

Operational Statistics for Improved Decision Making Policies

L. HANSEN

School of Computing and Mathematics
University of Western Sydney
Locked Bag 1797 Penrith South DC, NSW 1797, Australia

ABSTRACT

The primary purpose of this paper is to outline the new developments in the area of finding better decision making policies. First it introduces the operational statistics approach and then the BMIT concept showing how operational statistics fits into the big framework of integrated efforts in achieving better decision making policies. Secondly this paper introduces the idea of operational statistics and show how it differs from traditional approaches. In operational statistics, one integrates the estimation and the optimization tasks so as to estimate the optimal policy directly. When using traditional methods and when the data available to estimate the distributional assumptions are limited, the loss in expected gain can be significant. We demonstrate the power of operational statistics concept by showing that the traditional approach which separates the estimation and the optimization tasks in inventory control problem with an ambiguous demand can lead to a suboptimal solution. When the data available on the random demand is limited, the loss in expected profit can be substantial. Finally this paper introduce the concept of BMIT and illustrate and highlight the roles in different key stakeholders of BMIT and show how they integrate to give a dynamic and powerful framework which will revolutionize the research efforts and capabilities of many application domains where there are huge volumes of data and information is available. This knowledge discovery process has assumed many identities and forms in recent years such as E-Research and Analytics. Coming from the applications arena, such examples include, E-Health, E-Science etc. and the general theme could assume names such as Data Informatics, Operational Statistical Informatics or Decision Making Informatics to name a few. Further contributing to education and curriculum development, the success of BMIT and the lessons learned will be discussed. Future developments underway, is also noted.

Key Words: Data informatics, Decision making informatics, Demand ambiguity. Estimation and optimization, Model uncertainty. Operational statistical informatics, Operational statistics

INTRODUCTION

Operational statistics integrates operational research procedures with statistical methodologies for deriving better decision making policies. This notion was first discovered when developing the Bachelor of Mathematics and Information Technology (BMIT) integrated degree (which was designed in 1999 and first offered in 2000) with the focus of transforming data into knowledge. In contrast to the traditional mathematics and computer science integrated courses, BMIT was designed for the purpose of achieving data driven optimum decision making policies and to solve real life problems in a larger scale leading to fulfilling the community needs. This is one of the first such courses of this type, where the integration of these complementary skills statistics, operations research, simulation techniques, data base, data integration and processing, was done through data mining. Since data to knowledge was the key and with the additional data integration capabilities contributed by IT, BMIT build the foundation for a flexible research platform with a multiple and dynamic research possibilities for the vast application domains from which the data is derived.

In integrating operations research and statistics, Liyanage and Shanthikumar (2005) proposed a new approach - operational statistics - which derives a decision rule that maximises the performance uniformly for all possible values of the unknown input parameters. This technique generalises the traditional approach of separating the parameter estimation and the maximization of the expected outcome which leads to suboptimal decision making policies in comparison. Operational statistics provides a better solution by integrating the estimation and the optimization tasks together in deriving at an improved optimal decision making policy. Its effectiveness is demonstrated by applying this theory to the newsvendor inventory control problem with an ambiguous demand.

This paper gives the overview of the development of this integrated approach, its successes and lessons learned. The key features includes, data base, data integration, data mining, machine learning, statistics, operations research, operational statistics, simulation techniques and most importantly the application domains such as, health, public health, science, environment, climate change, bio-security, psychology, education, business, management, marketing, finance and lots more.

OPERATIONAL STATISTICS

Traditional Approaches: Separating Estimation and Optimization

We will first introduce the traditional way we use statistics and operations research. First one would formulate the problem assuming that a full stochastic characterization is available. It will then be solved (that is, analyzed and optimized). Finally, the parameters needed for implementing the optimal policy will be estimated using standard statistical techniques such as maximum likelihood, least squared deviation and method of moments estimation procedures. In classical statistics, one hopes that such procedures will result in an unbiased and at the least a strongly consistent estimator. Note the dependence of the input distribution $F(x; \theta)$ on its parameter(s) θ .

Traditional approach separates the estimation and the optimization tasks. When the input distribution is unknown, one may either use a parametric approach where it is assumed that the input distribution belongs to a parametric family of distributions or a non-parametric approach where no assumption regarding the parametric form of the unknown input distribution is made.

