The reliability of the Shotgun Stochastic Parameter Estimation Algorithm for analysing rank ordered data

Justin Thomas Grant

A research report submitted to the Faculty of Commerce, Law, and Management, University of the Witwatersrand, in partial fulfilment of the requirements for the degree of Master of Business Administration

Johannesburg, 2010

(VERSION AUGUST 2010)

ABSTRACT

The introduction of the Shotgun Stochastic Parameter Estimation Algorithm (the Algorithm) in 2006 by A. Stacey offered a simple method of introducing interval level data into rank ordered observations thus allowing the application of more powerful statistical analysis on ordinal level data. As this is a relatively new development in the debate regarding treating ordinal level data as if it were interval the Algorithm's mechanics (i.e. its internal relationships) remain relatively unknown.

This study, by means of a series of simulations, looks at the standard errors of the sample means simulated by the algorithm after two variables, intrinsic to the Algorithm, are systematically changed. The variables changed are the number of items ranked and the size of the internally generated interval level based sample.

What was found was that as the number of items ranked increases so too does the inconsistency of the algorithm, conversely as the size of the simulated sample increases the consistency improves, initially at an accelerated rate but then slowing rapidly into very small changes for large increases in the sample size.

The results were then compared with the sample error that results from using observed sample sizes of the same magnitude as those of the simulated samples in the study and it was discovered that the Algorithm has comparatively much less error in it.

The study concludes that the Algorithm is a reliable method to use and offers guidelines for the safe use of the algorithm that hope to ensure consistent and reliable generation of outputs.

DECLARATION

I, Justin Thomas Grant, declare that this research report is my own work except as indicated in the references and acknowledgements. It is submitted in partial fulfilment of the requirements for the degree of Master of Business Administration in the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination in this or any other university.

Justin Thomas Grant

Signed at ……

On the …….. day of …… 2008

DEDICATION

This piece of work would not be possible without the love and devotion of my beautiful wife Elmarie, I hope that this and what comes after makes up for all the nights you stayed awake with me and all the extra work needed to keep our home going while I was working on it. Thank you for your encouragement and faith in me.

This is also for my new baby girl, Kayla you are only seven months old now so you will not be able to read this for many years to come but hopefully the effect of it will be felt in your life until then.

FOREVER

ACKNOWLEDGEMENTS

Corneile Britz, thank you for the personal time you sacrificed programming the Algorithm into an application that effectively reduced the time it would take to do this study by at least a factor of thirty.

Dr. Anthony Stacey, I appreciate the patience and guidance you have given me in spite of experiencing a significant (no hypothesis test required) amount of personal turmoil in your own life.

Lyn Elliot and Ryan Walker, I am very grateful for all the times you have willingly and unwillingly assisted in reading my work and for patiently explaining to me why my English is quite terrible.

Vulcan Catering Equipment and specifically my Manager Rui Barros, thank you for allowing me the odd moment here and there to put down thoughts that absolutely could not wait until I got home.

Margaret Crawford my Managing Director, this study but would never have taken place had it not been for the amazing opportunity you have given me to grow my own horizon, the world will never look the same again and for that I am eternally thankful.

Alan Walker, a man I consider my mentor and friend, your influence in my life and in my studies has always been a positive one that has evolved into a moral cornerstone that will forever assist me in the decisions I have yet to make, this gift is priceless.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

1.1 Purpose of the study

The purpose of this research is to analyse the influence that the number of variables and the simulated sample size have when using the Shotgun Stochastic Parameter Estimation Algorithm (the Algorithm) developed by A.G. Stacey (2006).

By defining this influence it is intended to start the process of designing a set of best practices to enable the effective and appropriate use of the algorithm (Stacey 2006).

1.2 Context of the study

The debate on performing statistical analysis beyond those prescribed by Spearman (1904) with relation to rank ordering, has been continuing since the seminal work by Thurstone (1927).

The main contention has always been the appropriate distance to apply between the subjects being ranked, i.e. “A may differ from B five times as much as B differs from C” (Spearman 1904: 80).

The use of rank ordering as a research tool appears in fields as varied as psychology, sociology, marketing, sports and econometrics (Yu 2000), in marketing it is mainly used for attribute mapping, product positioning, and finding ideal brand points (Greatorex 2007). In marketing, brands ranked according to attributes they might have in varying degrees, will assist in gaining

deeper knowledge about the positioning of a brand relative to customers preferences and to other brands (Unknown 2009)

Rank ordering also has a use in group decision making by increasing the consideration of alternatives thereby improving the decisions made by that group (Hollingshead 1996).

