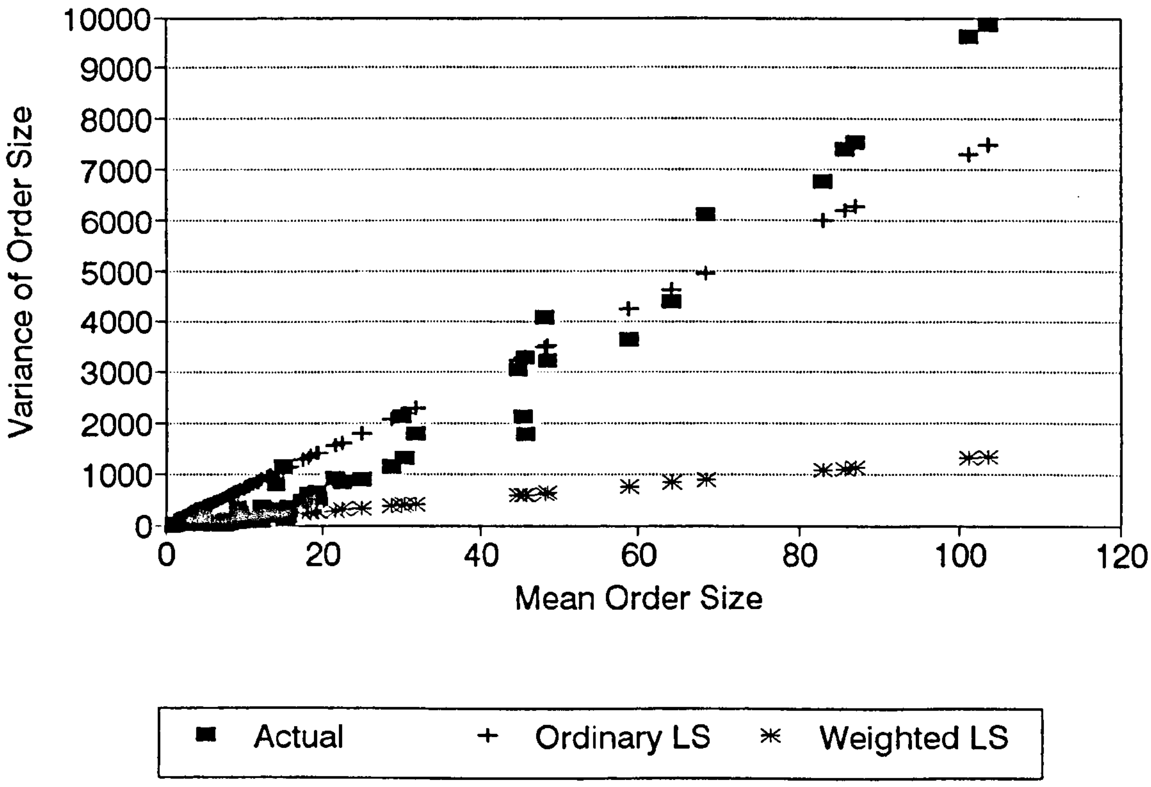
The estimates derived by weighted regression are shown in Table 13.9, alongside the estimates obtained by ordinary regression, to aid comparison :

TABLE 13.9
Order-size variance - mean parameter estimates
(weighted least squares)

| | *Ordinary least squares* | | *Weighted least squares* |
|---|---|---|---|
| | *All data* | *Excluding two highest means* | *All data* |
| Linear | 83.724 | 72.114 | 12.754 |
| Quadratic | 0.517 | 1.012 | 0.941 |

The linear relationships using OLS on all data excepting the two highest means and WLS on all data, are compared below in Figure 13.10 :

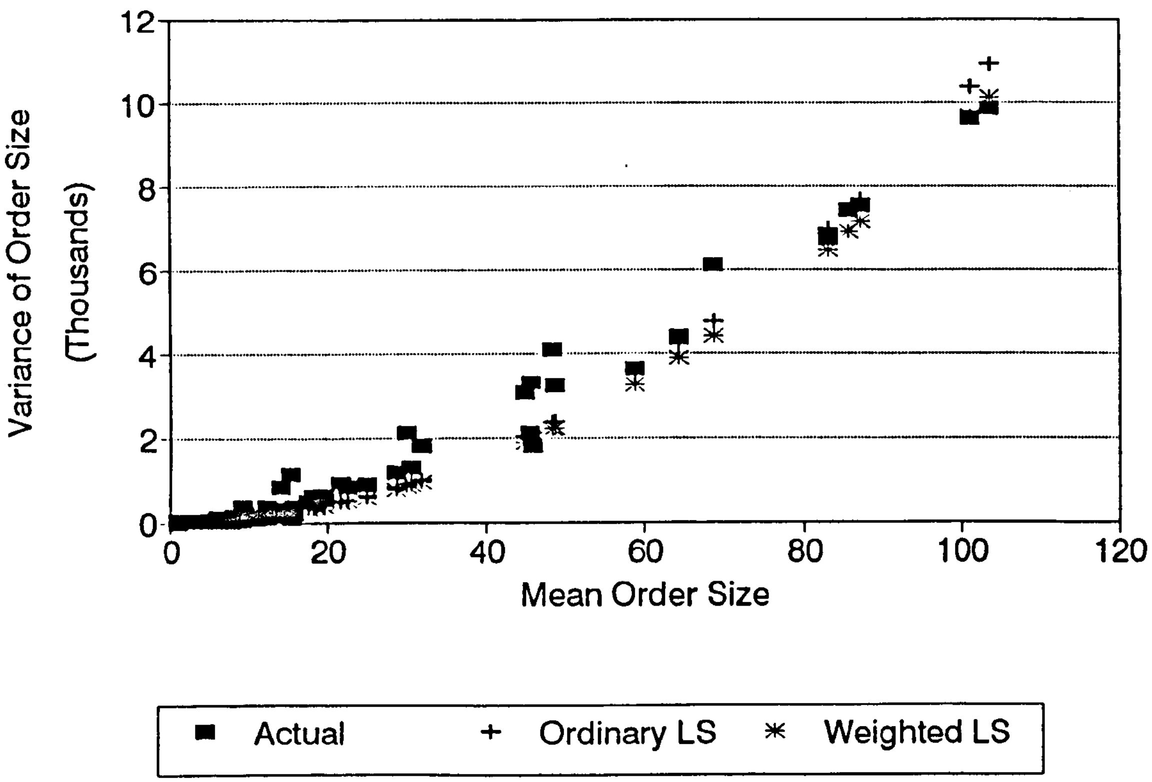
Figure 13.10
Order-size variance - mean linear relationships



The graph shows the straight line based on ordinary least squares to perform well on the high mean order-sizes, but to be consistently too high for the majority of SKUs with mean order-sizes under forty. This is also shown clearly in the residual plot of Figure 13.8. Conversely, the straight line based on weighted least squares performs well for SKUs with small mean order sizes but poorly for high mean order-sizes.

The quadratic relationships, using OLS on all data excepting the two highest means, and WLS on all data, are compared below in Figure 13.11 :

Figure 13.11
Order-size variance - mean quadratic relationships




In this case, there is very little difference between the two lines of fit. The line based on weighted least squares is not quite so heavily influenced by the SKUs with highest order-size variances than the OLS line. This results in lower residuals for high mean order-sizes, thus combatting the effects of heteroscedasticity.

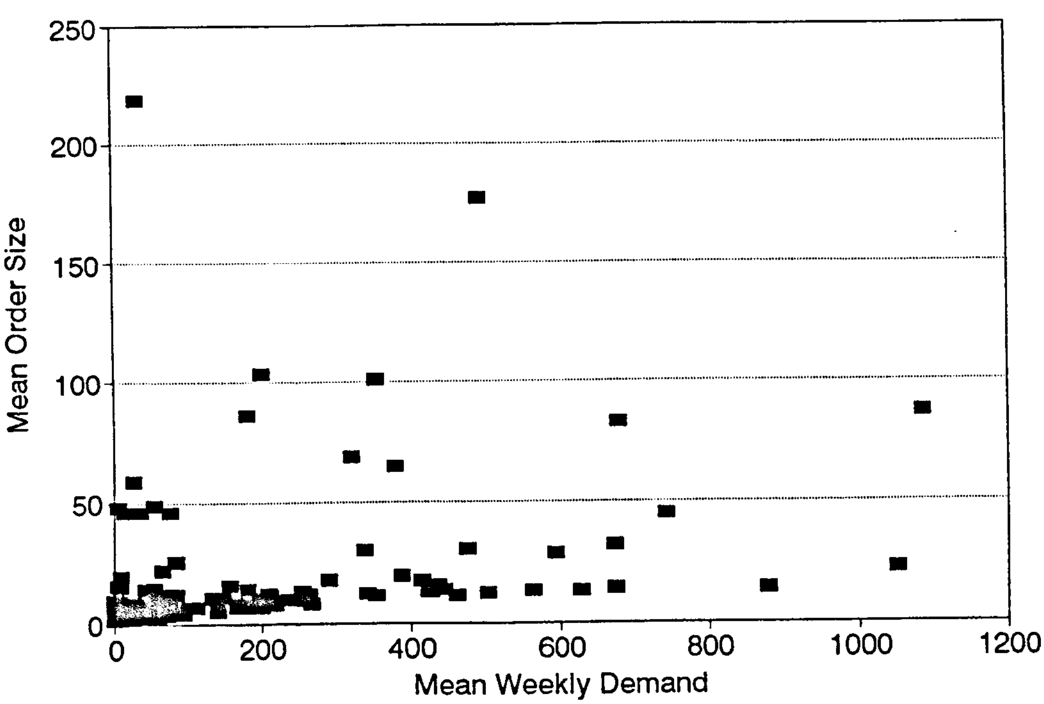
The above graphs confirm the conclusion of the ordinary least squares analysis, with outliers excluded, that the fit of the quadratic relationship is better than the linear relationship.

## 13.8 Estimation of the relationship between mean order-size and mean demand

### 13.8.1 Ordinary least squares estimation

A relationship of the form : Mean order-size = 1 + B * (Mean demand)$^{C}$ was postulated in the previous chapter. It was noted that this relationship is necessary for a 'general variance law' linking the variance of demand to its mean. A 'quadratic variance law', on the other hand, assumes that mean order-size and mean demand are independent. A scatter-chart of PES data is shown in Figure 13.12 :

Figure 13.12
Mean order-size and mean demand for PES data




The scatter-chart shows that any functional relationship between these two variables will have little predictive power. A range of relationships was examined for goodness of fit. The following relationship was estimated :

$$\text{Mean order-size} = A + B * (\text{Mean demand})^C$$

The A parameter is not constrained to take the value of unity, as postulated in the previous chapter. Powers of C = 0.3, 0.4, 0.5, 0.6 and 0.7 were examined. The following $R^2$ results were obtained, allowing the constant parameter to deviate from unity (as stipulated above), as shown in Table 13.10 :

TABLE 13.10
Mean order-size - mean demand $R^2$ measure

| *Power (C)* | *R - squared* |
|---|---|
| 0.3 | 14.5 % |
| 0.4 | 17.2 % |
| 0.5 | 17.3 % |
| 0.6 | 17.2 % |
| 0.7 | 16.9 % |

The $R^2$ values are very low for each of the powers analysed. However, the corresponding correlation coefficients are sufficiently high to be deemed statistically significant. Hence, it may be concluded that there is some evidence of a relationship, although the relationship is very weak.

A more detailed investigation of the strongest fitting relationship (C = 0.5) reveals the following results, shown in Table 13.11 :

TABLE 13.11
Mean order-size - mean demand parameter estimates

| | *Parameter estimate* | *Standard error* |
|---|---|---|
| *A* | 1.753 | 1.445 |
| *B* | 1.394 | 0.202 |

The B parameter is significantly different from zero, confirming the conclusion that a relationship between the variables does exist. Also, the A parameter is not significantly different from the theoretical value of unity. However, the A parameter must be positive and therefore, it is of some concern that the estimate is not significantly higher than zero. This indicates that the relationship is not strong enough to permit accurate estimation of the A parameter. Consequently, it will be constrained to unity, in line with its theoretical value, to ensure that all predicted order-sizes are greater than or equal to one.

If the relationship is re-estimated, constraining the A parameter to unity, then the following result is obtained :

TABLE 13.12
Mean order-size - mean demand parameter estimates
(constant term constrained to unity)

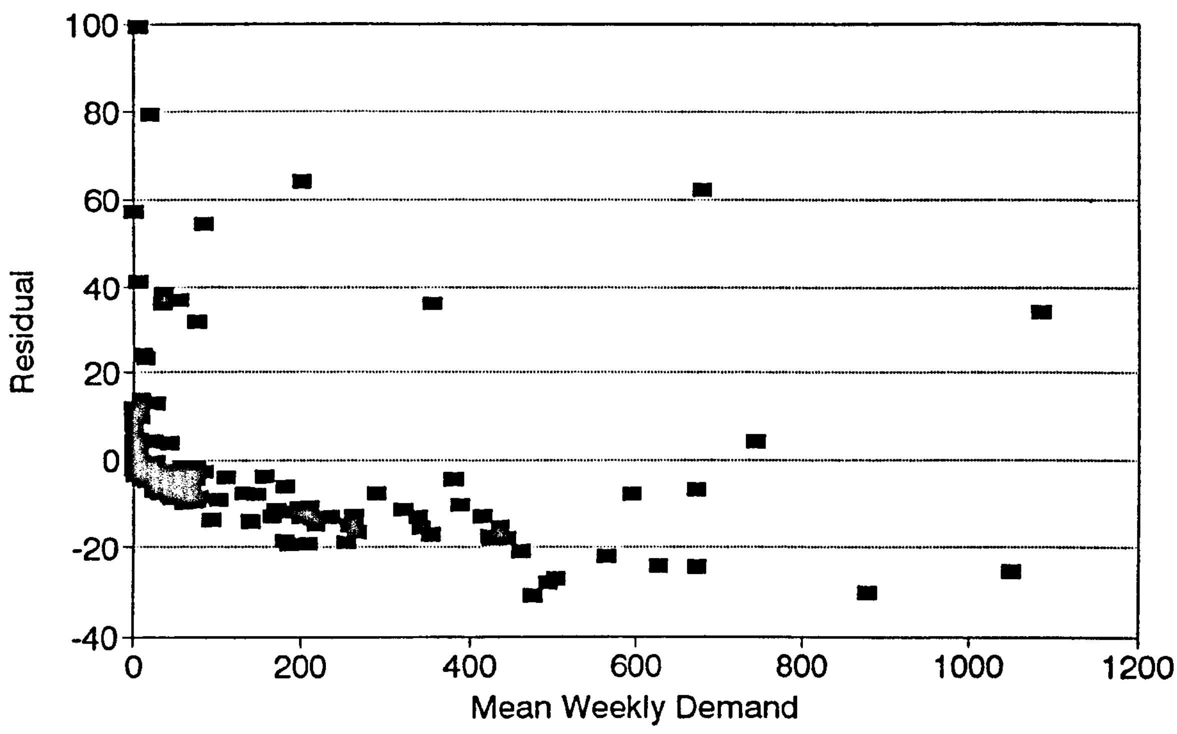
| | *Parameter estimate* | *Standard error* |
|---|---|---|
| *A* | 1.000 | |
| *B* | 1.447 | 0.145 |

As may be expected, the standard error of the B parameter reduces when the A parameter is forced to take the value of unity.

### 13.8.2 Weighted least squares estimation

The residuals from the relationship with the constant term forced to unity are shown in Figure 13.13 below :

Figure 13.13
Residuals from mean order-size - mean demand relationship (OLS)




As in previous analyses, the graph shows some evidence of heteroscedasticity. Therefore, the relationship will be re-estimated using weighted least squares.

The weighting factors may be obtained by the following argument, under the usual assumptions of normality and a constant sample size, n :

$$\text{sampling variance of } b_1 \quad = \quad \frac{b_2}{n}$$

By the postulated relationship $b_2 / b_1^2 = \gamma$,

$$\text{sampling variance of } b_1 = \frac{\gamma\, b_1^2}{n}$$

$$= \frac{\gamma}{n} \left( 1 + B\, c_1^{C} \right)^2$$

Since $\gamma$ is a constant and all SKUs have the same number of data points, the weighted sum of squares will be weighted by the reciprocal of $(1 + B\, c_1^{C})^2$. This result is exact, under the conditions given above, if the mean order size, $b_1$, is estimated using the sample mean.

The results from re-estimation of the relationship are summarised in Table 13.13, alongside the ordinary least squares results.