Parametric Approach: Separated Parameter Estimation and Optimization

Under the parametric approach, one may choose to estimate the unknown parameters θ or choose a prior distribution for the unknown parameters and apply the Bayesian approach to incorporating the input data available. For example Scarf (1959) and Karlin (1960) consider a Bayesian framework for the unknown demand distribution. Specifically, assuming that the demand distribution belongs to the family of exponential distributions, the demand process is characterized by the prior distribution on the unknown parameter. Further extension of this approach is presented in Azoury (1985). Application of the Bayesian approach to the censored demand case is given in Ding *et al.* (2002) and Lariviere and Porteus (1999). Parameter estimation is first considered in Conrad (1976) and recent developments are reported in Agrawal and Smith (1996) and Nahmias (1994).

Further these nonparametric Bayesian approaches learn from the past to predict future performance, these approaches becomes part of the BMIT framework. Finding better strategies in making improved decisions lead to discovering operational statistics – approach.

Non Parametric Approach: Separated Estimation of Empirical Distribution and Optimization

Within the non-parametric approach, either the empirical distribution (see Liyanage and Shanthikumar, 2005, Section 3) or the bootstrapping method (e.g. see Bookbinder and Lordahl, 1989) can be applied with the available demand data to obtain an inventory control policy.

A third alternative to dealing with the unknown input distribution is when the input variable is partially characterized by its moments. For example, when the unknown demand distribution is characterized by the first two moments, Scarf (1958) derives a robust min–max inventory control policy. Further development and review of this model is given in Gallego and Moon (1993).

Further these nonparametric approaches also learn from the past to predict future performance, these approaches also complements and added to the BMIT framework. Similarly, finding better strategies in making improved decisions and the availability of faster and cheaper disk storage technology, lead to discovering operational statistics – approach (Berry and Linoff 2000; Hand *et al.*, 2000).

Definition of Operational Statistics

In operational statistics, (as mentioned before) one integrates the estimation and the optimization tasks so as to estimate the optimal policy directly.

Liyanage and Shanthikumar (2005) introduced operational statistics and demonstrated it for exponentially distributed input variables. It is shown that the optimal policy derived using operational statistics is strictly better (in a priori expected profit) than those derived using either the parametric or the non-parametric approaches.

In this paper we consider the news vendor inventory control problem, where it is known that the demand distribution function belongs to a parameterized family of distribution functions. However, unlike in the Bayesian approach, we do not assume any prior knowledge on the parameter values. We can, however, infer the parameter values using the demand data $\{D_k, k = 1, 2, \dots, n\}$.

Application: The Newsvendor Inventory Control Problem with an Ambiguous Demand

To demonstrate the concept of - Operational Statistics -, we will consider the single period newsvendor inventory control problem. Items are purchased at c per unit and sold at s ($s > c$) per unit. Without loss of generality, it will be assumed that the salvage value of the unsold items is zero. $\{D_k, k = 1, 2, \dots, n, n + 1\}$ is the sequence of *i.i.d.* demand for the items with distribution function FD . We will assume, for ease of exposition, that the demand distribution is continuous. We have observed the demand $\{D_k, k = 1, 2, \dots, n\}$ for the first n periods and we wish to find the optimal order quantity for the $(n + 1)$ th period. The demand distribution FD is unknown. Analysis of inventory control problems with unknown demand distribution is not new. Several important papers have appeared in the literature. When the demand distribution is unknown, traditionally one may either use a parametric approach (where it is assumed that the demand distribution belongs to a parametric family of distributions) or a non-parametric approach (where no assumption regarding the parametric form of the unknown demand distribution is made). Next we introduced the improved – operational statistics – approach and show it generates better outcome.

Inventory Control Using Operational Statistics

In operational statistics, an optimal order quantity, rather than the parameters of the distribution, is directly estimated from the data. The optimal order quantity is estimated in such a way that a priori expected profit is maximized. For this, we define a statistic S of the data $\{D_1, D_2, \dots, D_n\}$ parameterized by some optimization variables, say z , within an acceptable range Z . Then a priori expected profit is then maximized with respect to z . Let

$$\hat{X}(z) = S(D_1, D_2, \dots, D_n; z)$$

be the order quantity estimated from the data $\{D_1, D_2, \dots, D_n\}$ with the optimization parameters z . Let $\eta(z)$ be the a priori expected profit for the order quantity $\hat{X}(z)$. That is,

$$\eta(z) = E[\phi(\hat{X}(z), \theta)].$$

If $z^* = \arg \max \{ \eta(z) : z \in Z \}$, then the optimal order quantity for the class of ordering policies $\hat{X}(z)$ is $\hat{X}(z^*)$. Motivated by the two order policies estimated with the sample mean given in Liyanage and Shanthikumar (2005) in Section 2 and empirical distribution in Section 3, we first applied the operations statistics idea with the following two classes of statistics:

$$\hat{X}(z) = z\bar{D}, \quad z \geq 0 \quad \text{and} \quad \hat{X}(r, a) = D_{\lfloor r-1 \rfloor} + a(D_{\lfloor r \rfloor} - D_{\lfloor r-1 \rfloor}).$$

$$r \in \{1, \dots, n\}, a \geq 0$$

Operational statistics using three different approaches are considered, (i) using the sample mean, (ii) using the empirical distribution, and (iii) using sample spacing in Liyanage and Shanthikumar (2005) and have proved that the priority expected profit is optimum compared to the traditional approaches.

Though we have illustrated our idea of operational statistics for exponentially distributed demand, this approach can be applied to any demand distribution, where the only unknown parameter is the scale parameter. For example we can apply this approach to demand with a (i) Gamma distribution with an unknown scale parameter (with a known shape parameter), (ii) normal distribution with an unknown variance (with a known mean) and (iii) uniform distribution on $[0, b]$ with an unknown b . Extensions of this approach to other distributions and to dynamic decision making is reported in Lim *et al.*, (2006). Operations statistics is similar to decision theory (eg see Azoury, 1985; Lariviere and Porteus, 1999). The differences in these two approaches are given in Lim *et al.*, (2006). Applications of operational statistics to static and dynamic problems with non i.i.d. historic data and the fact that the proposed approach can be implemented using a computationally efficient simulation makes this a key component of the BMIT concept -data to knowledge for better decision making -. The link of operational statistics to BMIT or decision making informatics framework is illustrated in the next section.

BMIT Motivation and Key Components: Machine Learning, Data Mining, Operational Statistics and Data Base, Simulation Techniques with other Application Domains.

Motivation of BMIT and the significance of Data Mining and Visualization within BMIT

Data Mining or Machine Learning was developed around mid 1990's. In addition to the technology improvements the development of the world-wide-web added to the availability and accessibility of massive sets of data (Berry and Linoff, (2000); Grossman *et al.*, 1999; Hand *et al.*, 2000). This united operations research scientists, applied mathematicians and statisticians who traditionally learn from data with the newly formed data mining or machine learning practitioners by the common goal 'change data into information' (Friedman, 1997; Hestie *et al.*, 2001). This led to a need to integrate these approaches. Statisticians provide the theoretical basis for understanding existing machine learning algorithms and for developing new methodologies (Hand, 1999). Computing and IT specialist, providing the data base, data storing, warehousing and programming skills needed to carry out all phases of data mining tasks well (Berry and Linoff, 2000).

For statisticians, operations researchers or applied mathematicians data mining approaches became the first step in extracting information from data. Hence Statistics, Operations Research and Applied Mathematics skills enhanced the capabilities of a data mining specialist providing opportunity to take the modeling and analysis tasks further into greater depth (Babinec, 2000; Friedman, 1997; Hand, 1998; Hand, 1999; Hestie *et al.*, 2001). However, without skills such as data base and data warehousing a statistician or mathematician will become restricted in applying data mining approaches well. Clearly Data Mining became the integrating unit of these skills while highlighting the strength in this integration.

Thus Data Mining and Visualization played a significant role in acting as the common and integrating unit within the BMIT. After completion BMIT students are able to perform all six phases, business/research understanding, data understanding, data preparation, modeling and deployment of the CRISP-DM process (Berry and Linoff, 2000). In the deployment phase, students in the BMIT have additional skills needed to produce Decision Support Tools specially designed for the given applications based on the obtained results from the modelling phase. Hence the Data Mining unit sits within the BMIT, demonstrating how well the Mathematics and IT skills integrate and thus promoting the BMIT concept. Further BMIT provides a course to train well rounded data mining specialists needed by industry promoting Data Mining concept.

BMIT Concept

BMIT concept can be described as New Methods to analyze

- Massive
- Multi-dimensional
- Multi-source
- Time-varying

information streams **to make effective decisions** in **time critical** situations. The major components of this knowledge discovery process are illustrated in the Figure 1.