The problem with rank ordering is attempting to make sense of the observations through the correct application of statistical methods. In many instances, the incorrect method, despite its inappropriateness is utilised in an attempt to gain further insight (Stevens 1946; Abelson and Tukey 1963; Mayer 1971; Hensler and Stipak 1979; Knapp 1990; Velleman and Leland 1993; Hollingshead 1996; Yao and Bockenholt 1999; Yu 2000).

It is into this turmoil that the Algorithm is introduced in an effort to estimate the unknown parameters of the variables that constitute a rank ordering. By applying the computational power currently available the Algorithm estimates the population means and standard deviations of the underlying scale for a quantity of survey items from which the observed data sample was drawn (Stacey 2006).

The Algorithm offers a simple method that could potentially yield powerful results, the gap that exists is the very knowledge as to how powerful a tool it really is and the purpose of this research is to start the verification of the robustness of the method.

1.3 Problem statement

1.3.1 *Main problem*

What would the result be of changing the quantities of variables and simulated sample size within the Algorithm, would it in any way affect the accuracy of it over a set of iterations.

1.3.2 *Sub-problems*

The first sub-problem is to analyse the effect that changing the quantity of ranked variables has on the accuracy of the Algorithm.

The second sub-problem is to analyse the effect that changing the simulated sample size has on the accuracy of the Algorithm.

1.4 Significance of the study

The study fills a gap in that currently there is no literature in support or denial of the Algorithm neither is there any literature beyond the initial publishing of the algorithm to substantiate its possible widespread use (Stacey 2006).

The study hopes to provide guidance in enabling researchers to make a decision as to the applicability of the Algorithm tool relative to their research by dealing with two important variables that constitute part of the Algorithm. The aim is to either give confidence in its use or rule it out as a possible method going forward.

Upon completion, this study will be of use to marketing researchers wishing to clarify brand positioning (Greatorex 2007) as well as managers requiring assistance in group decision making as in the election of the American Psychologist Association's president (Diaconis 1989). For Marketers it means knowing not only that their brand is better than the opposition's but also by how much, this could influence pricing strategies in a profound way. In decision-making, it means knowing that the first, second and third ranked choices were much closer than was originally apparent, and therefore the decision is not as simple as it appeared to be.

1.5 Delimitations of the study

This study deals with simulated data in an attempt to isolate methodological error and therefore excludes the effects that empirical data would also have on the results.

The Algorithm is stated to be able to usefully transform partially ranked variables (Stacey 2006) the use of the Algorithm in this capacity is excluded from this research.

The study looks only at the internal relationships that exist within the Algorithm and as a result external applicability to the various disciplines mentioned in sub heading **1.4** are not included in the analysis. The resulting recommendations are therefore for general use of the Algorithm.

1.6 Definition of terms

The Algorithm will always refer to Shotgun Stochastic Parameter Estimation Algorithm as developed by Stacey (2006).

A distinction exists between the population sample and the simulated population sample. Population sample refers to the actual observed data as gathered by responses to a request to rank order a set of variables. This will involve actual human beings in the decision process of assigning those ranks, whereas simulated population sample refers to the computer generation of a population calculated from a set of normally distributed standardised numbers of overall variance equalling unity randomly sorted.

1.7 Assumptions

It is assumed that the hypothetical observed population samples used in the research are free from sample error. This assumption is reasonable considering the infinite amount of possible variations that the observed population sample could take it is safe to assume that at some point the samples used in this study could be a true representation of an existing population somewhere in the infinite universe. Given that it is only an investigation of the Algorithm that is at stake in this study the effect of any sample error that might occur is of no matter as this study looks internally at the algorithm's mechanics and not its external sources of information.

It is assumed that in spite of the small sample size of data collected that there is enough data to point the next researcher in the right direction. In this study a

small sample of four variations of the simulated sample is used instead of the widely recognised minimum of 30 (Albright, Winston and Zappe 2008) however, the Algorithm is applied to each variation a minimum of 30 times. The sensitivity of the research outcome to this limitation could be great however; this study provides a good starting point for further research.

CHAPTER 2: LITERATURE REVIEW

2.1 Introduction

The purpose of this literature review is to provide contextual background to the analysis of rank ordered data, outlining the history behind the developments in this field as well as expose the causes of disagreement that still exist.