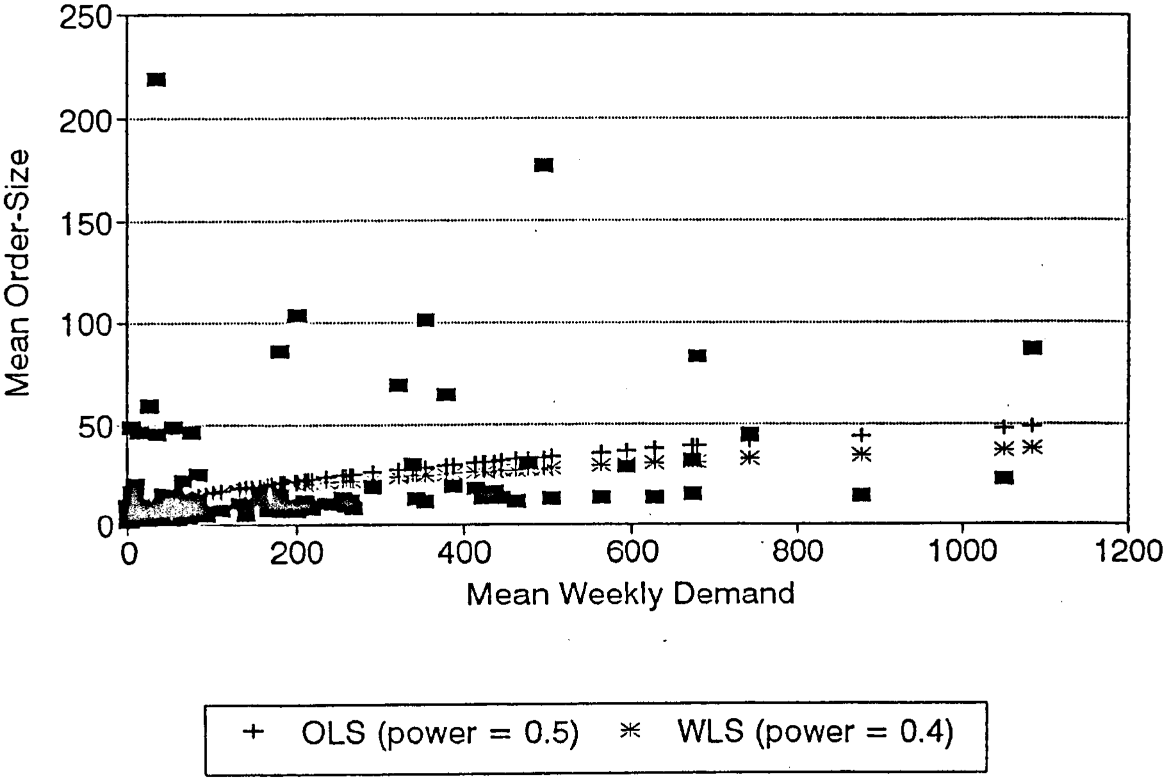
<u>TABLE 13.13</u>
<u>Mean order-size - mean demand parameter estimates</u>
<u>(ordinary and weighted least squares)</u>

| | *Ordinary least squares* | *Weighted least squares* |
|---|---|---|
| *A* | 1.000 | 1.000 |
| *B* | 1.447 | 2.215 |
| *C* | 0.500 | 0.400 |

Note : For both OLS and WLS estimation, powers of C = 0.3, 0.4, 0.5, 0.6 and 0.7 were examined.

The effect on the line of best fit is shown in Figure 13.14 :

Figure 13.14
Mean order-size - mean demand relationships




The weighted least squares fit is less influenced by those SKUs with highest mean order-sizes and is, therefore, more robust to mean order-sizes unduly affected by one very large order.

## 13.9 Estimation of the relationship between the mean and variance of demand

### 13.9.1 Ordinary least squares estimation

In the previous chapter, a 'general' demand variance law was postulated which contained linear, power and quadratic terms. In section 13.8, weak power relationships between mean order-size and mean demand were found, of the form:

Mean order-size = A + B * (Mean demand)$^{0.5}$ (ordinary least squares)

Mean order-size = A + B * (Mean demand)$^{0.4}$ (weighted least squares).

To maintain consistency with the first relationship shown above, ordinary least squares analyses will be conducted to estimate the parameters in the following two relationships linking the mean and variance of demand :

*Quadratic* Variance = A * Mean + B * (Mean)$^2$

*General* Variance = A * Mean + B * (Mean)$^{1.5}$ + C * (Mean)$^2$

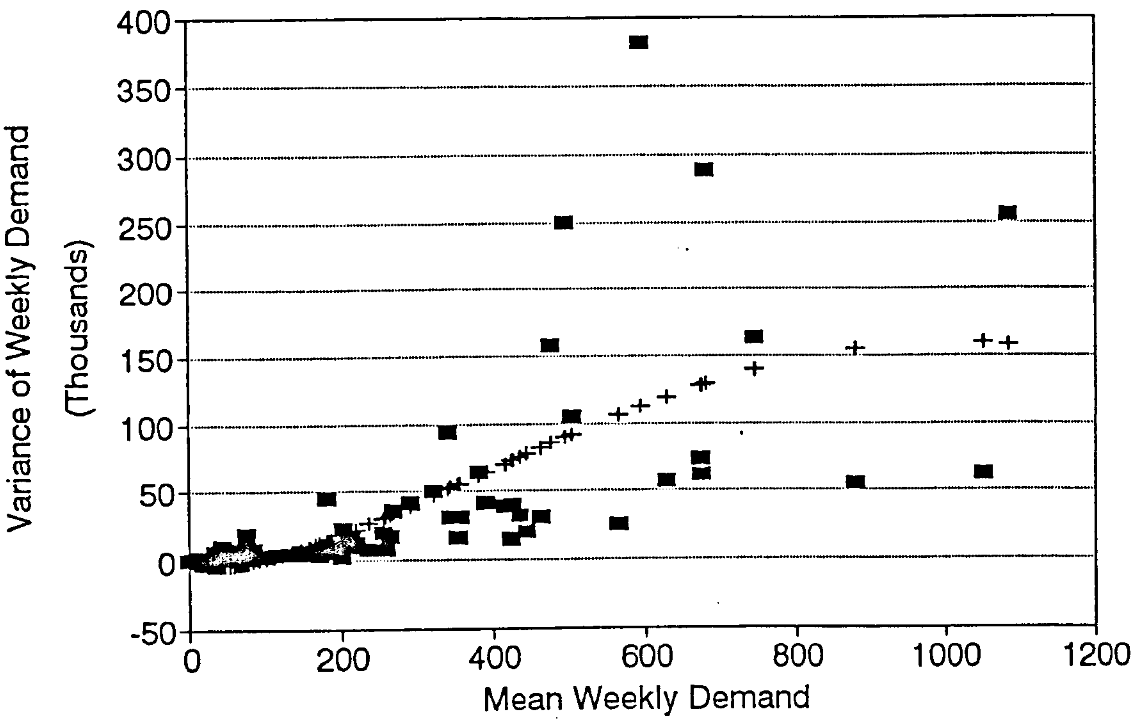
The $R^2$ values for the two relationships are given in Table 13.14 :

TABLE 13.14
Demand variance - mean $R^2$ measure

| | *R - squared* |
|---|---|
| Quadratic relationship | 50.7 % |
| General relationship | 54.2 % |

Although the general relationship shows a slight improvement on the quadratic relationship, neither shows the strength of the mixed quadratic relationship for demand incidence. The approximate nature of the general relationship is shown in Figure 13.15 :

Figure 13.15
General demand variance - mean relationship (OLS)




The graph shows that the relationship is strongly heteroscedastic, confirming the findings of Stevens (1974). The variability of the data suggests that it may be difficult to distinguish between the general and quadratic variance laws. The parameter estimates and their standard errors for both variance laws are shown in Table 13.15:

TABLE 13.15
Demand variance - mean parameters (OLS estimates)

| | *Linear Term* | *Power Term (Power = 1.5)* | *Quadratic Term* |
|---|---|---|---|
| | Estimate (Standard Error) | Estimate (Standard Error) | Estimate (Standard Error) |
| Quadratic | 129.465 (23.491) | | 0.047 (0.033) |
| General | -322.998 (110.618) | 40.435 (9.676) | -0.796 (0.204) |

The table shows that, in the general variance relationship, two of the parameters are negative and are significantly less than zero. Since this is inconsistent with the theory propounded in chapter 12, and results in negative variance estimates for many SKUs, the 'general variance relationship' obtained by OLS may be disregarded for the PES data. For the quadratic relationship, only the linear parameter is significantly higher than zero (5% significance level). This is inconsistent with the finding in section 13.5 that the quadratic term in the variance - mean relationship is significantly greater than zero.

### 13.9.2 Weighted least squares estimation

It has been shown that ordinary least squares estimation of the parameters in the general relationship yields nonsensical results. As strong heteroscedasticity was evident, it would be useful to re-estimate the parameters using weighted least squares. Indeed, it was the variance-mean demand relationship which prompted Stevens (1974) to use WLS to counter the effects of heteroscedasticity.

It is intended to maintain consistency with the relationship obtained between mean order-size and mean demand using weighted least squares, namely :

$$\text{Mean order-size} = A + B * (\text{Mean demand})^{0.4} .$$

Therefore, analyses will be conducted to estimate parameters for the following 'general' relationship :

$$\text{Variance} = A * \text{Mean} + B * (\text{Mean})^{1.4} + C * (\text{Mean})^{2} .$$

The results of the weighted least squares analyses for both quadratic and general relationships are shown in Table 13.16 :

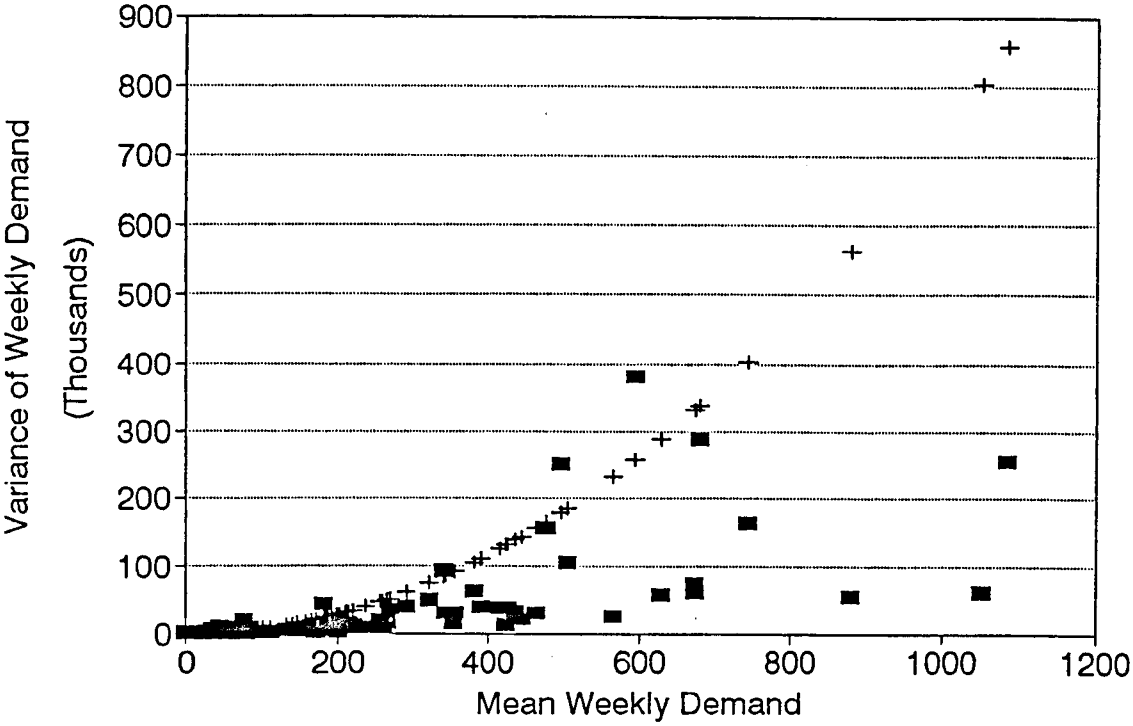
TABLE 13.16
Demand variance - mean parameters (WLS estimates)

| | *Linear Term* | *Power Term (Power = 1.4)* | *Quadratic Term* |
|---|---|---|---|
| Quadratic | 2.872 | | 0.726 |
| General | -0.467 | 4.524 | 0.249 |

The use of weighted least squares does not entirely overcome the difficulty of negative parameter estimates for the general relationship. Therefore, the quadratic relationship, using WLS estimation, is to be preferred.

The approximate fit given by the quadratic variance law, estimated using weighted least squares, is shown in Figure 13.16 :

Figure 13.16
Quadratic demand variance - mean relationship (WLS)




The use of WLS has resulted in estimates which are reasonably accurate for most slow or medium moving SKUs but which are highly inaccurate for the fastest moving items.

## 13.10 Conclusions

### 13.10.1 Variance of demand incidence

The empirical results from PES data were entirely consistent with the theory developed earlier in chapter 12 for demand incidence. The quadratic variance-mean relationship showed a good fit ($R^2$ = 87.6%). Both linear and quadratic parameters were significantly higher than zero, and the linear parameter was not significantly different from its theoretical value of unity. When the linear parameter was constrained to unity, the goodness of fit was barely affected.

### 13.10.2 Independence of mean order-size and mean demand incidence

It was shown in chapter 12 that independence between mean order-size and mean demand incidence is a necessary condition for both the quadratic and general variance laws. For the PES data, two relationships were tested: linear and inverse (reciprocal). In both cases, the $R^2$ value was very low, confirming that the independence condition is met for this data.

### 13.10.3 Variance of order-sizes

Analysis of the PES data showed a very good fit for a quadratic relationship between the variance and mean of order-sizes ($R^2$ = 96.9 %), if the two SKUs with the highest mean order-sizes are excluded from the calculation. Under ordinary least squares estimation, the quadratic coefficient was very sensitive to the presence of the two outlying SKUs (estimate of 0.517 for all data, and 1.012 with two outliers excluded). This difficulty was overcome by weighted least squares estimation, which yielded an estimate of 0.941. The closeness of the parameter estimate to unity indicates a shape parameter of one for the hypothesised gamma distribution of order-sizes. This is simply the negative exponential distribution. Its monotonic decreasing density function is consistent with an order-size of one being the most common, a pattern observed for most of the SKUs in the sample. Although this is intuitively reasonable, it must be noted that no formal test on distributions of order-sizes have been conducted.

### 13.10.4 Mean order-sizes

Empirical analysis revealed a weak relationship of the form $b_1 = A + B\ c_1^{0.5}$ (where $b_1$ is mean order-size and $c_1$ is mean demand). The $R^2$ value was low, at 17.3 %, but the B parameter was significantly higher than zero. Ordinary and weighted least squares analyses were also conducted for relationships of similar form but with the A parameter fixed at unity, for consistency with the theory developed in chapter 12. Similar results were obtained; the best power index was found to be 0.4 with weighted least squares. Overall, the results showed a disappointingly weak linkage between mean order-size and mean demand. It is possible that better results would be obtained if pack-sizes were taken into account. However, pack-size data was not available on the extracted files used for the empirical analysis in this thesis.

### 13.10.5 Variance of demand

The general variance law, $c^2 = A\ c_1 + B\ c_1^C + D\ c_1^2$ and the quadratic variance law, $c_2 = A\ c_1 + B\ c_1^2$, were compared using the PES data. An $R^2$ value of 54.2 % was obtained for the goodness of fit of the general law, an improvement on the

50.7 % obtained for a quadratic relationship. However, neither ordinary least squares nor weighted least squares analyses yielded consistently positive parameters for the general law, as required to be consistent with the theory developed in chapter 12. It would, therefore, seem more appropriate to use the simpler quadratic relationship since it gave consistently positive parameter estimates.

# CHAPTER 14

## *Application of Variance Laws in Centralisation Modelling*

### 14.1 Introduction

In this chapter, the implications of variance law theory on inventory centralisation modelling are assessed. This investigation proceeds in a number of stages. Firstly, the 'stationary mean' model is revisited and its application in a centralised demand setting is examined. In particular, the effect of centralisation on the distribution of demand is assessed. If it is known that demand is compound Poisson distributed at decentralised depots, does a compound Poisson also apply at the centralised depot ?

The effect of centralisation on each of the parameters in a quadratic variance law for demand are investigated, namely :

1. Variance of demand incidence.
2. Mean order-size.
3. Variance of order-size.

Methods for estimating the effect of centralisation on each of these parameters are presented.

Finally, the effect of centralisation on demand is assessed by integrating the three parameters into an overall model. Although this model contains linear and quadratic mean demand components, the parameters are not necessarily independent of mean demand, except in certain circumstances.

## 14.2 The stationary mean model for centralised demand

In this section, the stationary mean model of chapter 12 is revisited. The model allows for non-constant order-sizes and for variation by the mean demand incidence. If it is assumed that order-size is independent of order-incidence, then the demand equations are :

*Equation 1* $$y_{it} = \Sigma_{j=1}^{n_{it}} A_{ijt}$$

*Equation 2* $$n_{it} = \lambda_{it} + \varepsilon_{it}$$

*Equation 3* $$\lambda_{it} = \lambda_i + \eta_{it}$$

where $A_{ijt}$ is the size of the *j*th order in time period t for the *i*th SKU
$n_{it}$ is the number of orders in time period t for the *i*th SKU
$\varepsilon_{it}$ ) are stochastic disturbance terms.
$\eta_{it}$ )

If demand is centralised, and it is assumed that :

1. Ordering patterns from customers do not change
2. Total demand is unchanged,

then the centralised demand for the *i*th SKU in time t may be represented by :

*Equation 1* $$\Sigma_k y_{ikt} = \Sigma_{j=1}^{\Sigma n_{ikt}} A_{ijkt}$$

*Equation 2* $$\Sigma_k n_{ikt} = \Sigma_k (\lambda_{ikt} + \varepsilon_{ikt})$$

*Equation 3* $$\Sigma_k \lambda_{ikt} = \Sigma_k (\lambda_{ik} + \eta_{ikt})$$

where the notation is as before, with the k suffix added to denote the numeric label of the decentralised depot location.