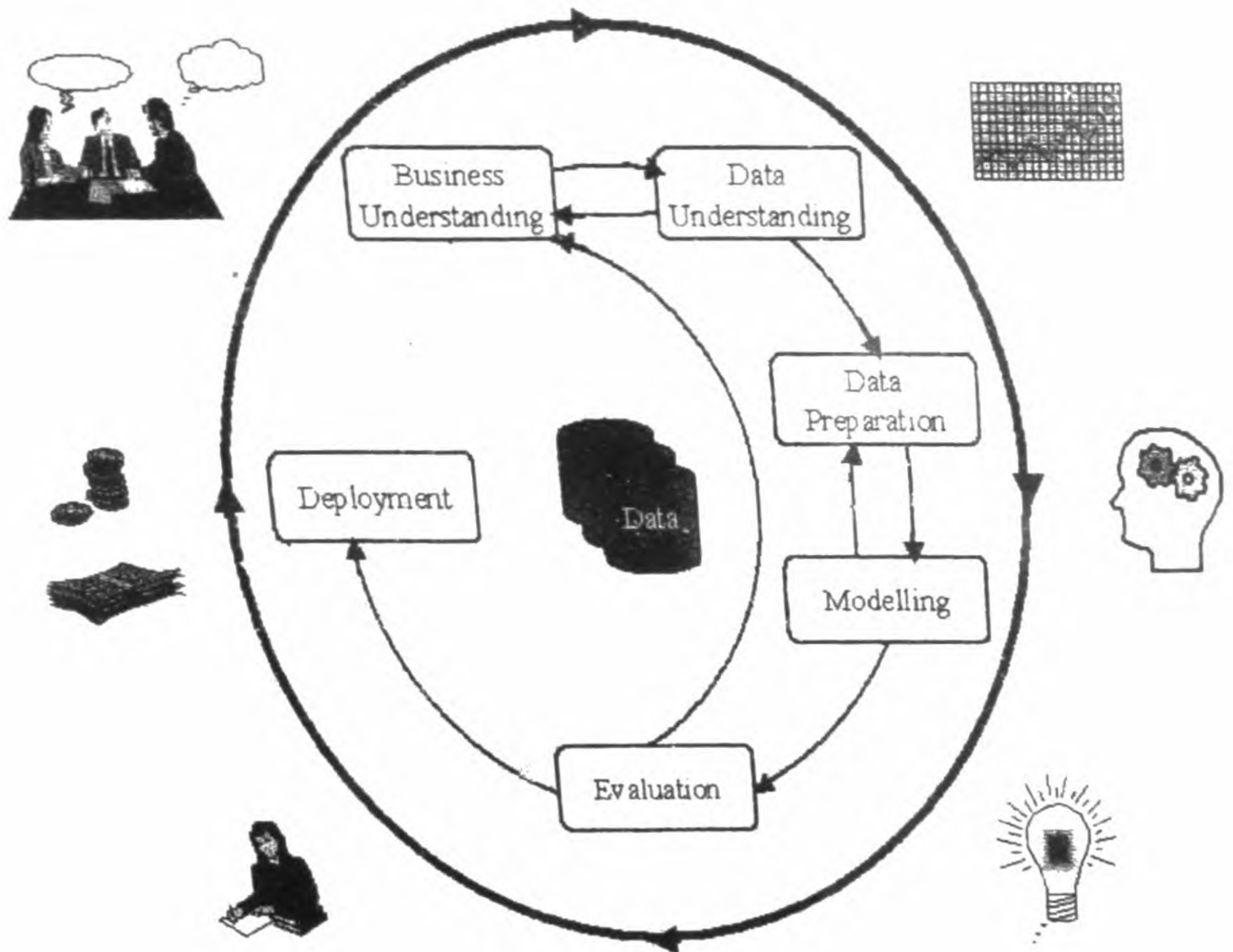


Figure 1: Larose, D. T. (2006) "Data Mining methods and models"

How these complementary approaches evolved and how Data Mining acts as the integrated unit within the BMIT is further explained in proceeding sections. First, Machine Learning, Statistical Learning and Operational Statistics are defined and then show how they complement in changing data into knowledge.