2.2 Quantitative Estimates for Sensory Events

Spearman (1904) in his seminal work “The Proof and Measurement of Association between Two Things” made the statement that on the basis of rank ordering alone the distance between two variables is indeterminate and went on further to say that the attractiveness of using rank order data is that any disparity between the two variables being compared is eliminated.

Thurstone (1927), realised the power such a comparison method could have. Thurstone took it upon himself to define the distance between two sources of physical stimulus on a “psychological continuum” (Thurstone 1927: 273) in his work entitled “A Law of Comparative Judgement” which has become the cornerstone of all attempts at defining the underlying scale of judgement existing in rank ordered data. It is based on the assumption that the variables follow a multivariate normal distribution function (Yao and Bockenholt 1999). The problem with Thurstonian ranking models is the “high-dimensional integration” (Yao and Bockenholt 1999, p 79) that is required when attempting to estimate the model parameters and as such is limited to ranking probabilities with up to four options (Yao and Bockenholt 1999). Bell (2005) also states that

caution should be used in particular when dealing with Thurstone scales as they require careful consideration.

Stevens (1946), in response to a call to answer the question of what exactly is meant by measurement and the variety of forms that it may exist in, came up with the scale names as well as the various attributes of data as described in his table below.

Scale	Basic Empirical Operations	Mathematical Group Structure	Permissible Statistics (invariantive)
NOMINAL	Determination of equality	*Permutation group* $x' = f(x)$ $f(x)$ means any one-to-one substitution	Number of cases Mode Contingency correlation
ORDINAL	Determination of greater or less	*Isotonic group* $x' = f(x)$ $f(x)$ means any monotonic increasing function	Median Percentiles
INTERVAL	Determination of equality of intervals or differences	*General linear group* $x' = ax + b$	Mean Standard deviation Rank-order correlation Product-moment correlation
RATIO	Determination of equality of ratios	*Similarity group* $x' = ax$	Coefficient of variation

Table - 1 A Classification of Scales of Measurement (Stevens, 1946)

Stevens (1946) went on to say that rank ordering gives rise to the ordinal scale of data and in his mind condemned it to the realm of non-parametric statistics but he also stated that utilising “illegal” statistics could provide useful results. It is the last statement that has since created all the controversy surrounding analysis of rank ordered data, as Knapp(1990: 121) succinctly put it “ both sides agree that a certain variable is ordinal. But they part company when analyzing the data generated by that variable.”

There are those that conservatively agree with Stevens that only non-parametric methods of analysis are truly applicable and defendable as a

method to interpret ordered data. Siegel (1957), also acknowledges that a continuum may exist underlying the observed data but maintains the use of non parametric techniques are still the most reliable for analysis on such data as seen by his recommendations in the table below.

LEVEL OF MEASUREMENT	NONPARAMETRIC STATISTICAL TEST					NONPARAMETRIC MEASURE OF CORRELATION
	One-Sample Case	Two-Sample Case		k-Sample Case		
		Related Samples	Independent Samples	Related Samples or Randomized Blocks	Independent Samples	
Nominal	Binomial test χ^2 test	McNemar test	Fisher exact probability test χ^2 test	Bowker test Cochran Q test	χ^2 test	Contingency coefficient
Ordinal	Kolmogorov-Smirnov test Runs test	Sign test Wilcoxon test	Festinger test Kolmogorov-Smirnov test Mann-Whitney U test Median test Moses test of extreme reactions Wald-Wolfowitz runs test White test Wilcoxon test	Friedman test Wilson test (15)	Extension of the median test Jonckheere test Kruskal-Wallis test Mood runs test Mosteller slippage test Whitney extension of the U test	Cureton biserial rank correlation (1) Kendall rank correlation coefficient Kendall partial rank correlation coefficient Kendall coefficient of concordance Moran multiple rank correlation Spearman rank correlation coefficient
Interval	Walsh test Randomization test	Walsh test Randomization test	Randomization test	Randomization test	Randomization test	Olmstead-Tukey corner test Randomization test

Table 2 – Nonparametric Statistical Tests and Measures of Correlation for Various Designs and Various Levels of Measurement (Siegel, 1957)

Mayer (1971) also agreed that ordinal data should be left in its observed state and in fact treating it as though it were interval would possibly result in a radically false analysis and made recommendations along the lines of greater research being applied to ordinal data without the assumptions of interval spacing.