In the above formulation, the $\varepsilon_{ikt}$ disturbance terms are the independent components of variance; the $\eta_{ikt}$ disturbance arise from 'family effects', as described in sub-section 12.2.2, and so will not be independent between inventory locations.

Since the stochastic disturbance terms, $\varepsilon_{ikt}$ (1 = 1, ...n), are independently distributed, the centralised demand incidence is Poisson distributed about the mean $\Sigma_k \lambda_{ikt}$. The 'mixing' equation (Equation 3) specifies the disturbance term, $\Sigma_k \eta_{ikt}$, to the mean. Equation 1 represents the 'compounding' that arises from pooling all of the orders which would have been received by the individual decentralised depots. Thus, the above formulation is consistent with compound Poisson demand which has arisen from both 'mixing' and 'compounding' processes.

From the analysis in chapter 12, treating mean demand incidence as a continuous variable over time, it is known that the following result holds for compound Poisson demand :

$$c_2 = \left( \frac{b_2}{b_1^2} + 1 \right) b_1 c_1 + \frac{a_2}{a_1^2} c_1^2$$

where the notation is unchanged.

It is important to note that the parameters in the above equation are not necessarily invariant under centralisation. The linear and quadratic coefficients must be investigated for the effect of centralisation. This examination will be conducted in the next three sections, beginning with the quadratic parameter and concluding with the two components of the linear parameter.

## 14.3 Variance of demand incidence

### 14.3.1 Non-correlated demand between depots

The models of Maister (1976) and Das (1978) were reviewed in chapter 2. In both cases, it is assumed that demand is not correlated between any pair of depots. For such models, the estimate of the variance of demand incidence at the centralised depot is simply the sum of the variance estimates at each decentralised depot.

Since there is no correlated component, the demand equations simplify to :

*Equation 1* $$\Sigma_k\, y_{ikt} = \Sigma_{j=1}^{\Sigma n_{ikt}} A_{ijkt}$$

*Equation 2* $$\Sigma_k\, n_{ikt} = \Sigma_k\, (\lambda_{ikt} + \varepsilon_{ikt})$$

and since the second term represents the sum of independent Poisson streams of demand incidence,

$$\mathrm{Var}(\Sigma_k\, n_{ikt}) = \Sigma_k\, \lambda_{ikt}\,.$$

Thus, the special case of a linear variance law translates to a centralised setting quite naturally.

### 14.3.2 Correlated demand between depots

In the models reviewed in chapter 2 which assume correlated demand, such as Zinn, Levy and Bowersox (1989) and Tallon (1993), estimates are required of the correlation of demand between each pair of depots. However, in the circumstances outlined above, it may prove difficult to estimate the correlations which are required for the models.

Suppose that the size of a 'disturbance term' is proportional to mean demand incidence and that this proportionality extends between depots. Then a ratio of disturbance terms may be written :

$$\frac{\eta_{it}}{\eta_{kt}} = \frac{\lambda_i}{\lambda_k} \qquad \text{for } i = 1, \ldots, n$$

where $\eta_{it}$ is the disturbance to the mean demand incidence of a particular SKU at depot i, at time t,

$\lambda_i$ is the underlying mean demand incidence of the same SKU at depot i.

The suffix notation has changed to denote depots rather than SKUs, with the subscript for SKUs being dropped in order to avoid a proliferation of suffices. Then, the variance of the demand incidence at depot i, $n_{it}$, is given by :

$$Var(n_{it}) = \lambda_i + \frac{Var(\eta_{kt})}{\lambda_k^2} \lambda_i^2$$

The relationship applies at each decentralised depot, with the same quadratic parameter. This may be tested empirically from the estimated relationships at each depot. If they are consistent, as predicted by the theory outlined above, then the covariance between demand incidence at two depots i and j is given by :

$$Cov(n_{it}, n_{jt}) = E(\eta_{it}\, \eta_{jt}) = \frac{E(\eta_{kt}^2)}{\lambda_k^2} \lambda_i \lambda_j$$

assuming that the other stochastic terms ($\varepsilon_{it}$ for i =1, ..., n) are independent of each other and of the $\eta_{it}$ terms.

Therefore, the variance of demand incidence at the centralised depot location will be given by :

$$\begin{aligned}\mathrm{Var}(\Sigma_j n_{jt}) &= \Sigma_j \lambda_j + \frac{E(\eta_{kt}^2)}{\lambda_k^2} [\Sigma_j \lambda_j^2 + 2\, \Sigma\Sigma_{i<j} \lambda_i \lambda_j] \\ &= \Sigma_j \lambda_j + \frac{E(\eta_{kt}^2)}{\lambda_k^2} [\, \Sigma_j \lambda_j \,]^2\end{aligned}$$

Hence, if the same quadratic relationship applies at each decentralised depot, then this relationship may also be used at the centralised depot to predict the variance of demand incidence. The formula automatically takes into account correlation between demand at different depots without needing to use sample correlations as in the analysis of Zinn, Levy and Bowersox (1989).

The quadratic term in the variance-mean relationship for demand, rather than demand incidence, is independent of order-size and is a function of demand incidence alone. Therefore, the same quadratic term may also be used in the demand variance-mean relationship at the centralised depot.

The above analysis is restricted to those situations where the disturbance term applies to all depots in proportion to the size of mean demand. However, situations may arise where there is a correlation effect but the disturbance may be greater or lesser at some decentralised depots, according to a set of scaling factors. This may be represented as follows :

$$\frac{\eta_{it}}{\eta_{kt}} = \alpha_i \frac{\lambda_i}{\lambda_k} \qquad \text{for } 1 = 1, \ldots, n.$$

In this case, the following quadratic relationships apply at the decentralised depots :

$$\mathrm{Var}(n_{it}) \quad = \quad \lambda_i \quad + \quad \alpha_i^2 \; \frac{\mathrm{Var}(\eta_{kt})}{\lambda_k^2} \; \lambda_i^2 \qquad \text{for } i = 1, \ldots, n.$$

Although each relationship is of the same form, the quadratic parameter may differ between depots. Using similar arguments to those presented above, it may be shown that :

$$\mathrm{Var}(\Sigma_j n_{jt}) \quad = \quad \Sigma_j \lambda_j \; + \; \frac{E(\eta_{kt}^2)}{\lambda_k^2} \; [\, \Sigma_j \, \alpha_j \, \lambda_j \,]^2$$

Thus, the relationship is *not* a quadratic function of $\Sigma_j \lambda_j$ at the centralised depot. The third requirement for a quadratic relationship, namely that $a_2/a_1^2$ is independent of mean demand, does not apply in this case.

The above analysis shows how a consideration of the theory underpinning a quadratic functional relationship prevents the relationship being misapplied in centralisation models. Two cases have been identified. In the first, where proportionality of the disturbance effect extends across depots, the quadratic relationship is identical at all decentralised depots and at the centralised depot. In the second, where proportionality may be modified by scaling factors, the relationship at the centralised depot is no longer quadratic. However, if each of the decentralised relationships are known, then an estimate the variance of centralised demand incidence may be found using the formula which has been derived.

## 14.4 Mean order-size

### 14.4.1 Customers' ordering patterns

To establish a model for centralised demand, it will be necessary to make some assumptions about customers' ordering patterns. In this sub-section, the effect of centralisation on different modes of customer ordering will be assessed.

The first case to consider is when each customer is dedicated to a particular decentralised depot. 'Dedicated' means that all of the orders are placed on the same depot. In this situation, the frequency and size of orders are not affected by centralisation, although the destination of the orders may change.

The second case is when customers 'share out' their orders. There are a number of ways in which customers may share their orders between depots. Using the nomenclature of chapter 7, three possibilities are :

1. Sharing an order for different SKUs to different depots, ensuring that all units for an SKU are ordered from the same depot.

2. For a particular SKU, sharing the order-lines over time, but ensuring that no particular order-line is split between depots.

3. Splitting an order-line, for a particular SKU, between depots.

In the first and second situations, the size and frequency of orders for a particular SKU are not affected by centralisation although, as before, the destination of orders may change.

In the third situation, the pattern of orders would be affected by centralisation since the splitting of orders would cease. However, in this case, it would be possible to achieve most of the inventory benefits which would accrue from centralisation by re-allocating the splitting of orders to the optimal, as identified by Evers and Beier (1993), or as near as possible to the optimal as may be achieved in practice.

To summarise, the pattern of orders from customers is unaffected by centralisation if customers are 'dedicated' to decentralised depots. If they are not 'dedicated', then the patterns of orders are unaffected unless customers split order-lines. However, for this case, an inventory centralisation model is not required. Therefore, for the remainder of this chapter, it will be assumed that customers' ordering patterns are unaffected by centralisation.

### 14.4.2 Inappropriateness of EOQ based relationships

Regardless of the strength of its goodness of fit, the relationship

$$b_1 = 1 + \alpha c_1^{\beta}$$

is not applicable to centralised demand. In chapter 12, it was shown that the motivation behind this relationship is the Economic Order Quantity (EOQ). EOQ theory predicts higher order-sizes for higher mean demands. However, order-size will not necessarily change from individual customers after centralisation. If total demand from a consumer is not affected by centralisation, then there is no reason why order-sizes should rise. Therefore, the EOQ predictions will not apply.

### 14.4.3 Algebraic formula

In this case, the individual order-sizes do not change and there is a 'pooling effect' of taking different mean demands from the decentralised depots. The overall average mean order-size is simply the weighted mean :

$$(b_1)_0 = \frac{\Sigma \lambda_j (b_1)_j}{\Sigma \lambda_j} = \Sigma_j p_j (b_1)_j$$

where $(b_1)_j$ is the mean order-size at decentralised depot j
$(b_1)_0$ is the mean order-size at the centralised depot
$\lambda_j$ is the mean demand incidence at decentralised depot j

$$p_j = \lambda_j / \Sigma_j \lambda_j,$$

and the summation extends over all decentralised depots.

Thus, a simple formula may be used which requires no more information than the mean demand incidences and order-sizes at each decentralised depot.

## 14.5 Variance of order-sizes

### 14.5.1 Non-identical distributions of order-sizes

If there is no change in the orders from customers, then the effect of centralisation is one of 'pooling' the order-sizes from each of the decentralised depots. Following the model developed in chapter 12, it will be assumed that there is no correlation between the sizes of individual orders at different depots. If this is the case, then the variance of order-sizes at the centralised depot is given by the following 'pooling' formula :

$$(b_2)_0 = \Sigma_j \, p_j \, (b_2)_j + \Sigma_j \, p_j \, [\, (b_1)_j - \Sigma_k \, p_k \, (b_1)_k \,]^2$$

where $(b_2)_j$ is the order-size variance at decentralised depot j
$(b_2)_0$ is the order-size variance at the centralised depot.

If the distributions of order-sizes are not identical at each of the decentralised depots, and if the relationship $(b_2)_j / [(b_1)_j]^2 = \gamma_j$ holds (for i =1, ..., n) in keeping with the theory developed in chapter 12, then the above expression for centralised variance simplifies to :

$$(b_2)_0 = \Sigma_j \, p_j \, (\gamma_j + 1) \, [(b_1)_j]^2 - [\, \Sigma_j \, p_j \, (b_1)_j \,]^2$$

from which it follows that the variance to mean-squared ratio is given by :

$$\frac{(b_2)_0}{[(b_1)_0]^2} = \frac{\Sigma_j \, p_j \, (\gamma_j + 1) \, [b_1)_j]^2}{[\, \Sigma_j \, p_j \, (b_1)_j \,]^2} - 1$$

It is apparent that this ratio does not, in general, take a constant value. This result is not surprising since the pooling of n non-identical gamma distributions will not, in general, yield a gamma distribution, for which the ratio is constant.

## 14.5.2 Identical distributions of order-sizes

In the special case that the n shape parameters $\gamma_1, ..., \gamma_n$ are identical and the n mean-order sizes $(b_1)_1, ..., (b_1)_n$ are also identical, then the ratio becomes :

$$\frac{(b_2)_0}{[(b_1)_0]^2} = (\gamma + 1) \frac{\Sigma_j p_j [b_1]^2}{[\Sigma_j p_j (b_1)]^2} - 1$$

Since the $b_1$ terms cancel and since $\Sigma_j p_j = (\Sigma_j p_j)^2 = 1$, this simplifies to :

$$\frac{(b_2)_0}{[(b_1)_0]^2} = \gamma$$

So, in the special case of identical gamma distributions, the pooling of order-sizes results in the same gamma distribution, and the same variance to mean-squared ratio applies.

## 14.6 Variance of demand

### 14.6.1 General variance law for centralised demand

If the different components of centralised demand variance are combined, then the general formula is given by :

$$(c_2)_0 = \frac{\Sigma_j\, p_j\, (\gamma_j + 1)\, [(b_1)_j]^2}{\Sigma_j\, p_j\, [(b_1)_j]^2} \quad \Sigma_j\, p_j\, (b_1)_j \;\; \Sigma_j\, (c_1)_j$$

$$+ \; \frac{E(\eta_k^2)}{\lambda_k^2} \; \frac{[\Sigma_j\, \alpha_j\, \lambda_j\,]^2}{[\,\Sigma_j\, \lambda_j\,]^2} \; [\Sigma_j\, (c_1)_j\,]^2$$

where $(c_1)_j$ is the mean demand at depot j (depot 0 is the centralised depot)
$(c_2)_j$ is the variance of demand at depot j.
and the remaining notation is unchanged.

The expression is quite complicated as it represents an amalgam of the most general form of relationship for each component analysed in this chapter. However, the expression contains no sample variance estimates: only mean estimates and parameters of functional relationships are required for the overall demand variance estimate.

### 14.6.2 Centralised quadratic variance law as a special case

The expression in sub-section 14.6.1 is not identical to the quadratic expressions obtained at each of the decentralised depots. Hence, decentralised variance laws do not translate, in general, to centralised variance laws. In the following special case, the variance expressions are the same for centralised and decentralised scenarios :

1. Shape parameters, $\gamma_j$, are identical at decentralised depots.
2. Mean order-sizes, $(b_1)_j$, are identical at decentralised depots.

3. Scaling factors, $\alpha_j$, are all equal to unity.

In this case, the centralised demand variance expression collapses to :

$$(c_2)_0 = (\gamma + 1)\, b_1 \Sigma_j (c_1)_j + \frac{E(\eta_k^2)}{\lambda_k^2} [\Sigma_j (c_1)_j]^2$$

which is identical to the decentralised variance expressions. Hence, the quadratic variance law applies at the centralised depot location as a special case.

### 14.6.3 Testing the centralised variance expression

It has not been possible to perform an empirical comparative analysis on decentralised and centralised variance expressions using data from Pillar Engineering Supplies (PES) Ltd. In PES, only one depot was operational, namely the National Distribution Centre at Leicester. Although the company operated a branch network, there were an insufficient number of SKUs with an adequate volume of sales at branch level, to enable the variance laws to be estimated. The empirical testing of decentralised and centralised variance laws is a key area of further research. A suitable company for such an analysis is currently being sought.

## 14.7 Conclusions

The implications of the variance laws, developed theoretically in chapter 12 and tested empirically in chapter 13, have been examined in this chapter.

It was shown that if ordering patterns from customers remain unchanged, then the stationary mean model may be extended to centralised demand. If demand is compound Poisson distributed at decentralised locations, then centralised demand will also be compound Poisson. The quadratic variance-mean relationship for demand incidence was found to take inter-depot demand correlations into account automatically, if disturbances to mean demand incidences are assumed to be proportional across depots. The same quadratic function which applies at decentralised depots also holds at the centralised depot. If the proportionality of disturbances needs to be modified by scaling factors, then it was shown that the centralised demand variance function is no longer quadratic.

The effect of centralisation on order-sizes was also assessed. It was shown that the EOQ approach is not appropriate for this assessment. Instead, algebraic relationships may be used to find the centralised mean and variance of order-size in terms of the mean and variance of order-size at the decentralised depots.

The overall effect of centralisation on the variance of demand was expressed in terms which involved no sample variance estimates. Mean estimates and estimates of coefficients in functional relationships were required. The centralised model has not yet been tested empirically; this represents the next stage of research.

# PART V

# CONCLUSIONS

# CHAPTER 15

## *Conclusions, Limitations and Further Research*

In this chapter, the main threads of the research presented in the thesis are drawn together and the principal conclusions are summarised in a concise form. The main limitations of the work are assessed and, where appropriate, avenues of further research are suggested.