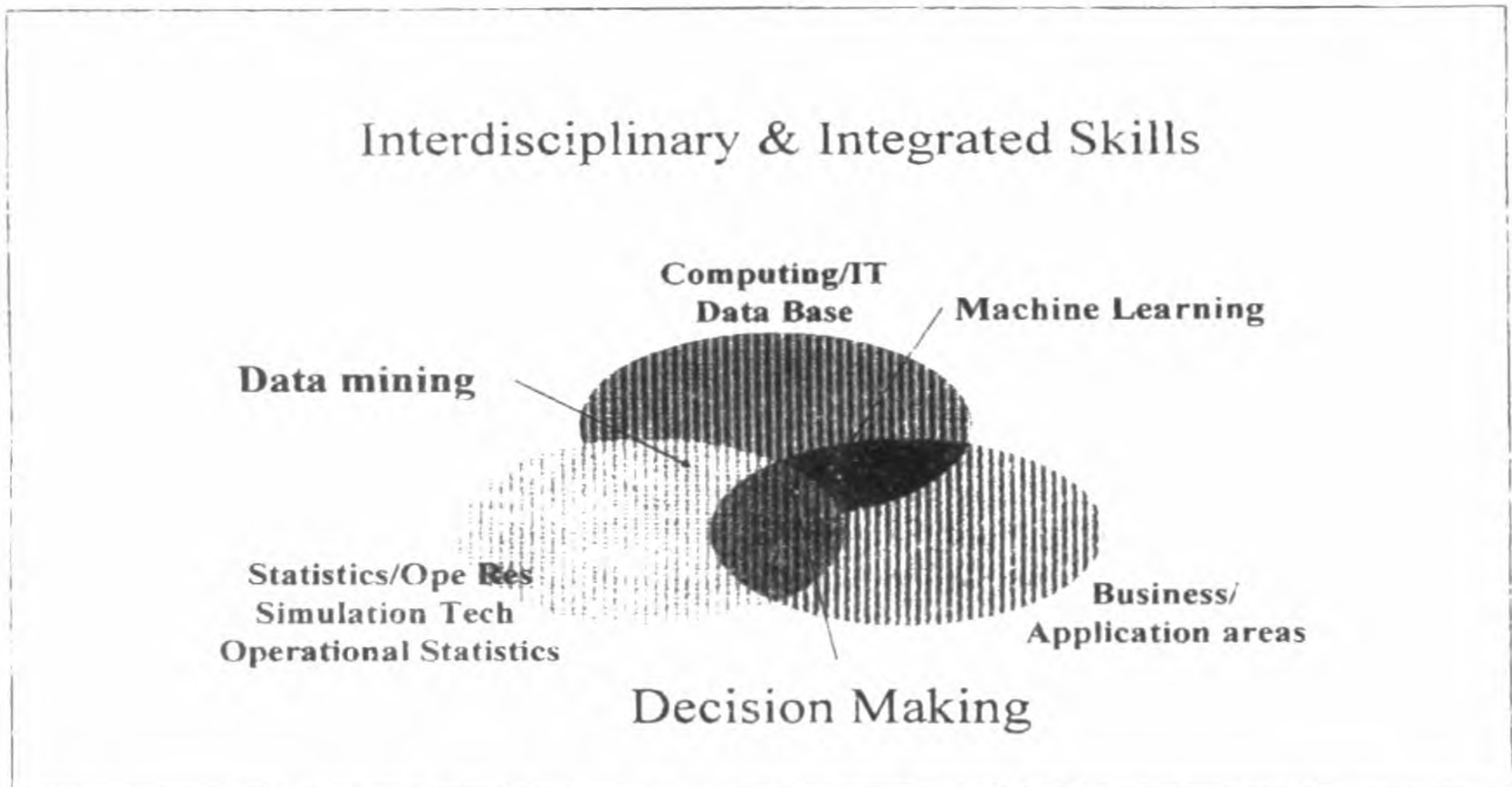


Figure 2: Key components of the BMIT conceptual framework

Machine Learning and Data Mining

Machine Learning or Data Mining was rooted from computer practice, computer science or artificial intelligence. It generates computer programs to automate the capability of learning from data by example rather than based on theoretical grounds (Berry and Linoff, 2000). This automation of data processing resulted from the technology development as mentioned above. Further new data base management systems (DBMS) software provided easier access to data. Increased memory and especially increased processor speed made possible the tackling of problems that were formerly too difficult to be feasible. Trends in networking include higher speed, greater connectivity, and greater bandwidth (Babinec, 2000).

This development of technology to support huge databases led to the rapid spread of computerization in all industries, government organizations and business and presented users with the problem of having to interpret vast amounts of data. This urgent and rapidly growing need to discover knowledge from these massive data sets

led to the need of systematic, fast and automatic ways in discovering knowledge from data (Kettening, 1994). Further rapid development in internet technology and hence the availability of vast amounts of data added to the demand for such Machine Learning techniques.

Statistical Learning

Traditionally Statistics is the science dealing with the collection, organizations, analysis and eventual interpretation of numerical data. Experiments and surveys are carried out in order to answer, or investigate, certain questions. Once the investigations have been carried out and the data have been collected conclusions, or assertions, were made, based on that data. What is common to data mining and statistics is that they both “Learn” from data or turn data into knowledge.