Despite Mayer's (1971) warning there were still statisticians that believed in the additional comment that Stevens (1946) made when he said that useful results could still be arrived at when applying "illegal" statistics to ordinal data. A year before Mayer's warning, Labovitz (1970) argued for the treatment of ordinal data as if it were interval stating that there were three benefits to doing so, more powerful statistical tools could be used, greater levels of data retention could be achieved and access to greater statistical manipulation and representation of data. However, Labovitz (1970) as well as Hensler and Stipak (1979) add the need for caution when using interval level assumptions on ordinal data especially with relation to the pre-existing category values chosen to represent them.

The dawning of a new age arrived where considerable computational power could be applied to problems with minimal effort (Stacey 2006), amongst those to develop methods utilising this benefit were Young (1981), Yu (2000), Yao and Bockenholt (1999), Efron and Tibshirani (1986). These academics focussed on fully ranked data and the transformation of ordinal level data for enhanced analysis. In 2006, Stacey decided to develop an algorithm that could also incorporate the analysis of partially ranked data and came up with the Algorithm as detailed below.

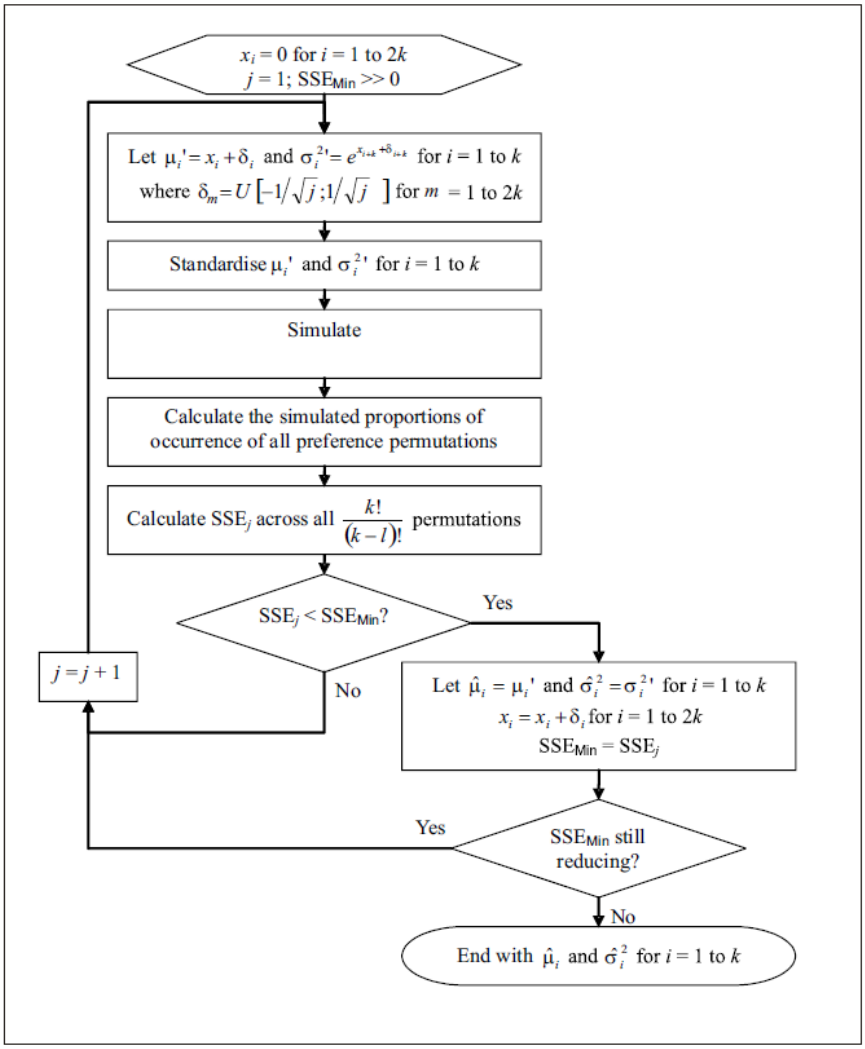


Figure 1 - The Shotgun Stochastic Parameter Estimation Algorithmic Flow

Stacey (2006: 29) explains the process (which has been interpreted into a Microsoft® Excel® example for 6 ranks in Appendix A for ease of understanding) as follows:

1. Start with initial guesses for the item means, $\mu_i = 0$, and item variances, $\sigma_i^2 = 1$ for $i = 1\ to\ k$.
2. Standardise the means and variances such that the total variance is equal to 1; i.e. $\sigma_{Total}^2 = \sigma_\mu^2 + \mu_{\sigma^2} = 1$

3. Numerically simulate the population rank ordering responses by taking a very large simulated sample.
4. Calculate the sum of squared errors, (SSE) between the observed proportions and the simulated population proportions across all preference permutations.
5. Compare the SSE to that obtained using previous estimates of item means and variances. If the SSE is less than the previously recorded minimum SSE, then the current guesses become the best estimates of the items' means and variances.
6. Introducing uniformly distributed random perturbations to the guesses of the item means and variances, which decrease in magnitude with the square root of the number of iterations, carried out, repeat steps 2 to 6 until the SSE is no longer reducing with multiple successive iterations.