### 15.1 Introduction

The purpose of this thesis is to operationalise centralisation models, rather than present new models. It would not be difficult to continue developing models along the lines of such authors as Zinn, Levy and Bowersox (1989), Tallon (1993) and Evers and Beier (1993). However, there seems little point in doing so until the limitations of their work are addressed. The papers in the literature discuss the algebra of centralisation models, but do not adequately address two issues, namely:

1. The treatment of service levels.
2. The estimation of demand variance and covariance.

Cost-reduction is often a driving force behind inventory centralisation, but another motivating factor may be the achievement of a common service goal to all customers, using a common measure. This issue was addressed in part II of the thesis.

Variance estimates are required for all centralisation models, unless it is assumed that the variances are identical at all decentralised depots. Similarly, the effect of covariance must be assessed unless it is assumed that demands at all pairs of depots are un-correlated. However, the estimates may be difficult to obtain for slow-moving items, products which have been recently launched or which have experienced a step-change in mean demand. In these cases, a reliable mean estimate is easier to obtain and a relationship linking the variance to the mean of demand would be useful. The theory underpinning these models was discussed in parts III and IV. In this chapter, the implications of this work for the decision of centralising or not centralising an SKU are examined.

Demand distributions and variance laws has been tested empirically but inventory amalgamation models have not. The next stage of research is to test such models empirically.

## 15.2 Conclusions from part I : inventory amalgamation models

### 15.2.1 Classification of inventory amalgamation models

No previous attempts have been made to classify inventory amalgamation models. The multitude of fixed location inventory models and the relative paucity of inventory amalgamation models is a factor in explaining the omission. Although few in number, examination of published inventory amalgamation models revealed some mis-classifications by the authors. In chapter 1, it was argued that the use of a classification system would alleviate this problem and provide a useful framework for the operationalisation of amalgamation models. Classifying the key model assumptions facilitates the correct choice of model.

Since Chikán's (1990) classification of fixed-location inventory models was found to be the most comprehensive, it was used as the basis for the development of a new classification system for inventory amalgamation models. Some limitations of Chikán's system were addressed and adaptations made for amalgamation models. A simple classification system was developed, based on five factors:

1. Single item / multi-item system
2. System objective
3. Ordering rule
4. Character of lead-time
5. Correlation of demand between depots.

Since the classification system was designed with published models in mind, it is hardly surprising that the system works well for these models. The weaknesses of the system will become apparent as further models are published in the literature. If limitations to the system are revealed, then these may be remedied by the addition or removal of appropriate factors.

### 15.2.2 Safety stock centralisation models

A literature review revealed that in most papers on safety stock centralisation modelling, scant attention has been paid to the estimation of demand variance and covariance between depots. Since there are many circumstances in which these variables may be difficult to estimate, this neglect is surprising. Accurate estimation of these variables is particularly difficult for slow-moving SKUs. As will be shown later in this chapter, it is these SKUs which have the greatest potential for stock-saving by centralisation.

Using the new classification system, it was shown that Zinn, Levy and Bowersox's (1989) centralisation model should be classified as a service driven model and not, as claimed by the authors, a cost-driven model. Ronen's (1990) criticism that the use of the 'probability of stock out during lead time' service criterion was inappropriate was noted. In chapter 2, it was argued that centralisation models should have the capacity to deal with all service measures in common use, a theme developed in part II of the thesis.

Tallon's (1993) model, extending safety stock centralisation models to variable lead-times, was generalised in chapter 2. The revised formulation includes variables which had been omitted in the original paper. It was noted that the variance estimation problems identified for constant lead-times are magnified by lead-time variability.

The classification system was used to analyse previously published models, but no attempt was made to identify gaps in model provision, nor to fill those gaps with new models. In this thesis, the view has been taken that variance estimation issues should be resolved before extending the range of available models. Although a conscious limitation of the research, the lack of new safety stock models is a limitation, nevertheless. For example, cost-based models with variable lead-times have not been analysed. A systematic analysis of models, to cover all feasible combinations of model assumptions according to the classification, must await future research.

### 15.2.3 Safety stock consolidation models

A literature review of consolidation models prompted a discussion on transport costs. The issue of transport modelling was neglected by Evers and Beier (1993) but was included in the Portfolio Cost Effect model of Mahmoud (1992). This matter will be considered again in section 15.6.

Evers and Beier (1993) advanced the proposition that all of the safety stock benefits of centralising to one depot may be achieved by consolidating to m depots ($m > 1$). Their argument rests on the assumption that demands may be re-allocated to depots in such a way that demands at all depots are perfectly correlated. The authors had not considered the operationalisation of their findings, however. Re-allocation may be financially undesirable when transport costs are taken into consideration. Also, if demand is sufficiently slow, further stock savings may be made by centralising to one depot, whether or not demand is correlated.

Mahmoud's (1992) treatment of 'sub-consolidations' versus 'super-consolidation' preference was corrected. Explicit preference rules were given. It was shown that the indifference function is not linear, as suggested by Mahmoud, but quadratic.

In addition to the specific criticisms of consolidation models summarised above, it was argued in chapter 3 that the general problem of variance estimation applies with as much force to consolidation models as to centralisation models.

As in the case of centralisation models, previously published models were classified and some errors were corrected. However, no attempt was made to identify gaps in the provision of models, nor to design new models to fill those gaps. This work must await further research.

### 15.2.4 Cycle stock centralisation models

Maister (1976) claimed that the 'square root law' is not sensitive to the assumption of equal mean demands at all depots, but Das (1978) gave a counter-example to demonstrate that the opposite may be true. In chapter 4, the circumstances under which robustness of the square root law breaks down were identified. Indeed, formulae were provided which enable the proportion of maximum cycle stock savings to be calculated when the mean demands are not identical at all decentralised depots. The formulae reveal that the robustness of the cycle stock square root law improves as the number of inventory locations to be centralised increases.

Cycle stock centralisation, under the standard Economic Order Quantity (EOQ) formula, has been analysed exhaustively. Future research will concentrate on the situations when the EOQ assumptions are not met.

### 15.2.5 Robustness to Economic Order Quantity assumptions

The treatment of cycle stock centralisation in this thesis, as in the literature generally, is based on the EOQ approach. However, criticisms of the EOQ have not been previously addressed in the field of centralisation; in chapter 5, the implications of those criticisms for centralisation models were assessed.

It was shown that cycle stock square root laws are surprisingly robust to deviations from the EOQ model assumptions. These findings are limited, however, to deviations from an individual assumption, with all other assumptions holding. No investigations have been made into 'interaction effects', ie the implications of two or more assumptions being violated at the same time. This is a topic which must await further research.

It was also argued that difficulties in parameter estimation are often irrelevant to centralisation models since the parameters cancel out (assuming the parameters are unaffected by centralisation). This applies for many approximate results (eg Silver (1976)) but is rarely true for exact results (eg Goyal (1974)).

It was noted that the square root law is sensitive to the replacement of a 'minimum cost' criterion by an 'ROI maximisation' criterion. Upper and lower bounds were derived for the centralised to decentralised cycle stock ratio. The results are limited to those situations when parameters other than equity and fixed costs are assumed to be identical at decentralised depots and unchanged by centralisation. Further research is required into the more general case.

The criticism that the EOQ leads to an 'optimising' approach rather than a 'managing' approach was noted, particularly with reference to 'managing down' the set-up costs. However, it was argued that, in the context of inventory centralisation, the formulae could be used to assess the impact of centralisation with concomitant reduction of set-up costs if they were possible, or without if they were not.

## 15.3 Conclusions from part II : inventory service models

### 15.3.1 Benefits and disbenefits of inventory centralisation

Research into the effect of centralisation on service level to customers was reviewed. Much of the work in the literature is very mathematical in its focus and little attention has been given to the practical implications of the findings.

It would seem an intuitively obvious result that, if total stocks are unchanged by centralisation, then service must improve to all customers. However, Stulman's (1987) assertion, that centralised inventory need never be higher than decentralised inventory to achieve at least the same service level, was shown to be untrue by Chen and Lin (1990). However, Chen and Lin's counter-example is artificial since the probability of a stock-out is so high (86%) that no organisation could prosper with such a poor inventory service level.

A general inequality for service benefits / disbenefits was presented in chapter 6. Using this inequality for two depots, each with one item in stock and with identical mean demands, it was shown that disbenefits occur if mean demands are greater than a 'critical mean' value of 2.149. At this level of mean demand, with one item in stock, service levels are below 37%. Since this level is unrealistically low, it may be concluded that, under the given conditions, service disbenefits will not arise in practice.

Investigation into limiting values led to a new conjecture for the more general case of n decentralised depots ($n \geq 2$), each with one unit in stock, assuming that mean demand and assurance of service constraints are identical at each depot :

*The probability of stock-out never deteriorates by centralisation if the total stock is unchanged, provided that the mean demand at a decentralised depot is less than e-1. The corresponding 'probability of stock-out' constraint at decentralised depots is no less than* $1-2^{2-e}$.

The analysis was extended to the case where there are m items ($m \geq 1$) in stock rather than one item in stock at each decentralised depot, but retaining the assumptions of identical mean demands and 'assurance of service' constraints. The above conjecture was modified to the following :

*The probability of stock-out never deteriorates by centralisation if the total stock is unchanged provided that the mean demand at a decentralised depot is less than m + ln(2). The corresponding 'probability of stock-out' constraint at decentralised depots is no less than 1/2.*

The two conjectures are consistent for the special case of one in stock at each decentralised depot, but the first conjecture is sharper than the second for this case. Since 50% is an unrealistically low service level, the problem of disbenefits will not arise in practice if the conjectures are true.

A limitation of the results presented is that complete formal proofs have not yet been provided. A further weakness is that the restrictive assumption of equal mean demands at decentralised depots has been maintained in order to keep the analysis tractable. Further research will be directed towards finding formal proofs of the conjectures and investigating the more general case.

### 15.3.2 Relationships between service-level measures

A number of measures are in common use and there is no agreement about which measure is most appropriate, as illustrated by the debate between Ronen (1990) and Zinn, Levy and Bowersox (1989). However, this lack of agreement has not always been reflected in the centralisation literature where 'probability of stock out' is generally the assumed measure. It is therefore important that amalgamation models should have the capacity to deal with all measures in common use. With this capacity, the model can also cater for situations where different measures are in use at different decentralised depots.

A set of relationships was presented, in chapter 7, between six of the most commonly used service level measures :

* order-fill (complete orders)
* order-fill (partial orders)
* line-fill (complete order lines)
* line-fill (partial order-lines)
* units satisfied
* value satisfied.

The relationships were adapted for the situation when historical records of filled orders / order-lines are unavailable, but records of unfilled orders / order-lines are available.

This adaptation was required for an implementation of this model at Unipart Group Ltd (Boylan (1986)) and would be necessary for any organisation which did not keep historical transaction files of filled orders.

The analysis is limited to just six measures, although a similar approach could be used to extend the results to other measures if necessary.

## 15.4 Conclusions from part III : demand distribution families

As noted earlier, centralisation models require variance estimates unless the variance of demand is equal at all centralised depots. Since there are difficulties in obtaining accurate estimates of demand variance, particularly for slow-moving SKUs or if there have been step-changes in mean demand, a linkage between the variance and mean of demand may prove useful. To understand this relationship, it is necessary to take a step backwards and examine the underlying demand distribution and the distribution of time between purchases.

### 15.4.1 Demand distribution models

A distinction was drawn between demand modelling and demand representation. If a probability distribution may be justified in terms of an underlying mechanism, then it is said to model demand. If it is based on empirical evidence with no theoretical foundation, then it is said to represent demand.

Kwan (1991) provided empirical evidence to support the log-zero Poisson (lzP) distribution as a suitable representation of demand. In chapter 8, it was noted that the lzP is a member of the family of compound binomial distributions. An underlying model was suggested for the lzP based on a Bernoulli process of order-line arrivals and a logarithmic Poisson distribution of the number of items per order-line. However, it was argued that there is no natural justification for the use of the logarithmic Poisson in this context.

The negative binomial distribution (NBD) has received extensive support in the literature. It is well known that the distribution may be justified on the basis of random demand incidence and a 'mixing model' of mean levels varying over time but with each order for a single item. It is also known that the NBD distribution may be based on random demand incidence and a 'compounding model' of constant mean order-incidence but with varying order-sizes.

By researching the literature on insurance mathematics it was found that, for random demand incidence, the combined effect of 'mixing' and 'compounding' is modelled by the family of compound Poisson distributions. This finding is particularly significant since, in many real-life demand processes, both factors will be at work simultaneously.

No direct theoretical support was found for the gamma distribution. However, it was noted that the gamma is the continuous analogue of the NBD, a distribution which has support through the processes of 'mixing' or 'compounding'.

The work on demand distribution modelling presented in chapter 8 is by no means comprehensive. It deals with the distribution of demand over a fixed lead-time and ignores the various distributions which have been advocated for variable lead-times, such as the logarithmic Poisson gamma (Nahmias and Demmy (1982)). Further research is needed to extend the findings to variable lead-times. However, the research confirms the findings of Kwan (1991) that the negative binomial is an appropriate distribution to represent demand. Therefore, the commonly made assumption in the centralisation literature of normal demand (eg Eppen (1979)) should be questioned.

### 15.4.2 Inter-purchase time distribution models

The assumption of exponential inter-purchase times has been questioned by some authors. The criticism that many consumer purchases are often immediately followed by a 'dead period' (Kahn and Morrison (1989)) - in contradiction to the shape of the negative exponential distribution - was noted.

A Poisson process, where every second event is 'marked', yields an Erlang-2 inter-purchase distribution. However, it was argued that the marking may be more random in practice. By the 'colouring theorem' (see Kingman (1993)), such random markings yield a negative exponential inter-purchase distribution.

The inverse Gaussian (IG) distribution was proposed by Banerjee and Bhattacharyya (1976) as a model of inter-purchase time. By considering the movement of stocks between receipts, it was argued in chapter 9 that an underlying Brownian motion process is not an appropriate model.

Banerjee and Bhattacharyya's empirical study was replicated and it was shown that the superiority of the IG-based distribution over the NBD depends on the use of an inferior estimation method for the NBD. When maximum likelihood is used for the

NBD, it out-performed the IG-based distribution.

### 15.4.3 Efficiency and bias of estimation methods

Estimation methods for inter-purchase, demand incidence and demand distributions were reviewed in chapter 10, in preparation for the empirical analysis of demand data from Pillar Engineering Supplies. The estimation of exponential, geometric and Erlang-2 inter-purchase distributions is straightforward, since the moments and maximum likelihood (ML) estimates coincide. Similarly, the moments and ML estimates coincide for the Poisson and binomial demand incidence distributions. The estimates differ for the condensed Poisson distribution.

Potential efficiency gains were identified for the negative binomial and gamma distributions by using the method of maximum likelihood (ML). However, it was noted that ML may introduce bias problems for the gamma if the sample-size is small. It was also noted that the method of factorial moments may be highly inefficient for the log-zero-Poisson distribution and that there are large potential efficiency gains, although the bias properties of estimation methods are not well understood for this distribution.

Much further work remains to be done in this area. Computationally efficient methods for finding the maximum likelihood estimates have not been identified for the condensed negative binomial and the log-zero-Poisson distribution. Also, the bias properties of ML need to be examined for both of these distributions.

### 15.4.4 Empirical analysis of demand distributions

In chapter 11, an analysis of the goodness-of-fit of inter-purchase time, demand incidence and demand distributions was conducted using demand data from Pillar Engineering Supplies. The analysis showed strong support for the Poisson demand incidence model across all movement categories. The results for the condensed Poisson were not as good as for the Poisson distribution.

The analysis showed strong support for the negative binomial distribution (NBD), particularly when estimated using maximum likelihood. The results for the gamma distribution were also encouraging, although not as good for slow-moving SKUs as the NBD. The log-zero-Poisson distribution performed poorly when estimated by factorial moments; its goodness-of-fit was much improved by ML estimation.

## 15.5 Conclusions from part IV : demand variance laws

### 15.5.1 Variance law models

Variance estimates are required for all centralisation models, unless it is assumed that the variances are identical at all decentralised depots. However, the estimates may be difficult to obtain for products which are slow-moving, or which have been recently launched or which have experienced a step-change in mean demand. In these cases, a reliable mean estimate is easier to obtain and a relationship linking the variance to the mean of demand would be useful.

Stevens (1974) provided a theoretical justification for the quadratic variance law based on Poisson demand. However, his model was limited by the assumption of constant order-sizes. In chapter 12, a model was proposed based on non-constant order-sizes. A quadratic variance law was obtained under the following conditions :

* The ratio of the variance of mean demand incidence to the square of the expected mean demand incidence is independent of mean demand.
* Mean order-size is independent of mean demand incidence.
* The ratio of the variance of order-size to the square of mean order-size is independent of mean demand.
* The mean order-size is independent of mean demand.