The main goal of Statistical Learning Theory is to “provide a framework for studying the problem of inference that is of gaining knowledge, making predictions, making decisions or constructing models from a set of data in a statistical framework” (Friedman, 1997; Hand, 1998; Hestie *et al.*, 2001). Leading from the statistics discipline Statistical Learning was also developed in response to this technology development, increase availability of large volumes of data and the need to discover knowledge from the data. This illustrates the complementary nature of Machine Learning and Statistical Learning and the strength of integrating the two approaches.

Operational statistics

Traditionally, in Operations Research, a stochastic model assumes a certain distributional form and use statistics in estimating the distributional parameters. In contrast to this, in Operational Statistics, first takes real data and uses statistics and operations research techniques to build the model based on this data. Then the same data will be used to demonstrate that it will do better under revised conditions. This extends this model to forecast future behaviour. Once again this is a result of technology development and the availability of massive sets of data. The Statistical part is mainly intended to predict what will happen in the future. Further to show that if the same statistical nature was to continue or persist, one could predict and illustrate how the future behaviour and the improvement we expect under revised conditions will perform. Building the input distribution based on data rather than making usual distributional assumptions makes Operational Statistics, Machine Learning and Statistical Learning share the common task of “Learn” from data.

This describes an Operational Statistics Model and Forecast which will give a snapshot of what would have happened under specified conditions. In this procedure the output distribution is estimated based on statistics rather than the input distribution as traditionally done. Finally due to the technology advancements this model can be used to develop an effective interactive Decision Support System as a deployment tool (Liyanage and Shanthikumar, 2005). This demonstrates the role of statistics; operations research and simulation techniques can be used in complement and how it leads from Statistical Learning. Thus the BMIT integrates Machine Learning, Statistical Learning, and Operational Statistics through integrated skills gained and further extends to addressing industry needs through applications.

Simulation Techniques

Simulation Techniques plays a significant role in developing prototypes, decision support systems, processes modeling where theoretical solutions are hard and when real situations are too complex to model. Further these skills will enhance theoretical based solutions to make them dynamic and give real time solutions.

Majors and sub-majors in BMIT

In addition to the growing need to turn data into information, availability of data from many different disciplines and application areas where these skills can be utilized became immense (Berry and Linoff (2000); Berry and Linoff, (1997); Grossman *et al.*, 1999; Hand, 1999; Ketting, 1994). They include, medicine, biology, business, finance, economics, marketing, environmental science and even social sciences and much more. This has opened the door for students in the BMIT to apply their combined computing and mathematics/statistics/operations research skills to any application area of their choice. Thus the flexibility within the degree having open elective units provided students the opportunity to tailor the course to suit their particular interests. Within the BMIT, students had the option of taking a major or a sub-major from a variety of disciplines.

Popular computing and mathematics majors and sub majors within the BMIT are Business Information Systems, E-Business, Software Development, Systems Analysis and Design, Data Communications, Web Development, Software Development, Statistics and Mathematics. Using elective units within the course, students were able to select majors/sub-majors available outside the school of computing and mathematics from application areas relevant to their own interests. An exhaustive list of majors available for the BMIT students consist of commercial law, finance law, property law, international economics, accountancy, asian studies,

banking, E-business management, E-business information systems, economics, economic policy, finance, finance law, financial management, health services management, horticultural food and wine, hospitality management, international economics, languages (Chinese, Japanese, etc.), marketing, urban and regional economics etc. New additional and proposed majors and sub-majors for the BMIT are, Knowledge Discovery and Data Mining and Intelligent Business and Information Systems. These two majors are strictly integrated majors and particularly designed for BMIT students.

Hence the BMIT not only combines two, logically related subject matters computing and quantitative techniques through data mining into a single offering it also enables students to apply these integrated skills into many application areas through majors or sub majors in many other disciplines and offer more exclusive and broad skills to employers. This attracted the students with higher UAI we were loosing to other disciplines while producing graduates valued by employers (See the Figure 3 given below for increasing demand). Further this fulfilled the earlier mentioned national need highlighted by The National Committee for Mathematics of The Australian Academy of Science in their review report “Mathematical Sciences Adding to Australia” (The National Committee for Mathematics, 1996).