In simplified terms, this means that the algorithm estimates the means of the observed population set and by using a Latin Hypercube, it calculates a ranked simulated population set.

At this juncture the concept of Latin Hypercube Sampling (LHS) (McKay, Beckman and Conover 1979) needs to be introduced. LHS ensures that all variations of a variable are represented regardless of its level of possible importance (McKay et al. 1979; Cheng and Druzdzel 2000). Cheng and Druzdzel (2000) concluded that using LHS would lead to significant improvements in performance in sampling algorithms with regard to convergence rates. The quality of improving convergence rate makes LHS

extremely useful for the Algorithm. An example of how to generate a LHS for the purposes of the algorithm appears in Appendix A.

The Algorithm then compares these ranks to the actual observed set of ranks gathered through field research by looking at the Sum of the Squared Errors (SSE) between observed and simulated sample sets. The Algorithm then repeats the whole process but only accepts new guesses for the means if the SSE reduces. Stacey (2006) found that 1000 iterations would suffice to conclude an outcome. When the Algorithm concludes, what the researcher has is the estimated standardised means of the observed set of ranked items with respect to one another based on normally distributed interval level data.

It is the Algorithm that this research will attempt to shed light on, as its creator has pointed out there are many properties that are unexplored such as the dual sample error that could occur as well as other relationships that remain undefined. It is this uncertainty that leads this research into attempting to quantify some of the relationships that exist (Stacey 2006).

As mentioned in Chapter One, this study will deal with two important variables within the Algorithm and therefore it is necessary at this point to list the variables that affect the performance of the Algorithm.