The same form of result is obtained for one step ahead forecasts in the steady-state model, in which the mean demand incidence follows a random walk. The more general situation, in which n-step ahead forecasts are required in a steady state model, requires further research.

### 15.5.2 Empirical analysis of variance laws

In chapter 13, each of the four assumptions outlined above were tested empirically using the same data which was used for testing demand distributions. The first three assumptions were found to hold for the PES data. The fourth assumption did not hold; there was a weak, but statistically significant, correlation between mean order-size and a power of mean demand. Nevertheless, the quadratic variance law proved to be more robust than a more general variance law which included a term to allow for correlation between mean order-size and mean demand.

### 15.5.3 Application of variance laws in centralisation modelling

The implications of the variance laws for the centralisation of an SKU were examined. It was shown that the effect of mean order-size should not be assessed by using an Economic Order Quantity approach but simply by calculation of a weighted mean. The variance of centralised demand should be estimated by calculation of the pooled variance. These parameters may then be substituted in a formula involving linear and quadratic terms of mean demand, thus allowing an estimate of centralised demand variance to be made.

## 15.6 Trade-offs between inventory and transport costs

Leaving aside considerations of weight and volume, annual transport costs are greater for a fast-moving SKU than for a slower moving SKU, and the transport cost penalty of centralising inventory is correspondingly greater. The financial benefit in stock-reduction may be assessed using the methods suggested in this thesis. It was shown in chapter 14 that if the assumptions re-iterated in sub-section 15.5.1 hold and, more specifically, it is also assumed that :

1. The variance to mean-squared ratio of order-size, $\gamma_j$, for item *j* is the same at all decentralised depots

2. Mean order-sizes are identical at decentralised depots

3. Disturbance factors to mean demand incidence are directly proportional in size to the mean demand incidence itself

then the centralised demand variance is equal to :

$$(c_2)_0 \quad = \quad (\gamma + 1)\, b_1 \,\Sigma_j\, (c_1)_j \; + \; \frac{E(\eta_{kt}^2)}{\lambda_k^2} \; [\Sigma_j\, (c_1)_j]^2$$

| | | |
|---|---|---|
| where | $(c_2)_0$ | is the variance of centralised demand |
| | $(c_1)_j$ | is the mean demand at decentralised depot j |
| | $(b_1)$ | is the mean order-size at each decentralised depot |
| | $\gamma$ | is the variance to mean-squared ratio of order-size |
| | $\lambda_k$ | is the mean demand incidence at decentralised depot j |
| | $\eta_{kt}$ | is the disturbance to $\lambda_k$ at time t. |

The first term arises as a result of the assumed Poisson distribution of demand incidence. It represents the element of demand which is un-correlated between depots. The second term arises from the disturbance factors assumed to be acting in concert across depots but proportional in size to the mean demand incidence at each depot. This represents the element of demand which is correlated between depots.

As the mean demand rises, the second term in the expression becomes greater relative to the first. Now, the first term represents the part of demand which is uncorrelated between depots and from which a 'portfolio effect' predicts stock-savings by centralisation. On the other hand, the second term represents the correlated part of demand from which there are no stock-savings by centralisation. Hence, as the mean demand rises, the proportional stock-savings decrease.

The expression on the previous page allows the stock-savings by centralisation to be predicted for each SKU. This may then be compared to the additional incremental transport costs incurred by centralising the SKU, taking weight and volume into account if necessary. Thus, a decision rule may be developed, providing an operationalisation of variance law theory in inventory centralisation.

Estimation of safety stock savings may be made if a particular distribution is assumed, such as negative binomial. The square root law applies only as a special case, under the following conditions :

* demand is sufficiently fast and symmetrically distributed to be approximated by the normal distribution
* a safety factor may be applied to the standard deviation of demand, such as is the case when the 'probability of a stock out' is the service measure and demand is normally distributed, or when another measure is used but the safety factor approach yield a good approximation to the required safety stock
* demand is not correlated between decentralised depots
* the mean demand is the same at all decentralised depots.

The above assumptions are very restrictive, thus showing the need to examine safety stock centralisation in more depth than the square root law if a more accurate estimate is required.

## 15.7 Limitations and further research

### 15.7.1 Theoretical work

It is useful to summarise the limitations of the thesis which apply specifically to the theoretical work :

* The new classification system has not yet been used for models which were not included in its design.

* No work was done on filling gaps of model provision in the fields of safety stock centralisation and consolidation.

* The examination of the robustness of cycle stock centralisation formulae was restricted to cases where the EOQ assumptions hold.

* The general discussion of the robustness of the EOQ approach, in the context of cycle stock centralisation, was restricted to deviations from individual assumptions. No attempt was made to examine 'interaction effects'.

* The conjectural result on service benefits and disbenefits from centralisation has been only partially proved. Moreover, the result was restricted to cases where mean demands are equal at decentralised depots.

* The relationships between service-level measures were restricted to six measures.

* The analysis of demand distributions was limited to the NBD, CNBD, gamma

and lzP. Variable lead-time distributions such as the logarithmic-Poisson-gamma were not considered.

* Computationally efficient methods of identifying maximum likelihood estimators were not found for the CNBD and lzP distributions. No attempt was made to analyse the bias properties of these distributions.

* Variance law theory was limited to the 'stationary mean' model and the one step ahead 'steady state model'. The n-step ahead 'steady state' model was not analysed.

The final limitation must be addressed in order to incorporate the auto-correlation of demand over time, as well as the correlation of demand between depots. This is an important avenue of further research.

### 15.7.2 Empirical work

Demand incidence and demand distributions have been tested empirically using data from Pillar Engineering Supplies. Similarly, the components of a variance law, and the variance law itself, have been tested empirically. However, the predictions of a variance law for centralised stock-holding requirements have not been tested.

In the next stage of research, it is intended to compare the centralised stock-holdings predicted using methods presented in this thesis against the actual requirements of a company after centralisation.

# BIBLIOGRAPHY

*Note : The method of referencing used in this thesis follows the 'Harvard' system'*


ABRAHAMSSON, M (1993) Time based distribution, in *Quick response in the supply chain*, Unicom Seminars Ltd, Brunel Science Park, Uxbridge.

AGGARWAL, SC (1974) A review of current inventory theory and its applications. *Int J Prod Eng*, **12**, 443-482.

ANSCOMBE, FJ (1948) The transformation of Poisson, binomial and negative-binomial data. *Biometrika*, **35**, 246-254.

ANSCOMBE, FJ (1950) Sampling theory of the negative binomial and logarithmic series distributions. *Biometrika*, **37**, 358-382.

AUCAMP, DC (1982) Nonlinear freight costs in the EOQ problem. *EJOR*, **9**, 61-63.

AXSÄTER, S (1994) Using the deterministic EOQ formula in stochastic inventory control. *Paper presented at EURO XIII, Glasgow, July 1994.*

BAGCHI, U, HAYYA, JC and ORD, JK (1983) The Hermite distribution as a model of demand during lead-time for slow-moving items. *Decision Sciences*, **14**, 447-466.

BANERJEE, AK and BHATTACHARYYA, GK (1976) A purchase incidence model with inverse Gaussian interpurchase times. *J Amer Statist Ass*, **71**, 823-829.

BONDESSON, L (1978) On infinite divisibility of powers of a Gamma variable. *Scand Actuarial J*, **1978**, 48-61.

BOOKBINDER, JH and CHEN, VYX (1992) Multicriteria trade-offs in a warehouse / retailer System. *J Opl Res Soc*, **43**, 707-720.

BOYLAN, JE (1986) *Initial service-level targets for INFOREM*. Unpublished report, OR Centre, Unipart Group Ltd, Oxford.

BOYLAN, JE and JOHNSTON, FR (1994) Relationships between service level measures for inventory systems. *J Opl Res Soc*, **45**, 838-844.

BOYLAN, JE and JOHNSTON, FR (1996) Variance laws for inventory management. *Int J Production Economics*, **45**, 343-352.

BRETSCHNEIDER, S (1986) Estimating forecast variance with exponential smoothing. *Int J Forecasting*, **2**, 349-355.

BROWN, RG (1959) *Statistical forecasting for inventory control.* McGraw Hill, New York.

BROWN, RG (1963) *Smoothing, forecasting and prediction.* Prentice-Hall, Englewood Cliffs, New Jersey.

BROWN, RG (1967) *Decision rules for inventory management.* Holt, Reinhart and Winston, Chicago.

BÜHLMANN, H and BUZZI, R (1971) On a transformation of the weighted compound Poisson process. *ASTIN Bulletin,* **6**, 42-46.

BURGIN, TA (1975) The Gamma distribution and inventory control. *Opl Res Q,* **26**, 507-525.

BURGIN, TA and WILD, AR (1967) Stock control - experience and usable theory. *Opl Res Q,* **18**, 35-52.

BURNHAM, JM and MOHANTY, BB (1990) Requiem for EOQ - but Unified Order Quantity is alive and well. *Prod & Inv Mgmt J,* **31**, 80-83.

BURY, KV (1975) *Statistical models in applied science.* Wiley, New York.

BUZACOTT, JA (1975) Economic Order Quantities with inflation. *Opl Res Q,* **26**, 553-558.

CARLSON, ML and ROUSSEAU, JJ (1989) EOQ under date-terms supplier credit. *J Opl Res Soc,* **40**, 451-460.

CHANG, P-L and LIN, C-T (1991) On the effect of centralization on expected costs in a multi-location newsboy problem. *J Opl Res Soc,* **42**, 1025-1030.

CHATFIELD, C, EHRENBERG, ASC and GOODHARDT, GJ (1966) Progress on a simplified model of stationary purchasing behaviour [with discussion]. *J R Statist Soc, A,* **129**, 317-367.

CHATFIELD, C and GOODHARDT, GJ (1973) A consumer purchasing model with Erlang inter-purchase times. *J Amer Statist Ass,* **68**, 828-835.

CHEN, M-S and LIN, C-T (1989) Effects of centralization on expected costs in a multi-location newsboy problem. *J Opl Res Soc,* **40**, 597-602.

CHEN. M-S and LIN, C-T (1990) An example of disbenefits of centralized stocking. *J Opl Res Soc,* **41**, 259-262.

CHIKÁN, A [Ed] (1990) *Inventory models.* Akadémiai Kiadó, Budapest.

CLARK, CE (1957) Mathematical analysis of an inventory case. *Opns Res*, **5**, 627-643.

DAS, C (1978) A reappraisal of the square root law. *Int J Phys Distrib & Mat Mgmt*, **8**, 331-336.

DUNN, R, READER, S, and WRIGLEY, N (1983) An investigation of the assumptions of the NBD Model as applied to purchasing at individual stores. *Appl Statist*, **32**, 249-259.

EASTBURN, MP and NOCKOLD, CRT (1986) *Austin Rover Group European parts distribution study. Phase I report.* Unpublished report, PA Computers and Telecommunications Ltd, London.

EHRENBERG, ASC (1959) The pattern of consumer purchases. *Appl Statist*, **8**, 26-41.

EHRENBERG, ASC (1972) *Repeat buying: theory and applications*. North Holland.

EHRHARDT, R (1979) The power approximation for computing (s,S) inventory policies. *Mgmt Sci*, **25**, 777-786.

EHRHARDT, R and MOSIER, C (1984) A revision of the power approximation for computing (s,S) policies. *Mgmt Sci*, **30**, 618-622.

EILON, S (1964) Dragons in pursuit of the EBQ. *Opl Res Q*, **15**, 347-354.

EPPEN, GD (1979) Effects of centralisation on expected costs in a multi-location newsboy problem. *Mgmt Sci*, **25**, 498-501.

EVANS, T (1993) Pan European business systems - results of a KPMG survey, in *European distribution and logistics: the IT strategy*, Unicom Seminars Ltd, Brunel Science Park, Uxbridge.

EVERS, PT and BEIER, FJ (1993) The portfolio effect and multiple consolidation points: a critical assessment of the square root law. *J Business Logist*, **14**, 109-125.

FELLER, W (1957) *An introduction to probability theory and its applications*. John Wiley and Sons.

FERRAR, WL (1938) *A text-book of convergence*. Oxford University Press, Oxford.

GOODHARDT, GJ and CHATFIELD, C (1973) Gamma distribution in consumer purchasing, *Nature*, **244**, 316.

GOYAL, SK (1974) Determination of optimum packaging frequency of items jointly replenished. *Mgmt Sci*, **21**, 436-443.

GRAHN, GL (1969) NBD model of repeat-purchase loyalty: an empirical investigation. *J Marketing Res*, **6**, 72-78.

GREENWOOD, JA and DURAND, D (1960) Aids for fitting the gamma distribution by maximum likelihood. *Technometrics*, **2**, 55-65.

GUPTA, S (1988) Impact of sales promotions on when, what and how much to buy. *J Marketing Res*, **25**, 342-355.

GUPTA, S (1991) Stochastic models of interpurchase time with time-dependent covariates. *J Marketing Res*, **28**, 1-15.

HARRISON, PJ (1967) Exponential smoothing and short-term sales forecasting. *Mgmt Sci*, **13**, 821-842.

HERNITER, J (1971) A probabilistic market model of purchase timing and brand selection. *Mgmt Sci*, **18**, P102-P113.

HILL, RM (1990) Measuring stock service levels. *Int J Retail & Distribution Mgmt*. **18** (3), 26-30.

HUQUE, F (1973) *Adaptive estimators for the log-zero-Poisson distribution*. Unpublished doctoral dissertation, University of Missouri-Columbia, USA.

INGELHART, DL (1963) Optimality of (s,S) policies in the infinite horizon dynamic inventory problem. *Mgmt Sci*, **9**, 259-267.

JACOBS, RA and WAGNER, HM (1989) Lowering inventory systems costs by using regression-derived estimators of demand variability. *Decision Sciences*, **20**, 558-574.

JEULAND, AP, BASS, FM and WRIGHT, GP (1980) A multibrand stochastic model compounding heterogeneous Erlang timing and multinomial choice processes. *Opns Res*, **28**, 255-277.

JOHNSTON, FR (1973a) *Two approximations for inverting an incomplete Gamma function and the incomplete standard normal*. Working Paper No 11, School of Industrial and Business Studies, University of Warwick.

JOHNSTON, FR (1973b) *A note on the expected lost sales function*. Working Paper No 12, School of Industrial and Business Studies, University of Warwick.

JOHNSTON, FR (1975) *Forecasting in a real time stock control system*. Working Paper No 24, School of Industrial and Business Studies, University of Warwick.

JOHNSTON, FR (1980) An interactive stock control system with a strategic management role. *J Opl Res Soc*, **31**, 1069-1084.

JOHNSTON, FR and HARRISON, PJ (1986) The variance of lead-time demand. *J Opl Res Soc*, **37**, 303-309.

JOHNSTON, FR and MILNE, DJ (1972) A comparison of two stock control models. *Opl Res Q*, **23**, 573-577.

JOHNSTON, FR, TAYLOR, SJ and OLIVERIA, RMMC (1988a) Setting company stock levels. *J Opl Res Soc*, **39**, 15-21.

JOHNSTON, FR, TAYLOR, SJ and OLIVERIA, RMMC (1988b) Corrigendum. *J Opl Res Soc*, **39**, 615.

JONES, PCT and MOLLISON, JE [with a statistical note by MH Quenouille] (1948) A Technique for the Quantitative Estimation of Soil Micro-Organisms. *J Gen Microbiology*, **2**, 54-69.

KAHN, B and MORRISON, DG (1989) A note on 'random' purchasing: additional insights from Dunn, Reader and Wrigley. *Appl Statist*, **38**, 111-114.

KATTI, SK (1966) Inter-relations among generalised distributions and their components. *Biometrics*, **22**, 44-52.

KATTI, SK and RAO, AV (1970) The log-zero Poisson Distribution. *Biometrics*, **26**, 801-813.

KENDALL, M and STUART, A (1979) *The advanced theory of statistics, Vol 2 (4th Edition)*. Griffin, London & High Wycombe.

KHEDR, MS (1980) *Statistical inference on the log-zero-Poisson distribution.* Unpublished doctoral dissertation, University of Missouri-Columbia, USA.

KINGMAN, JFC (1993) *Poisson processes*. Clarendon Press, Oxford.

KINGSMAN, BG (1983) The effect of payment rules on ordering and stockholding in purchasing. *J Opl Res Soc*, **34**, 1085-1098.

KINGSMAN, BG (1991) EOQ under date-terms supplier credit: a near-optimal solution. *J Opl Res Soc*, **42**, 803-809.

KWAN, H W (1991) *On the demand distributions of slow-moving items.* Unpublished PhD thesis, University of Lancaster.

LAMBERT, DM and STOCK, JR (1993) *Strategic logistics management, 3rd Edn.* Irwin, Homewood Illinois.

LAU, H-S (1989) Toward an inventory control system under non-normal demand and lead-time uncertainty. *J Business Logist*, **10**, 88-103.

LUNDBERG, O (1964) *On random processes and their application to sickness and accident statistics (2nd Edition).* Almqvist and Wiksells, Boktryckeri AB, Sweden.