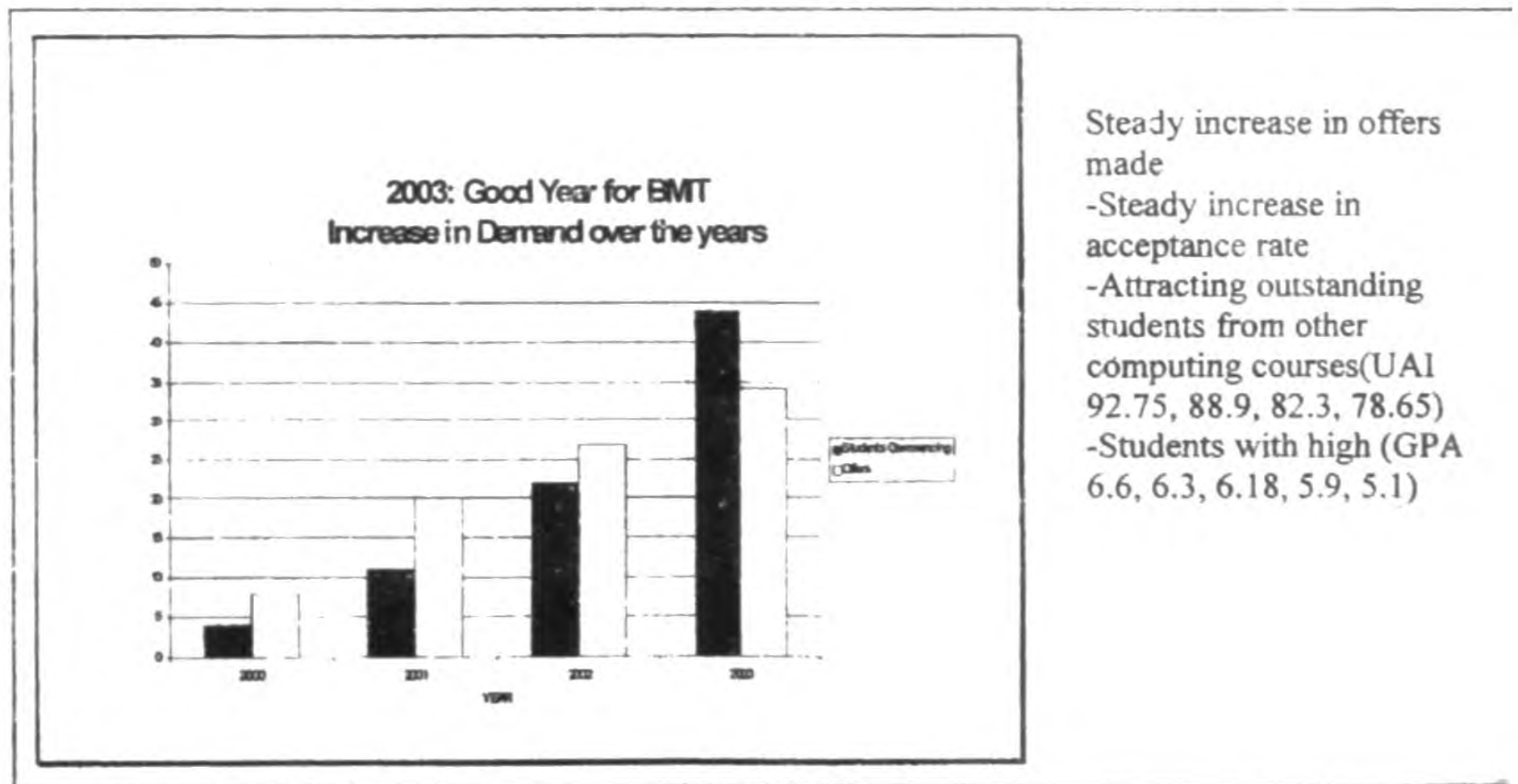


Figure 3: Mathematics and information technology demand curve

BMIT Success, Lessons Learned and Future Directions

The introduction of the BMIT with Data Mining as a package has been successful. The student satisfaction and the increasing demand in three years for the course (as shown in the graph above) confirmed this notion. It has been noted that the high achieving students have been well placed in the job market. Further it has developed new research strengths and industry links and broadens the integration with many other disciplines. New developments in research such as E-Research, E-Health, and E-Science are based on and build on this framework. Further most of the new – Analytics – domain covers what is derived from this framework. It is recommended that a course integrating Statistics, Operations Research and IT is suitable and necessary to produce Data Mining Professionals needed by industry to cater for current expectations and demand. Development of a suite of courses based on this concept, namely, three year single degree and a four year double degree (with B. Business and Commerce), bachelors honours degree, graduate certificate, post graduate masters are currently underway.