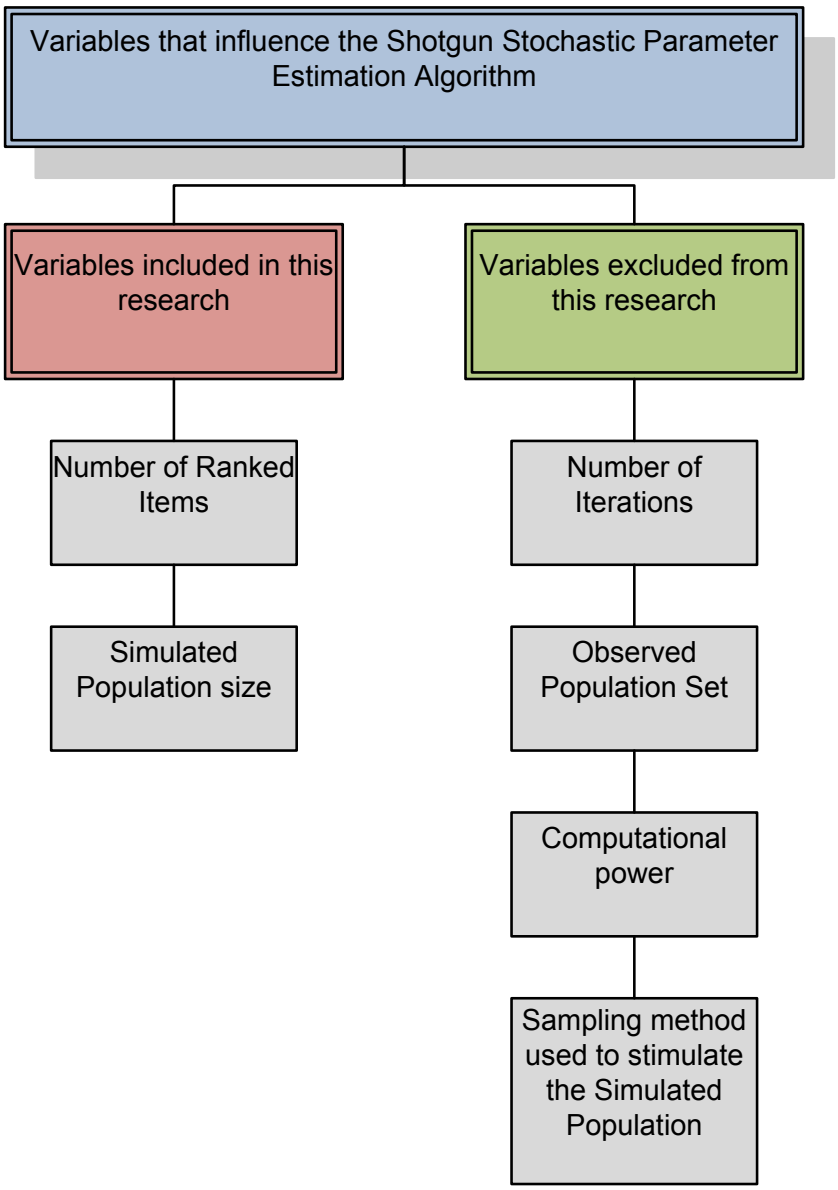


Figure 2 – Variables that Influence the Shotgun Stochastic Parameter Estimation Algorithm

2.3 Defining the Variable Variance

The initial problem with Thurstonian ranking involves the mathematical limitation on the number of variables to be ranked (Yao and Bockenholt 1999; Stacey 2006), so it is a logical choice to test the consistency of the Algorithm across a range of variables. Stacey (2006) makes mention that ranking a large number of

items tends to be quite difficult and therefore a logical limitation of from two to eight items will be used in this analysis.

Cheng and Druzdzel (2000) state that when precision is not important stochastic simulation methods are preferred to exact algorithms mainly because of execution time, which in the case of stochastic simulations relies on the number of samples.

2.3.1 Research Proposition One: varying the number of ranked items

It is proposed that the variance of estimates given for the means and standard deviations for a set of observations when the quantities of ranked items are increased will themselves also increases.

2.4 Defining the Sample Error in the Simulated Sample

The second sub-problem is to quantify the affect that changing the quantity of the simulated sample size has on the accuracy of the Algorithm, as Stacey (2006) stated in his article the effect that this type of sampling error has on the data is unknown and therefore requires research. Cheng and Druzdzel (2000) state that an increase in simulated samples will lead to greater precision.

2.4.1 Research Proposition Two: varying the simulated sample size

It is proposed that the standard error of the estimates given for the means for a set of observations will decrease when the number of simulated observations is increased.

2.5 Conclusion of Literature Review

The entire literature review shows that this area of study has many varying opinions and views on what constitutes the correct method for analysing ordinal level data in an attempt to gain meaningful insights to the observations made by the population sample. It does not provide insight into the specific method under investigation and the researcher is limited to using logic and the recommendations of the Algorithm's creator, Stacey (2006).

This study is pioneering in that it is the first attempt at defining the Algorithm further and therefore it is as a valuable contribution towards this area of research.

2.5.1 *Research Proposition One*: varying the number of ranked items

It is proposed that the variance of estimates given for the means and standard deviations for a set of observations when the quantities of ranked items are increased will themselves also increases.

(Thurstone 1927; Yao and Bockenholt 1999; Cheng and Druzdzel 2000; Stacey 2006)

2.5.2 *Research Proposition Two*: varying the simulated sample size

It is proposed that the standard error of the estimates given for the means for a set of observations will decrease when the number of simulated observations is increased.

(Cheng and Druzdzel 2000; Stacey 2006)

CHAPTER 3: RESEARCH METHODOLOGY

This section will clarify the design and limitations of the methodology employed. It will look at trying to minimise any restrictions towards the reliability and validity of the research and address any possible issues that may arise or at least highlight areas of concern.

3.1 Research methodology / paradigm

This research, despite its qualitative nature in that it deals with judgements that people make concerning items under comparison, utilises quantitative descriptions such as standard error as well as means and therefore the methodology utilised is of the same quantitative nature. Confirming its quantitative nature is the examination of cause and effect relationships between numeric values by employing an experimental approach. (Cresswell 2003)

The obvious assumption this methodology makes is that the introduced cause is the reason behind the effect and so measures must put in place to ensure that this is true by examining all the additional contributing factors that could possibly have an effect on the outcome. This assumption is relevant to this research because it operates on an iterative experimental design (Steinberg and Hunter 1984) whereby holding all other variables constant one is changed to gauge the impact it makes. The research therefore needs to consider all other influences that led to the result, possibly changing one of the other variables and then holding it constant once more could result in an entirely different set of outcomes after repeating the process.