MAHMOUD, MM (1992) Optimal inventory consolidation schemes: a portfolio effect analysis. *J Business Logist,* **13**, 193-214.

MAISTER, DH (1976) Centralisation of inventories and the square root law. *Int J Phys Distrib*, **6**, 124-134.

MORRISON, DG and SCHMITTLEIN, DC (1981) Predicting future random events based on past performance. *Mgmt Sci*, **27**, 1006-1023.

NADDOR, E (1966) *Inventory systems.* John Wiley, New York.

NAHMIAS, S and DEMMY, WS (1982) The logarithmic Poisson Gamma distribution: a model for leadtime demand. *Naval Res Logist Q*, **29**, 667-677.

ORD, JK and BAGCHI, U (1983) The truncated normal-Gamma mixture as a distribution for lead time demand. *Naval Res Logist Q,* **30**, 359-365.

PAULL, AE (1978) A generalized compound Poisson model for consumer purchase panel data analysis. *J Amer Statist Ass*, **73**, 706-713.

PIETERS, EP, GATES, CE, MATIS, JH and STERLING, WL (1977). Small sample comparison of different estimators of negative binomial parameters. *Biometrics*, **33**, 718-723.

PILLAR ENGINEERING SUPPLIES LTD (1993). *Purchasing and stock management policy. Set up and control procedures.* Unpublished report, Pillar Engineering Supplies Ltd.

POPPER, KR (1972) *The logic of scientific discovery.* Hutchinson, London.

PORTEUS, EL (1985) Numerical comparison of inventory policies for periodic review systems. *Opns Res*, **33**, 134-152.

PRICHARD, JW and EAGLE, RH (1965) *Modern inventory management.* John Wiley and Sons.

QUENOUILLE, MH (1949) A relation between the logarithmic, Poisson and negative binomial series. *Biometrics*, **5**, 162-164.

RAMANAJUN, S (1911) Questions for solution, number 294. *J Indian Math Soc*, **3**, 128.

RAMANUJAN, S (1912) Question 294. *J Indian Math Soc,* **4**, 151-152.

RAO, AV (1969) *The log-zero Poisson distribution.* Unpublished doctoral dissertation, Florida State University, Tallahasse, Florida, USA.

RITCHIE, E and KINGSMAN, BG (1985) Setting stock levels for wholesaling: performance measures and conflict of objectives between supplier and stockist. *EJOR*, **20**, 17-24.

ROBERTS, D (1962) Approximations to optimal policies in a dynamic inventory model, in *Studies in applied probability and management science*, by Arrow, K, Karlin, S and Scarf, H (Eds), Stanford University Press, California.

RONEN, D (1982) Measures of product availability. *J Business Logist*, **3**, 45-58.

RONEN, D (1990) Inventory centralization / decentralization - the "Square Root Law" revisited again. *J Business Logist*, **11**, 129-138.

SANI, B and KINGSMAN, BG (1995) *Selecting the next periodic inventory control and demand forecasting methods for low demand items.* Unpublished paper, Lancaster University.

SCARF, H (1960) The Optimality of (s,S) policies in the dynamic inventory problem, in *Mathematical methods in the social sciences*, by Arrow, KJ, Karlin, S and Suppes, P, (Eds), Stanford University Press, California.

SCHRÖDINGER, E (1915) Zur Theorie der Fall- und Steigversuche an Teilchen mit Brownscher Bewegung. *Physikalische Zeitschrift*, **16**, 289-295.

SELEN, WJ and WOOD, WR (1987) Inventory cost definition in an EOQ model application. *Prod & Inv Mgmt J*, **28**, 44-47.

SESHADRI, V (1993) *The inverse Gaussian distribution. A case study in exponential families.* Clarendon Press, Oxford.

SHENTON, LR and BOWMAN, KO (1977) *Maximum likelihood estimation in small samples.* Griffin, London & High Wycombe.

SHERBROOKE, CC (1968) Discrete compound Poisson processes and tables of the geometric Poisson distribution. *Naval Res Logist Q*, **15**, 189-203.

SHERBROOKE, CC (1992) *Optimal inventory modeling of systems.* John Wiley, New York.

SILVER, EA (1975) Modifying the Economic Order Quantity (EOQ) to handle co-ordinated replenishments of two or more items. *Prod & Inv Mgmt J*, **16**, 26-38.

SILVER, EA (1976) Establishing the order quantity when the amount received is uncertain. *INFOR*, **14**, 32-39.

SILVER, EA and PETERSON, R (1985) *Decision systems for inventory management and production planning (2nd Edition).* John Wiley, New York.

STEVENS, CF (1974) On the variability of demand for families of items. *Opl Res Q*, **25**, 411-419.

STULMAN, A (1987) Benefits of centralized stocking for the multi-centre newsboy problem with first come, first served allocation. *J Opl Res Soc*, **38**, 827-832.

SZEGÖ, G (1928) Über Einige von S Ramanajun Gestellte aufgaben. *J London Math Soc*, **3**, 225-232.

TALLON, WJ (1993) The impact of inventory centralization on aggregate safety stock: the variable supply lead time case. *J Business Logist*, **14**, 185-203.

TATE, TB (1964) In defence of the Economic Batch Quantity. *Opl Res Q*, **15**, 329-340. (With responses by JL Burbridge and WE Duckworth, 341-345.)

TAYLOR, T (1991) Responsive delivery ! *OR Insight*, **4**, 14-18.

TAYLOR, CJ (1961) Application of the negative binomial distribution to stock control problems. *Opl Res Q*, **12**, 81-88.

TEICHER, H (1955) An inequality on Poisson probabilities. *Annals of Mathematical Statistics*, **26**, 147-149.

THORIN, O (1977a) On the infinite divisibility of the lognormal distribution. *Scand Actuarial J*, **1977**, 121-148.

THORIN, O (1977b) On the infinite divisibility of the Pareto distribution. *Scand Actuarial J*, **1977**, 31-40.

THORNLEY, G (1978) National distribution in F W Woolworth. *J Opl Res Soc*, **29**, 419-425.

TRIETSCH, T (1995) Revisiting ROQ:EOQ for company-wide ROI maximization. *J Opl Res Soc*, **46**, 507-515.

WATSON, GN (1929) Theorems stated by Ramanujan (V): approximations connected with $e^x$. *Proceedings of the London Mathematical Society, Series 2*, **29**, 293-308.

WEISS, EN (1990) Lot sizing is dead : long live lot sizing. *Prod & Inv Mgmt J*, **31**, 76-78.

WHEAT, RD (1987) *Determining the regularity of consumers' inter-purchase times: modelling, measurement and estimation issues.* Unpublished PhD thesis, Columbia University, USA.

WHEAT, RD (1987) *Determining the regularity of consumers' inter-purchase times: modelling, measurement and estimation issues.* Unpublished PhD thesis, Columbia University, USA.

WHEAT, RD and MORRISON, DG (1990) Estimating purchase regularity with two inter-purchase times. *J Marketing Res*, **27**, 87-93.

WOOLSEY, G (1988) A requiem for the EOQ: an editorial. *Prod & Inv Mgmt J*, **29**, 68-72.

WRIGHT, GG (1991) *An empirical analysis and stochastic modelling of aggregate demand behaviour in a spare parts inventory system.* Unpublished PhD thesis, City University, London.

ZHENG, Y-S (1992) On properties of stochastic inventory systems. *Mgmt Sci*, **38**, 87-103.

ZINN, W, LEVY, M and BOWERSOX, DJ (1989) Measuring the effect of inventory centralization / decentralization on aggregate safety stock: the "Square Root Law" revisited. *J Business Logist*, **10**, **1**, 1-14.

# APPENDICES

# APPENDIX 1.1

## Chikán's Secondary Codes

The following 'secondary codes' were defined by Chikán (1990) in addition to the ten 'main codes' of his classification system.

*Item codes*

* number of items (more specifically than in the first main code)
* connection between the items (substitute, complementary etc)
* type of inventory (raw materials, work-in-progress, finished goods etc)
* changes in the value of the item stocked (no change, depreciation etc)
* measurability of the items (continuously or discretely measurable)

*Store codes*

* number of stores in the model (more specifically than in the second main code)
* connection between the stores (eg parallel connection to a central store)
* storing capacity (unlimited, minimum limit, maximum limit)

*Input codes*

* delivery dates (single lot, several lots, dates known in advance or not etc)
* probability distribution of delivery dates (specified or not, and, if specified, which distribution)
* quantity of items arriving at one time (single item, whole lot, discrete given quantity etc)
* probability distribution of the quantities delivered during the re-order period (specified or not, and, if specified, which distribution)

*Demand (output) codes*

* dates of demand occurrence (demand in one lot, several lots etc)
* probability distribution of when demands arise (specified or not, and, if specified, which distribution)
* volume of demand arising at one time (unit demand, discrete demand etc)
* probability distribution of the volume of total demand arising during the re-order period (specified or not, and, if specified, which distribution)

*Code for 'dynamics'*

* periodicity of decisions made in the model (a single decision or a sequence of decisions)

*Decision variables codes*

* list of decision variables
* character of the set of values of decision variables (all discrete, all continuous, mixture of discrete and continuous)
* decision variables with a pre-fixed value (none, some)

*Systems and mathematics codes*

* relation of the inventory system to systems of higher level
* special constraints of the model (eg maximum permissible volume of shortage)
* computerized version of the model (computational technique specified or not)
* predominant mathematical apparatus (eg heuristic, dynamic programming etc)

*Shortages and orders codes*

* prescription of the re-order period (determined by model or specified)
* length of re-order periods (necessarily equal or not)

*Cost codes*

* list of unit cost factors (holding, shortage, ordering and other costs)
* dimensions of unit cost factors (eg $/unit time/unit quantity)
* character of unit costs (constant or functions of inventory level or time)
* changing of the purchase price of items of stock (eg purchase price depends on quantity ordered or time of ordering)

*Additional codes*

* mathematical preconditions (existing or not)
* optimal solution (none given, closed formula, iterative procedure etc)

# APPENDIX 3.1

## Mathematical Proofs of Results in Chapter 3

By definition of the portfolio quantity effect (PQE), the safety stock savings for super-consolidation and for several sub-consolidations are given below :

$$PQE_S = \Sigma_{i\in S}\, k\sigma_i - [\Sigma_{i\in S} k^2\sigma_i^2 + 2\Sigma\Sigma_{i<j;i,j\in S} k^2\sigma_i\sigma_j\rho_{ij}]^{1/2}$$

$$\Sigma_m\, PQE_{S(m)} = \Sigma_m\, \{\Sigma_{i\in S(m)}\, k\sigma_i - [\Sigma_{i\in S(m)} k^2\sigma_i^2 + 2\Sigma\Sigma_{i<j;i,j\in S(m)} k^2\sigma_i\sigma_j\rho_{ij}]^{1/2}\}$$

$$= \Sigma_{i\in S}\, k\sigma_i - \Sigma_m\, \{[\Sigma_{i\in S(m)} k^2\sigma_i^2 + 2\Sigma\Sigma_{i<j;i,j\in S(m)} k^2\sigma_i\sigma_j\rho_{ij}]^{1/2}\}$$

where $S = \cup\, S(m)$, $k$ is the common safety factor and the other variables are defined in the main part of the paper. Since the PQE measures safety-stock *savings*, super-consolidation is preferable to several sub-consolidations if and only if :

$$- [\Sigma_{i\in S} k^2\sigma_i^2 + 2\Sigma\Sigma_{i<j;i,j\in S} k^2\sigma_i\sigma_j\rho_{ij}]^{1/2} > - \Sigma_m\, \{[\Sigma_{i\in S(m)} k^2\sigma_i^2 + 2\Sigma\Sigma_{i<j;i,j\in S(m)} k^2\sigma_i\sigma_j\rho_{ij}]^{1/2}\}$$

$$[\Sigma_{i\in S} k^2\sigma_i^2 + 2\Sigma\Sigma_{i<j;i,j\in S} k^2\sigma_i\sigma_j\rho_{ij}]^{1/2} < \Sigma_m\, \{[\Sigma_{i\in S(m)} k^2\sigma_i^2 + 2\Sigma\Sigma_{i<j;i,j\in S(m)} k^2\sigma_i\sigma_j\rho_{ij}]^{1/2}\}$$

Cancelling the $k^2$ terms and squaring both sides,

$$\Sigma_{i\in S}\, \sigma_i^2 + 2\Sigma\Sigma_{i<j;i,j\in S}\, \sigma_i\sigma_j\rho_{ij} \quad < \quad \Sigma_m\, C(m) \;+\; 2\, \Sigma\Sigma_{m<n}\, [C(m)\, C(n)]^{1/2}$$

where $\quad C(m) = \Sigma_{i\in S(m)}\sigma_i^2 + 2\Sigma\Sigma_{i<j;i,j\in S(m)}\sigma_i\sigma_j\rho_{ij}$

Since $\quad \Sigma_{i\in S}\, \sigma_i^2 \; = \; \Sigma_m\, \Sigma_{i\in S(m)}\, \sigma_i^2,$

$$2\Sigma\Sigma_{i<j;i,j\in S}\, \sigma_i\sigma_j\rho_{ij} - \Sigma_m\, [2\Sigma\Sigma_{i<j;i,j\in S(m)}\, \sigma_i\sigma_j\rho_{ij}] \; < \; 2\, \Sigma\Sigma_{m<n}\, [C(m)\, C(n)]^{1/2}$$

By definition,

$$\Sigma\Sigma_{i<j\,(INTER)}\, \sigma_i\sigma_j\rho_{ij} \; = \; \Sigma\Sigma_{i<j}\, \sigma_i\sigma_j\rho_{ij} \; - \; \Sigma_m\, [\Sigma\Sigma_{i<j;i,j\in S(m)}\, \sigma_i\sigma_j\rho_{ij}]$$

Hence, the general condition for a super-consolidation to be preferred to several sub-consolidations is given by :

$$\Sigma\Sigma_{i<j\,(\mathrm{INTER})}\, \sigma_i\sigma_j\rho_{ij} \;<\; \Sigma\Sigma_{m<n}\, [C(m)\, C(n)]^{1/2}$$

Similarly, the condition for preferring several sub-consolidations is given by replacing the 'less than' inequality by a 'greater than' inequality in the above relation. Also, the indifference curve is given by using an 'equals' to replace the inequality sign.

In the special case of equal standard deviations of demand at all depots, the terms in $\sigma$ all cancel, giving the following condition for preferring a super-consolidation to several sub-consolidations :

$$\Sigma\Sigma_{i<j\,(\mathrm{INTER})}\, \rho_{ij} \;<\; \Sigma\Sigma_{m<n}\, [C(m)\, C(n)]^{1/2}$$

where

$$C(m) = \Sigma_{i\in S(m)}\, 1 + 2\Sigma\Sigma_{i<j;i,j\in S(m)}\, \rho_{ij}$$

$$= s_m + 2\Sigma\Sigma_{i<j;i,j\in S(m)}\, \rho_{ij}$$

and

$s_m$ = number of depots in the set S(m).

Hence, the condition for preferring a super-consolidation when the standard deviations of demand are equal is given by :

$$\Sigma\Sigma_{i<j\,(\mathrm{INTER})}\, \rho_{ij} \;<\; \Sigma\Sigma_{m<n}\, [(s_m+2\Sigma\Sigma_{i<j;i,j\in S(m)}\rho_{ij})(s_n+2\Sigma\Sigma_{i<j;i,j\in S(n)}\rho_{ij})]^{1/2}$$

As before, this condition may be modified for sub-consolidation preference and for indifference.

# APPENDIX 5.1

## Mathematical Proofs of Results in Chapter 5

Result for identical and unchanged parameters

Suppose there are n decentralised depots. Then,

$$ROI = \frac{nD\,(M - S/Q) - F - (J - I)\,nCQ/2}{L + nCQ}$$

$$\frac{d(ROI)}{dQ} = \frac{(L+nCQ)\,[\,nDS/Q^2 - (J-I)nC/2\,] - nC\,[nD(M - S/Q) - F - (J-I)\,nCQ/2\,]}{(L + nCQ)^2}$$

At maximum, first derivative is zero,

$$L\,\frac{nDS}{Q^2} + 2\,n^2\,\frac{CDS}{Q} - n^2\,CDM - n(J - I)\,\frac{CL}{2} + nCF = 0$$

$$\left(n\,CDM - CF + (J - I)\,\frac{CL}{2}\right) Q^2 - 2\,nCDS\,Q - LDS = 0$$

Taking the positive root of this quadratic equation, gives the ROQ for each decentralised depot, $ROQ_D$ :

$$ROQ_D = \frac{n\,CDS + \{n^2\,C^2D^2S^2 + CLDS\,[nDM - F + (J-I)\,L/2]\}^{1/2}}{C\,[\,nDM - F + (J-I)\,L/2\,]}$$

The ROQ at the centralised depot, $ROQ_C$, may be obtained by substituting nD (demand at centralised depot) into Trietsch's equation :

$$ROQ_C = \frac{n\,CDS + \{n^2C^2D^2S^2 + n\,CLDS\,[nDM - F + (J-I)\,L/2]\}^{1/2}}{C\,[\,nDM - F + (J-I)\,L/2\,]}$$

And so the ratio of decentralised to centralised cycle stock is given by :

$$\frac{\Sigma\, ROQ_D}{ROQ_C} = \frac{n^2\, CDS + n\, \{n^2C^2D^2S^2 + CLDS\, [nDM - F + (J-I)\, L/2]\}^{1/2}}{n\, CDS + \{n^2C^2D^2S^2 + n\, CLDS\, [nDM - F + (J-I)\, L/2]\}^{1/2}}$$

This may be re-expressed as,

$$\frac{\Sigma\, ROQ_D}{ROQ_C} = n \left( \frac{n\, CDS + \{n^2C^2D^2S^2 + CLDS\, [nDM - F + (J-I)\, L/2]\}^{1/2}}{n\, CDS + \{n^2C^2D^2S^2 + n\, CLDS\, [nDM - F + (J-I)\, L/2]\}^{1/2}} \right)$$

For ROI maximisation to be meaningful, the 'profit' must be positive. Therefore, the bracketed term [ nDM - F + (J-I) L/2 ] = Profit + nDS/Q must also be positive.