The lack of fully integrated majors within BMIT has been noted as a drawback and which needs to be introduced in the future. The proposed integrated majors such as Knowledge Discovery and Data Mining, Computational Statistics and Operations Research and Intelligent Business Information Systems have been recommended as future directions. Further Majors such as Health Informatics, Environment and Climate Change, Bio-Security, Psychological assessment in areas of Health Management, Scientific studies, Business Intelligence, Financial Engineering and Social Sciences are also under consideration.

Having “Mathematics” in the name of the course had its advantages and drawbacks. One of the major drawbacks is having a generic name like – Mathematics - in the title. It restricted giving its intended focus of data driven and data to knowledge focus. Having Information Technology in the title has also raised concerns as the name is shared by other computer science related courses. Hence the recommendation is to use a name not used by other groups already. Data Informatics, Operational Statistical Informatics, Decision Making Informatics, Statistical Informatics and Analytics are few names and titles suitable to describe this BMIT concept and have been put forward for consideration for the new such courses underway.

REFERENCES

- Agrawal, N. and S.A. Smith (1996). Estimating negative binomial demand for retail inventory management with unobservable lost sales. *Naval Research Logistics*, 43: 839–861.
- Azoury, K.S. (1985). Bayes solution to dynamic inventory models under unknown demand distribution. *Management Science*. 31: 1150–1160.
- Babinec, A. (2000). Data mining and statistics in the age of progress. Technical Report SPSS.
- Berry, M.J.A. and G. Linoff (1997). *Data Mining Techniques for Marketing, Sales, and Customer Support*. John Wiley and Sons, New York.
- Berry, M.J.A. and G. Linoff (2000). *Mastering Data Mining*, Wiley, Hoboken, New York.
- Bookbinder, J.H. and A.E. Lordahl (1989). Estimation of inventory reorder level using the bootstrap statistical procedure. *IIE Trans*. 21: 302–312.
- Conrad, S.A. (1976). Sales data and the estimation of demand. *Operations Research Quart.* 27: 123–127.
- Ding, X., M.L. Puterman and A. Bisi (2002), The censored newsvendor and the optimal acquisition of information. *Operations Research*, 50: 517–527.
- Friedman, J.H. (1997). Data mining and statistics: What's the connection?, 29th Symposium on the Interface.
- Gallego, G. and I. Moon (1993). The distribution free newsvendor problem: review and extensions. *Journal of Operations Research Soc.* 44: 825–834.
- Hand, D.J. (1998). Data mining statistics and more? *The American Statistician*, 52(2).
- Hand, D.J. (1999). Statistics and data mining: Intersecting disciplines. *SIGKDD Explorations*, 1(1): 16-19.
- Hand, D.J., H. Mannila and P. Smyth (2000). *Principles of Data Mining*. MIT Press.
- Hestie, T., R. Tibshirani, and J. Friedman (2001). *The elements of statistical learning: Data mining, inference, and prediction*. Springer-Verlag, New York.
- Karlin, S. (1960). Dynamic inventory policy with varying stochastic demands, *Management Science*, 6: 231–258.

- Kettinging, J.R. (1994). What industry needs? Proceedings of Symposium of Modern Inter-disciplinary University Statistics Education, Commission on Engineering and Technical Systems (CETS).
- Lariviere, M.A. and E.L. Porteus (1999). Stalking information: Bayesian inventory management with unobserved lost sales. *Management Science*, 45: 346–363.
- Lim, A.E.B., J.G. Shanthikumar and Z.J.M. Shen (2006). Model uncertainty, robust optimization, and learning, tutorials (In) *Operations Research: Models, Methods, and Applications for Innovative Decision Making*, M.P. Johnson, B. Norman and N. Secomandi (eds.), 66-94.
- Liyanage, L.H. and J.G. Shanthikumar (2005). A practical inventory control using operational statistics. *Operations Research Letters*, 33: 341-348.
- Nahmias, S. (1994). Demand estimation in lost sales inventory systems, *Naval Research Logistics*, 41: 739–757.
- Scarf, H. (1958). *Amin–max Solution of an Inventory Problem*, *Studies in the Mathematical Theory of Inventory and Production*, Stanford University Press, Stanford, CA, (Chapter 12).
- Scarf, H. (1959). Bayes solutions of statistical inventory problem. *Ann. Math. Statist.* 30: 490–508.
- The National Committee for Mathematics (1996). The National Committee for Mathematics of The Australian Academy of Science, Mathematical Sciences Adding to Australia.