3.2 Experimental Research Design

This research makes use of the Algorithm in isolation and therefore the methodology utilised will be an experimental design using a series of computer simulations based on contrived data. Given that the research is looking at the internal mechanics of the Algorithm, no external sources of information are required.

A disadvantage associated with this approach will be that the computational time required to run all the simulations is prohibitive, conversely not needing to gather external sources of information prior to running the simulations saves time. The use of the computer environment makes the experimental approach much more controlled and all factors influencing the outcome more visible.

Steinberg and Hunter (1984) state that according to “Response Surface Designs” data which becomes rapidly available allows researchers to be adaptive in their search for answers and accordingly they can make the experiment proceed in stages with each preceeding stage providing insight to shape the following stage. The process followed in this research will be to put forward an initial model that holds one of the variables constant and changes the other in stepped increases, which could be incorrect, and then make adjustments to the final experiment model encompassing all the concerned variables based on the outcomes of the first initial model (Steinberg and Hunter 1984).

3.3 Population and Sample

3.3.1 *Population*

This research analyses the Algorithm using simulations, the Algorithm makes use of a random number generator in the creation of the LHS and this together with the limitless combinations of the different variables makes the population sets that could exit the Algorithm the result of an infinite range of all possibilities and thus themselves also infinite.

3.3.2 *Sample and sampling method*

Given the experimental approach (Steinberg and Hunter 1984) this research has adopted, an initial test run of a series of simulations will attempt to discover a stable version of the Algorithm for four ranked items this will determine the final sample sets used. The initial model will use sets of 30 runs of the Algorithm for each change, therefore depending on the number of changes done the final model's sample size will equal the number of variations multiplied by 30. From this initial model, a new adopted range of experimental parameters will be set and applied to 4, 6 and 8 ranked items; this will give rise to the final sample set for analysis.

In addition, there are two internal samples utilised within the Algorithm, the LHS/simulated sample is one and the set of observations for comparison is the other. The LHS's used will be guided by the outcome of the initial model; the observation set for each of the quantities of ranked items will be a contrived set of numbers. The proportions used here were manually inserted one by one

whilst paying attention to making them as fairly even a spread of percentages across each of the ranked Items and rank values as seen below.

		Item 1	Item 2	Item 3	Item 4
Ranks	1	0.257	0.1877	0.4139	0.1414
	2	0.2168	0.2266	0.3866	0.1701
	3	0.1658	0.3452	0.0658	0.4232
	4	0.3604	0.2405	0.1338	0.2653

Table - 3 Contrived Observed Data Set Four Ranks

		Item 1	Item 2	Item 3	Item 4	Item 5	Item 6
Ranks	1	0.2144	0.1877	0.2025	0.1414	0.1877	0.0663
	2	0.2168	0.2266	0.1116	0.1701	0.1433	0.1318
	3	0.1658	0.2659	0.1544	0.2685	0.0647	0.0808
	4	0.1423	0.1588	0.1338	0.2045	0.1844	0.1762
	5	0.1558	0.1155	0.2655	0.1653	0.1253	0.1727
	6	0.1050	0.0457	0.1323	0.0501	0.2947	0.3723

Table 4 - Contrived Observed Data Set Six Ranks

		Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Item 7	Item 8
Ranks	1	0.1986	0.1877	0.1817	0.1414	0.0479	0.0663	0.0717	0.1048
	2	0.1722	0.1986	0.1116	0.1598	0.1312	0.1318	0.0022	0.0926
	3	0.1410	0.0478	0.1246	0.2165	0.0647	0.0808	0.2590	0.0657
	4	0.1166	0.1588	0.1165	0.1812	0.1065	0.1762	0.0547	0.0895
	5	0.1369	0.1155	0.2099	0.1365	0.1253	0.1727	0.0212	0.0820
	6	0.1035	0.0457	0.1477	0.0501	0.2215	0.2565	0.0785	0.0966
	7	0.0458	0.1581	0.0012	0.0931	0.2336	0.0579	0.2855	0.1247
	8	0.0856	0.0879	0.1069	0.0211	0.0692	0.0579	0.2274	0.3440

Table 5 - Contrived Observed Data Set Eight Ranks

3.4 The research instrument

The Algorithm is the instrument that will generate the data. For the purposes of this research, the C# language converted the Algorithm into an executable

programme that allows it to run at a higher rate than the Appendix A Microsoft® Excel® example by utilising the multithreading capacity of a computer's processor. Multithreading allows for simultaneous calculations to run whereas Microsoft® Excel® only allows single threading which means one calculation before another and that results in much longer run times.