Hence,

$$\frac{\Sigma\, ROQ_D}{ROQ_C} < n$$

The ratio may also be expressed as,

$$\frac{\Sigma\, ROQ_D}{ROQ_C} = \sqrt{n} \left( \frac{n\sqrt{n}\, CDS + \{n^3C^2D^2S^2 + n\, CLDS\, [nDM - F + (J-I)\, L/2]\}^{1/2}}{n\, CDS + \{n^2C^2D^2S^2 + n\, CLDS\, [nDM - F + (J-I)\, L/2]\}^{1/2}} \right)$$

Hence,

$$\frac{\Sigma\, ROQ_D}{ROQ_C} > \sqrt{n}$$

# APPENDIX 6.1

## General Condition for Service Benefits / Disbenefits

***Theorem***

If demand is Poisson distributed with parameter $\mu$ at each decentralised depot and the conditions for Stulman's theorem are obeyed, then disbenefits arise from centralisation if and only if :

$$e^{-(k-1)\mu} \sum_{j=0}^{mk-1} \frac{(k\mu)^j}{j!}\left\{\left(\frac{k}{k-1}\right)^{mk-j} - 1\right\} > \left(\frac{k}{k-1}\right)^{mk} - \sum_{j=0}^{m} \frac{\mu^j}{j!}$$

where $k$ is the number of decentralised depots
$m$ is the number of items held in stock at each decentralised depot.

***Proof***

Let $P_C$ be the probability of a stock-out at the centralised depot
$P_D$ be the probability of a stock-out at a decentralised depot.

By Stulman's (1987) theorem,

$$P_C = \sum_{j=mk+1}^{\infty} \frac{e^{-k\mu}(k\mu)^j}{j!} - \left(\frac{k}{k-1}\right)^{mk} e^{-\mu} \sum_{j=mk+1}^{\infty} \frac{e^{-(k-1)\mu}[(k-1)\mu]^j}{j!}$$

And since demand is assumed to be Poisson,

$$P_D = \sum_{j=m+1}^{\infty} \frac{e^{-\mu}\mu^j}{j!}$$

These equations may be re-written :

$$P_C = 1 - \sum_{j=0}^{mk} \frac{e^{-k\mu}(k\mu)^j}{j!} - \left(\frac{k}{k-1}\right)^{mk} e^{-\mu}\left\{1 - \sum_{j=0}^{mk} \frac{e^{-(k-1)\mu}[(k-1)\mu]^j}{j!}\right\}$$

$$P_D = 1 - \sum_{j=0}^{m} \frac{e^{-\mu}\mu^j}{j!}$$

Disbenefits arise from centralisation if and only if $P_C > P_D$, ie:

$$-e^{-k\mu} \sum_{j=0}^{mk} \left\{ \frac{(k\mu)^j}{j!} - \left(\frac{k}{k-1}\right)^{mk} \frac{[(k-1)\mu]^j}{j!} \right\} - \left(\frac{k}{k-1}\right)^{mk} e^{-\mu} > - \sum_{j=0}^{m} \frac{\mu^j e^{-\mu}}{j!}$$

And so it follows that :

$$-e^{-(k-1)\mu} \sum_{j=0}^{mk} \left\{ \frac{(k\mu)^j}{j!} - \left(\frac{k}{k-1}\right)^{mk} \frac{[(k-1)\mu]^j}{j!} \right\} - \left(\frac{k}{k-1}\right)^{mk} > \sum_{j=0}^{m} \frac{\mu^j}{j!}$$

$$e^{-(k-1)\mu} \sum_{j=0}^{mk} \frac{\mu^j}{j!} \left\{ \frac{k^{mk}}{(k-1)^{mk-j}} - k^j \right\} > \left(\frac{k}{k-1}\right)^{mk} - \sum_{j=0}^{m} \frac{\mu^j}{j!}$$

$$e^{-(k-1)\mu} \sum_{j=0}^{mk} \frac{(k\mu)^j}{j!} \left\{ \left(\frac{k}{k-1}\right)^{mk-j} - 1 \right\} > \left(\frac{k}{k-1}\right)^{mk} - \sum_{j=0}^{m} \frac{\mu^j}{j!}$$

And since the j=mk term in the summation vanishes,

$$e^{-(k-1)\mu} \sum_{j=0}^{mk-1} \frac{(k\mu)^j}{j!} \left\{ \left(\frac{k}{k-1}\right)^{mk-j} - 1 \right\} > \left(\frac{k}{k-1}\right)^{mk} - \sum_{j=0}^{m} \frac{\mu^j}{j!}$$

# APPENDIX 6.2

## Limiting Conditions for Service Benefits / Disbenefits

*Theorem*

If demand is Poisson distributed with parameter $\mu$ at each decentralised depot and the conditions for Stulman's theorem are obeyed, then as the number of decentralised depots tends to infinity, disbenefits arise from centralisation if and only if

$$\sum_{j=0}^{m} \frac{\mu^j}{j!} > e^m$$

where m is the number of items held in stock at each decentralised depot.

*Proof*

To prove the theorem, we need to show that the term on the left hand side of the 'extended general condition' tends to zero as k tends to infinity. If this is proved, then the result also follows for the 'general condition' as a special case when m = 1. The left hand side of the 'extended general condition' is :

$$e^{-(k-1)\mu} \sum_{j=0}^{mk-1} \frac{(k\mu)^j}{j!} \left\{ \left( \frac{k}{k-1} \right)^{mk-j} - 1 \right\}$$

Since k/(k-1) > 1,

$$\left(\frac{k}{k-1}\right)^{mk-j} - 1 < \left(\frac{k}{k-1}\right)^{mk} - 1$$

and since $[k/(k-1)]^k$ is a monotonic decreasing function, $[k/(k-1)]^{mk}$ is also a monotonic decreasing function.

For all k > N,

$$\left(\frac{k}{k-1}\right)^{mk-j} - 1 < \left(\frac{N}{N-1}\right)^{mN} - 1 = \beta$$

Hence,

$$e^{-(k-1)\mu} \sum_{j=0}^{mk-1} \frac{(k\mu)^j}{j!} \left\{ \left(\frac{k}{k-1}\right)^{mk-j} - 1 \right\} < \beta\, e^{-(k-1)\mu} \sum_{j=0}^{mk-1} \frac{(k\mu)^j}{j!}$$

And now, using an argument based on Ferrar (1938),

$$e^{-(k-1)\mu} \sum_{j=0}^{mk-1} \frac{(k\mu)^j}{j!} = e^{\mu} \left\{ \frac{1}{e^{k\mu}} \sum_{j=0}^{mk-1} \frac{(k\mu)^j}{j!} \right\}$$

$$< e^{\mu} \left\{ \sum_{j=0}^{mk-1} \frac{M!}{(k\mu)^{\alpha}\, j!} \right\}$$

for an integer M which exceeds mk and $\alpha = M - mk$; it is also required that M is chosen so that $\alpha > 1$. And so,

$$e^{-(k-1)\mu} \sum_{j=0}^{mk-1} \frac{(k\mu)^j}{j!} < e^{\mu} \frac{mk\, M!}{(k\mu)^{\alpha}}$$

$$= \frac{m e^{\mu} M!}{\mu^{\alpha} k^{\alpha-1}}$$

As k increases indefinitely, so does $k^{\alpha-1}$ (since $\alpha > 1$). Since the other terms are independent of k,

As $k \to \infty$, $\quad e^{-(k-1)\mu} \sum_{j=0}^{mk-1} \frac{(k\mu)^j}{j!} \to 0$

And since $\beta$ is also independent of k,

As $k \to \infty$, $\quad e^{-(k-1)\mu} \sum_{j=0}^{mk-1} \frac{(k\mu)^j}{j!} \left\{ \left(\frac{k}{k-1}\right)^{mk-j} - 1 \right\} \to 0.$

Also ,

$$\lim_{k\to\infty}\left(\frac{k}{k-1}\right)^{mk} = \lim_{k\to\infty}\left(1+\frac{1}{k-1}\right)^{mk} = \lim_{k\to\infty}\left(1+\frac{1}{k}\right)^{mk} = e^{m}$$

Hence, as $k \to \infty$, disbenefits occur if and only if

$$\sum_{j=0}^{m} \frac{\mu^j}{j!} > e^m$$

If there is only one unit of stock at each decentralised depot (m=1), then disbenefits occur if and only if

$$\mu > e - 1.$$

# APPENDIX 6.3

## Identification of a Limiting 'Break-Point'

### *Theorem*

If demand is Poisson distributed with parameter $\mu$ at each decentralised depot and the conditions for Stulman's theorem are obeyed, then as the number of decentralised depots tends to infinity, the 'break point' for benefits / disbenefits, $\mu(\infty,m)$, satisfies the inequality :

$$\mu(\infty,m) \; > \; m + \ln(2)$$

and, consequently,

$$\sum_{j=0}^{m} \frac{e^{-\mu(\infty,m)} \, [\mu(\infty,m)]^{j}}{j!} \; < \; \frac{1}{2}$$

where $m$ is the number of items held in stock at each decentralised depot.

### *Proof*

Define the function $g_m(\mu)$ as follows :

$$g_m(\mu) \; = \sum_{j=0}^{m} \frac{\mu^j}{j!} \; - \; e^m$$

By the theorem in the previous appendix, the solution to the equation $g_m(\mu) = 0$ is the limiting 'break point'. It is clear that $g_m(\mu)$ is an increasing function of $\mu$ for fixed m. Therefore, to prove that the break-point exceeds $m + \ln(2)$, it is sufficient to prove that: $g_m(m + \ln(2)) \; < 0$.

The proof now proceeds in two stages. The first stage of the proof covers the values of m which are less than or equal to 17. The second stage of the proof covers all values greater than or equal to 18. The reason for this split is that, for $m \geq 18$, the result may be proved algebraically. An algebraic proof for $m \leq 17$ has not been found, but the result may be demonstrated by direct evaluation.

Stage 1 : $m \leq 17$

The function $g_m(\mu)$ has been evaluated at $\mu = m+\ln(2)$ for $m = 1,2,3,\ldots,17$. The results are shown below :

| m | $g_m(m + \ln(2))$ |
|---|---|
| 1 | - 0.025 |
| 2 | - 0.069 |
| 3 | - 0.177 |
| 4 | - 0.450 |
| 5 | - 1.147 |
| 6 | - 2.939 |
| 7 | - 7.570 |
| 8 | - 19.591 |
| 9 | - 50.906 |
| 10 | - 132.742 |
| 11 | - 347.198 |
| 12 | - 190.547 |
| 13 | - 2393.536 |
| 14 | - 6304.793 |
| 15 | - 16637.749 |
| 16 | - 43977.067 |
| 17 | - 116410.877 |

Since all of the above values are negative, the result is established for all positive integer values of m not exceeding 17.

Stage 2 : m ≥ 18

Using a result of Teicher (1955),

$$\sum_{j=0}^{m} \frac{(m + \ln(2))^j e^{-(m+\ln 2)}}{j!} = \sum_{j=0}^{m} \frac{m^j e^{-m}}{j!} - \int_{m}^{m+\ln 2} \frac{\lambda^m e^{-\lambda}}{m!} \, d\lambda$$

Since the function $\lambda^m e^{-\lambda}/m!$ is a monotonic decreasing function in $\lambda$, for fixed m, over the interval [m, m + ln(2)], it follows that,

$$\sum_{j=0}^{m} \frac{(m + \ln(2))^j e^{-(m+\ln 2)}}{j!} < \sum_{j=0}^{m} \frac{m^j e^{-m}}{j!} - \ln(2) \left( \frac{(m + \ln(2))^j e^{-(m+\ln 2)}}{m!} \right)$$

Ramanajun (1911, 1912) proposed the following result :

$$\sum_{j=0}^{m-1} \frac{m^j e^{-m}}{j!} = \frac{1}{2} - y \frac{m^m e^{-m}}{m!}$$

where the limits of y are 1/2 and 1/3 when m is 0 and ∞ respectively. Unable to prove the result, Ramanajun (1912) commented that, "*... yet it is difficult to prove that y lies between 1/2 and 1/3*". Proofs were given independently by Szegö (1928) and Watson (1929). Using Ramanajun's result,

$$\sum_{j=0}^{m} \frac{(m + \ln(2))^j e^{-(m+\ln 2)}}{j!} < \frac{1}{2} + (1-y) \frac{m^m e^{-m}}{m!} - \ln(2) \left( \frac{(m + \ln(2))^j e^{-(m+\ln 2)}}{j!} \right)$$

Hence, it is required to prove that the sum of the second and third terms on the right hand side of the inequality are negative, to establish the theorem.

Since y > 1/3,

$$(1 - y) \frac{m^m}{m!} \leq \frac{2}{3} \frac{m^m}{m!}$$

$$(1 - y)\frac{m^m}{m!} \le \frac{2}{3}\left(\frac{m}{m + \ln(2)}\right)^m \frac{(m + \ln(2))^m}{m!}$$

$$(1 - y)\frac{m^m e^{-m}}{m!} \le \frac{2}{3}\left(\frac{m + \ln(2)}{m}\right)^m 2\,e^{-(m+\ln 2)} \left(\frac{m}{m + \ln(2)}\right)^{m+1} \frac{(m + \ln(2))^m}{m!}$$

Now, $\left(\frac{m}{m + \ln(2)}\right)^{m+1} = \left(1 + \frac{\ln(2)}{m}\right)^{-(m+1)}$

is a monotonic increasing function with limit $e^{-\ln(2)} = 1/2$.

Hence, $\left(\frac{m}{m + \ln(2)}\right)^{m+1} < \frac{1}{2}$

And so it follows that,

$$(1 - y)\frac{m^m e^{-m}}{m!} \le \frac{2}{3}\left(\frac{m + \ln(2)}{m}\right)\left(\frac{(m + \ln(2))^m e^{-(m+\ln 2)}}{m!}\right)$$

All integers $m \ge 18$ satisfy the inequality,

$$m \ge \frac{\ln(2)}{(3/2 \ln(2) - 1)}$$

or, re-writing the inequality,

$$\frac{2}{3}\left(1 + \frac{\ln(2)}{m}\right) \le \frac{1}{2}$$

Hence,

$$(1 - y)\frac{m^m e^{-m}}{m!} \le \ln(2)\left(\frac{(m + \ln(2))^m e^{-(m+\ln 2)}}{m!}\right)$$

And so it is proved that,

$$\sum_{j=0}^{m} \frac{(m + \ln(2))^j e^{-(m+\ln 2)}}{j!} < \frac{1}{2}$$

Or, equivalently,

$$\sum_{j=0}^{m} \frac{(m + \ln(2))^j}{j!} - e^m < 0$$

# APPENDIX 7.1

## Mathematical Proofs of Results in Chapter 7

Proofs of Alternative Versions of the 'Complete Fill' Relationships

By definition of $l_f$, $l_u$ and $l_d$,

$$l_d = l_f + l_u$$

$$\frac{l_d}{o_d} = \frac{l_f}{o_d} + \frac{l_u}{o_d}$$

$$= \frac{l_f}{l_d}\frac{l_d}{o_d} + \frac{l_u}{o_u}\frac{o_u}{o_d}$$

And so, by definitions of $L_c$ and $O_c$,

$$\frac{l_d}{o_d} = L_c \frac{l_d}{o_d} + \frac{l_u}{o_u}(1 - O_c)$$

$$L_c = 1 - (1 - O_c)\frac{l_u}{o_u}\frac{o_d}{l_d}$$

The proofs of the relationships between $L_c$ and U and between $L_c$ and V are of the same form as the above.

Proof of Relationship between 'Partial Fill' Measures

By definition of $L_p$ and $O_p$,

$$\frac{L_p}{O_p} = \frac{l_{cf} + \alpha\, l_{pf}}{l_{cf} + (\alpha+\beta)\, l_{pf} + l_u} \; \frac{o_{cf} + (\Theta+\phi)\, o_{pf} + o_u}{o_{cf} + \Theta\, o_{pf}}$$

$$= \frac{l_{cf} + \alpha\, l_{pf}}{o_{cf} + \Theta\, o_{pf}} \; \frac{o_d + [\Theta+\phi-1]\, o_{pf}}{l_d + [\alpha+\beta-1]\, l_{pf}}$$

And so, by definition of $r_{cf}$, $r_{pf}$, $s_{cf}$ and $s_{pf}$,

$$\frac{L_p}{O_p} = \frac{l_d}{o_d} \; \frac{r_{cf} + \alpha r_{pf}}{s_{cf} + \Theta s_{pf}} \; \frac{o_d}{l_d} \; \frac{o_d + [\Theta+\phi-1]o_{pf}}{o_d} \; \frac{l_d}{l_d + [\alpha+\beta-1]l_{pf}}$$

$$= \frac{r_{cf} + \alpha r_{pf}}{s_{cf} + \Theta s_{pf}} \; \frac{1 + [\Theta+\phi-1]\, s_{pf}}{1 + [\alpha+\beta-1]\, r_{pf}}$$

<u>Proofs of Relationships between 'Partial Fill' and 'Complete Fill' Measures</u>

By definition of $L_c$ and $L_p$,

$$\frac{L_c}{L_p} = \frac{l_f}{l_d} \; \frac{l_{cf} + (\alpha+\beta)l_{pf} + l_u}{l_{cf} + \alpha\, l_{pf}}$$

$$= \frac{l_f}{l_d} \; \frac{l_d + [\alpha+\beta-1]l_{pf}}{l_{cf} + \alpha\, l_{pf}}$$

$$= \frac{l_f}{l_{cf} + \alpha l_{pf}} \; \{\, 1 + [\alpha+\beta-1](l_{pf}/l_d) \,\}$$

And so, by definition of $r_{cf}$ and $r_{pf}$,

$$\frac{L_c}{L_p} = \frac{l_f}{l_d} \; \frac{1}{r_{cf} + \alpha r_{pf}} \; \{\, 1 + [\alpha+\beta-1]r_{pf} \,\}$$

$$= \frac{r_{pf}}{r_{cf} + \alpha r_{pf}} \; \{\, 1 + [\alpha+\beta-1]r_{pf} \,\}$$

The proof of the relationship between $O_c$ and $O_p$ is of the same form as the above.

# APPENDIX 10.1

## Maximum Likelihood Estimation for the Condensed Poisson Distribution

By definition of the condensed-Poisson distribution,

$$p_0(\lambda) = e^{-2\lambda} + \lambda e^{-2\lambda}$$

$$= e^{-2\lambda}(1 + \lambda)$$

and for $k \geq 1$,

$$p_k(\lambda) = \frac{1}{2} e^{-2\lambda} \frac{(2\lambda)^{2k-1}}{(2k-1)!} + \frac{e^{-2\lambda}(2\lambda)^{2k}}{(2k)!} + \frac{1}{2} e^{-2\lambda} \frac{(2\lambda)^{2k+1}}{(2k+1)!}$$

$$= \frac{e^{-2\lambda}(2\lambda)^{2k}}{(2k)!} \left( \frac{2k}{4\lambda} + 1 + \frac{2\lambda}{2(2k+1)} \right)$$

Suppose that n observations have been obtained, m of which are zero. Then the likelihood function, $L(\lambda \mid \{x_j = 0; j = 1,...m\} \cup \{x_j > 0; j = m+1,...,n\})$ is given by :

$$L = e^{-2m\lambda}(1 + \lambda)^m \prod_{j=m+1}^{n} \left( \frac{e^{-2\lambda}}{(2x_j)!} (2\lambda)^{2x_j} \right) \left( \frac{x_j}{2\lambda} + 1 + \frac{\lambda}{2x_j + 1} \right)$$

$$\log L = -2m\lambda + m \log(1+\lambda) - 2(n-m)\lambda - \log[\Pi(2x_j)!] + \Sigma(2x_j) \log(2\lambda)$$

$$+ \sum_{j=m+1}^{n} \log\left( \frac{x_j}{2\lambda} + 1 + \frac{\lambda}{2x_j+1} \right)$$

$$\frac{d(\log L)}{d\lambda} = -2n + \frac{m}{1+\lambda} + \frac{1}{\lambda} \sum_{j=m+1}^{n} 2x_j + \sum_{j=m+1}^{n} \frac{1/(2x_j + 1) - x_j/2\lambda^2}{\lambda/(2x_j+1) + 1 + x_j/2\lambda}$$

So the ML estimate is a root of the following equation in $\lambda$ :

$$2n - \frac{m}{1+\lambda} = \frac{2}{\lambda} \sum_{j=m+1}^{n} x_j + \sum_{j=m+1}^{n} \frac{1/(2x_j + 1) - x_j/2\lambda^2}{\lambda/(2x_j+1) + 1 + x_j/2\lambda}$$

Using a similar approach to Anscombe (1950), it is clear that the left hand side must be greater than the right hand side if $\lambda$ is large enough and the right hand side must be greater than the left hand side if $\lambda$ is small enough, provided that there are some non-zero observations ($m < n$). Since both sides are continuous functions for $\lambda > 0$, the equation must have at least one positive finite root. However, it has not been proven that there is only one root of the equation.

# APPENDIX 12.1

## Variance of distributions based on condensed Poisson demand incidence

In this appendix, the characteristic function of the condensed-Poisson distribution is derived for the case where the mean value is 'mixed over time', but with orders for single items only. The derivation is then extended to cover general order-size distributions.

By definition of the condensed-Poisson distribution,

$$p_0(\lambda t) = e^{-2\lambda t} + \lambda t\, e^{-2\lambda t}$$

and for $k \geq 1$,

$$p_k(\lambda t) = \frac{1}{2} e^{-2\lambda t} \frac{(2\lambda t)^{2k-1}}{(2k-1)!} + \frac{e^{-2\lambda t} (2\lambda t)^{2k}}{(2k)!} + \frac{1}{2} e^{-2\lambda t} \frac{(2\lambda t)^{2k+1}}{(2k+1)!}$$

Integrating over all possible values of the mean , the characteristic function, $\delta_t(u)$, is given by :

$$\delta_t(u) = \int [e^{-2\lambda t} + e^{-2\lambda t}\lambda t]\ dS(\lambda)$$

$$+ \Sigma_{k=1}^{\infty} \int e^{iuk} \left\{ \frac{1}{2} e^{-2\lambda t} \frac{(2\lambda t)^{2k-1}}{(2k-1)!} + \frac{e^{-2\lambda t} (2\lambda t)^{2k}}{(2k)!} + \frac{1}{2} e^{-2\lambda t} \frac{(2\lambda t)^{2k+1}}{(2k+1)!} \right\} dS(\lambda)$$

$$= \int e^{-2\lambda t} dS(\lambda) + \Sigma_{k=1}^{\infty} e^{iuk} \int e^{-2\lambda t} \frac{(2\lambda t)^{2k}}{(2k)!} dS(\lambda)$$

$$+ \frac{1}{2} \Sigma_{k=1}^{\infty} \int e^{-2\lambda t} \frac{(2\lambda t e^{iu/2})^{2k-1}}{(2k-1)!} e^{iu/2} dS(\lambda) + \frac{1}{2} \Sigma_{k=0}^{\infty} \int e^{-2\lambda t} \frac{(2\lambda t e^{iu/2})^{2k+1}}{(2k+1)!} e^{-iu/2} dS(\lambda)$$

[Note that the $e^{-2\lambda t}\lambda t$ term has been taken into the final summation, extending it from k=1 to k=0].

$$\delta_t(u) = \sum_{k=0}^{\infty} \int e^{-2\lambda t} \frac{(2\lambda t e^{iu/2})^{2k}}{(2k)!} dS(\lambda) + \frac{1}{2}(e^{iu/2} + e^{-iu/2}) \sum_{k=0}^{\infty} \int e^{-2\lambda t} \frac{(2\lambda t e^{iu/2})^{2k+1}}{(2k+1)!} dS(\lambda)$$

$$= \int e^{-2\lambda t} \cosh(2\lambda t\, e^{iu/2})\, dS(\lambda) \;+\; \cos(u/2) \int e^{-2\lambda t} \sinh(2\lambda t\, e^{iu/2})\, dS(\lambda)$$

To extend the derivation to a general order-size distribution, consider $\chi$, the characteristic function of the order-size distribution. Using a result proved by Bühlmann and Buzzi (1971), the overall characteristic function, $\phi_t(u)$, is given by :

$$\phi_t(u) = \delta_t \left\{ \frac{\log \chi(u)}{i} \right\}$$

Hence,

$$\phi_t(u) = \int e^{-2\lambda t} \cosh[2\lambda t \chi(u)^{1/2}]\, dS(\lambda) + \frac{1}{2}[(\chi(u)^{1/2} + \chi(u)^{-1/2}] \int e^{-2\lambda t} \sinh[2\lambda t \chi(u)^{1/2}]\, dS(\lambda)$$

Now the above result will be used to establish a relationship between the mean and variance of demand. The derivation will be based on the assumption of compounding, with no compounding as a special case. The result for another special case of interest will be given, namely gamma distributed mean demands over time.

Taking the first derivative of the characteristic function,

$$\phi_t'(u) = \chi(u)^{-1/2}\chi'(u) \int \lambda t\, e^{-2\lambda t} \sinh[2\lambda t\, \chi(u)^{1/2}]\, dS(\lambda)$$

$$+ \frac{1}{2} \left\{ \frac{1}{2} \chi(u)^{-1/2} \chi'(u) - \frac{1}{2} \chi(u)^{-3/2} \chi'(u) \right\} \int e^{-2\lambda t} \sinh[2\lambda t\, \chi(u)^{1/2}]\, dS(\lambda)$$

$$+ \frac{1}{2} [\chi(u)^{1/2} + \chi(u)^{-1/2}]\;\; \chi(u)^{-1/2} \chi'(u) \int \lambda t\, e^{-2\lambda t} \cosh[2\lambda t\, \chi(u)^{1/2}]\, dS(\lambda)$$

Evaluating the first derivative at zero, and substituting $\chi(0) = 1$ and $\chi'(0) = ib_1$, where

$b_1$ is the mean order-size, the second term vanishes and we have :

$$\phi_t'(0) = ib_1 \int \lambda t\, e^{-2\lambda t} \sinh(2\lambda t)\, dS(\lambda) + ib_1 \int \lambda t\, e^{-2\lambda t} \cosh(2\lambda t)\, dS(\lambda)$$

$$= ib_1 \int \lambda t\, dS(\lambda)$$

Hence,

$$\text{Mean} = b_1 \int \lambda t\, dS(\lambda)$$

Taking the second derivative of the characteristic function,

$$\phi_t''(u) = -\frac{1}{2} \chi(u)^{-3/2} \chi'(u)^2 \int \lambda t\, e^{-2\lambda t} \sinh[2\lambda t\, \chi(u)^{1/2}]\, dS(\lambda)$$

$$+ \chi(u)^{-1/2} \chi''(u) \int \lambda t\, e^{-2\lambda t} \sinh[2\lambda t\, \chi(u)^{1/2}]\, dS(\lambda)$$

$$+ \chi(u)^{-1} \chi'(u)^2 \int (\lambda t)^2 e^{-2\lambda t} \cosh[2\lambda t\, \chi(u)^{1/2}]\, dS(\lambda)$$

$$+ \frac{1}{4} \chi''(u) [\chi(u)^{-1/2} - \chi(u)^{-3/2}] \int e^{-2\lambda t} \sinh[2\lambda t\, \chi(u)^{1/2}]\, dS(\lambda)$$

$$+ \frac{1}{4} \chi'(u)^2 \left\{ -\frac{1}{2} \chi(u)^{-3/2} + \frac{3}{2} \chi(u)^{-5/2} \right\} \int e^{-2\lambda t} \sinh[2\lambda t\, \chi(u)^{1/2}]\, dS(\lambda)$$

$$+ \frac{1}{4} \chi'(u)^2 [\chi(u)^{-1/2} - \chi(u)^{-3/2}]\ \chi(u)^{-1/2} \int \lambda t\, e^{-2\lambda t} \cosh[2\lambda t\, \chi(u)^{1/2}]\, dS(\lambda)$$

$$+ \frac{1}{2} \chi''(u) [1 + \chi(u)^{-1}] \int \lambda t\, e^{-2\lambda t} \cosh[2\lambda t\, \chi(u)^{1/2}]\, dS(\lambda)$$

$$+ \frac{1}{2} \chi'(u)^2 [-\chi(u)^{-2}] \int \lambda t\, e^{-2\lambda t} \cosh[2\lambda t\, \chi(u)^{1/2}]\, dS(\lambda)$$

$$+ \frac{1}{2} \chi'(u)^2 [1 + \chi(u)^{-1}]\ \chi(u)^{-1/2} \int (\lambda t)^2 e^{-2\lambda t} \sinh[2\lambda t\, \chi(u)^{1/2}]\, dS(\lambda)$$

Evaluating the second derivative at zero, the fourth and sixth term vanish. The first and eighth terms combine; similarly, the second and seventh terms combine and the third and the ninth terms combine. After collecting terms together,

$$\phi_t''(0) = b_1^2\left\{\frac{1}{2}\int \lambda t \, dS(\lambda) - \int (\lambda t)^2 \, dS(\lambda)\right\} + \chi''(0)\int \lambda t \, dS(\lambda)$$

$$- \frac{1}{4} b_1^2 \int e^{-2\lambda t} \sinh(2\lambda t) \, dS(\lambda)$$

Hence,

$$\mathrm{Var} = b_1^2\left\{\int (\lambda t)^2 \, dS(\lambda) - \frac{1}{2}\int \lambda t \, dS(\lambda)\right\} - \chi''(0)\int \lambda t \, dS(\lambda)$$

$$+ \frac{1}{4} b_1^2 \int e^{-2\lambda t} \sinh(2\lambda t) - b_1^2 \left\{\int \lambda t \, dS(\lambda)\right\}^2$$

$$= [-\chi''(0) - \mu^2] \int \lambda t \, dS(\lambda) + \frac{1}{2} b_1^2 \int \lambda t \, dS(\lambda)$$

$$+ b_1^2 \int (\lambda t)^2 \, dS(\lambda) + \frac{1}{4} b_1^2 \int e^{-2\lambda t} \sinh(2\lambda t) \, dS(\lambda) - b_1^2\left\{\int \lambda t \, dS(\lambda)\right\}^2$$

$$\mathrm{Var} = \left(b_2 + \frac{b_1^2}{2}\right) E(\lambda t) + \frac{b_1^2}{4}\int e^{-2\lambda t} \sinh(2\lambda t) \, dS(\lambda) + b_1^2 \, \mathrm{Var}(\lambda t)$$

where $b_2$ is the variance of order-size.

This concludes the proof of the formula for the general case. The first special case to be examined is 'no compounding'. In this case, all orders are for single items; ie $b_1=1$ and $b_2=0$. Substituting in the general formula,

$$\mathrm{Var} = \frac{1}{2} E(\lambda t) + \frac{1}{4}\int e^{-2\lambda t} \sinh(2\lambda t) \, dS(\lambda) + \mathrm{Var}(\lambda t)$$

The second special case relates to gamma distributed mean demands. The following results for the gamma distribution are standard :

$$E(\lambda t) = \frac{k}{\alpha} t \qquad \mathrm{Var}(\lambda t) = \frac{k}{\alpha^2} t^2$$

where k is the shape parameter and $\alpha$ is the scale parameter.

In the case of the gamma distribution,

$$\int e^{-2\lambda t} \sinh(2\lambda t)\, dS(\lambda) = \frac{1}{2}\int dS(\lambda) - \frac{1}{2}\int e^{-4\lambda t}\, dS(\lambda)$$

$$= \frac{1}{2} - \frac{1}{2}\int e^{-4\lambda t}\left\{\frac{\alpha^k}{\Gamma(k)} e^{-\alpha\lambda}\lambda^{k-1}\right\} d\lambda$$

$$= \frac{1}{2} - \frac{1}{2}\left\{\frac{\alpha}{\alpha+4t}\right\}^k \int \frac{(\alpha+4t)^k}{\Gamma(k)} e^{-(\alpha+4t)\lambda}\lambda^{k-1}\, d\lambda$$

$$= \frac{1}{2} - \frac{1}{2}\left\{\frac{\alpha}{\alpha+4t}\right\}^k$$

Substituting the expression for $E(\lambda t)$ in the general formula for the mean,

$$\text{Mean} = b_1 \frac{k}{\alpha} t$$

Substituting the above expression and the expressions for $E(\lambda t)$ and $Var(\lambda t)$ in the general variance formula derived earlier in the appendix,

$$\text{Var} = \left(b_2 + \frac{b_1^2}{2}\right)\frac{k}{\alpha} t + \frac{b_1^2}{8}\left\{1 - \left(1+\frac{4t}{\alpha}\right)^{-k}\right\} + b_1^2 \frac{k}{\alpha^2} t^2$$