3.5 Procedure for data collection

The introduction of various LHS sizes based on the initial model into the Algorithm across the range of ranked items will give rise to the sample sets used for analysis. The C# executable programme releases the data in a comma separated value (csv.) file which is then stored for analysis.

3.6 Data analysis and interpretation

Willig (2008) states that a researcher needs to know how the data they want to collect will be analysed before they start to collect it. Bell (2005) recommends that when dealing with means, as are the outputs of the Algorithm which delivers sample means, standard deviation must be used to summarise dispersion. Albright (2008) clarifies further that the standard deviation of sample mean is in fact referred to as the Standard Error of Sample Mean (SE($\overline{\mathbf{X}}$)) which is calculated as $SE(\bar{X}) = s/\sqrt{n}$. Measuring the SE($\overline{\mathbf{X}}$) allows for the analysis of how much point estimates from different samples vary from one another and it will allow the research to determine how consistent the Algorithm is in reproducing results for a given set of inputs.

Bar charts, scatter plots as well as box plots will graphically present the results of the collective SE($\overline{\mathbf{X}}$) generated from the data. These will be analysed through observation as well as mathematical comparison.

3.7 Limitations of the study

Due to time constraints, a single focussed initial model will decide the final experiment's model, ideally several rounds of complete revised models would be preferred. This limitation could result in some unfocussed data, with assumptions made and possible incorrect conclusions drawn from it.

Time constraints might also limit the total number of variations allowed which could result in inferences made from very small sample sets that may or may not be true reflections of the population.

Using a single observed set of data for each of the ranked items could have a large effect on the outcome of the study however further study would need to confirm this.

3.8 Validity and reliability

"Reliability is concerned with consistency" (Kalof, Dan and Dietz 2008: 156) and interestingly enough that is the very purpose of this study, to test the consistency of the Algorithm and testing what could be the most stable version of it to deliver consistent results.

"Validity is concerned with congruence, or 'goodness of fit' between the details of the research, the evidence, and the conclusions drawn by the researchers"

(Kalof et al. 2008: 156). Validity therefore says that if the method used is correct then the results gained from it are correct and correct analysis can take place, in reverse if the analysis is to be correct it has to have had the correct method applied in the first place.

To achieve both validity and reliability is the reason for the particular experimental design adopted, as it will focus the research and give a definitive outcome by gathering data from as broad a spectrum as possible in the right direction.

3.8.1 *External validity*

External validity refers to the extent to which the conclusions drawn from the studies sample can have the same validity with respect to the population from which the sample was drawn (Kalof et al. 2008).

Due to the computational time that the Algorithm requires in order to run, the sample for each variation has been limited to n=30. A larger quantity would be more favourable for the study as well as add to the accuracy but in the absence of available resources, there exists a trade off with sacrificing greater external validity for available computational time. Also limited is the single observed set used for the study however, the impact of this, as stated before, may have to be decided by further study.

3.8.2 *Internal validity*

Internal validity refers to the studies ability to draw the correct conclusions from the analysis of the data that was gathered (Kalof et al. 2008).

The SE($\overline{\mathbf{X}}$) is an appropriate measure to determine consistency the only question remaining unanswered is whether the final experimental design eliminates the possibility of internal invalidity by gathering the correct size of samples from the simulations it recommends. Consideration of internal validity will guide the design of the final model and again the trade off between the ideal model and available computational time may take place.

3.8.3 *Reliability*

As stated above "reliability is concerned with consistency" (Kalof et al. 2008: 156) and in order to attain this within the experiments design what is done to the one variable will be replicated with the others so as to provide points of commonality and therefore enable comparison. The use of the same programme as well as computer for the experiment ensures a level playing field and observed data sets are roughly of the same nature ensuring further consistency.

CHAPTER 4: PRESENTATION OF RESULTS

4.1 Introduction

Results in this chapter will follow a pattern that allows for the sequential analysis of the data as it pertains to each proposition. Firstly, the discussion of probing model takes place along with its conclusion; this is what will drive the research to deliver the data in the remainder of this chapter.

4.2 Initial Model

The initial model is an attempt to see at which sizes of the simulated sample set (generated from and directly related to the size of the LHS) the Algorithm will deliver informative data for 4 ranked items. A LHS size of 2000 rows for each item generated the first set of data as seen below.

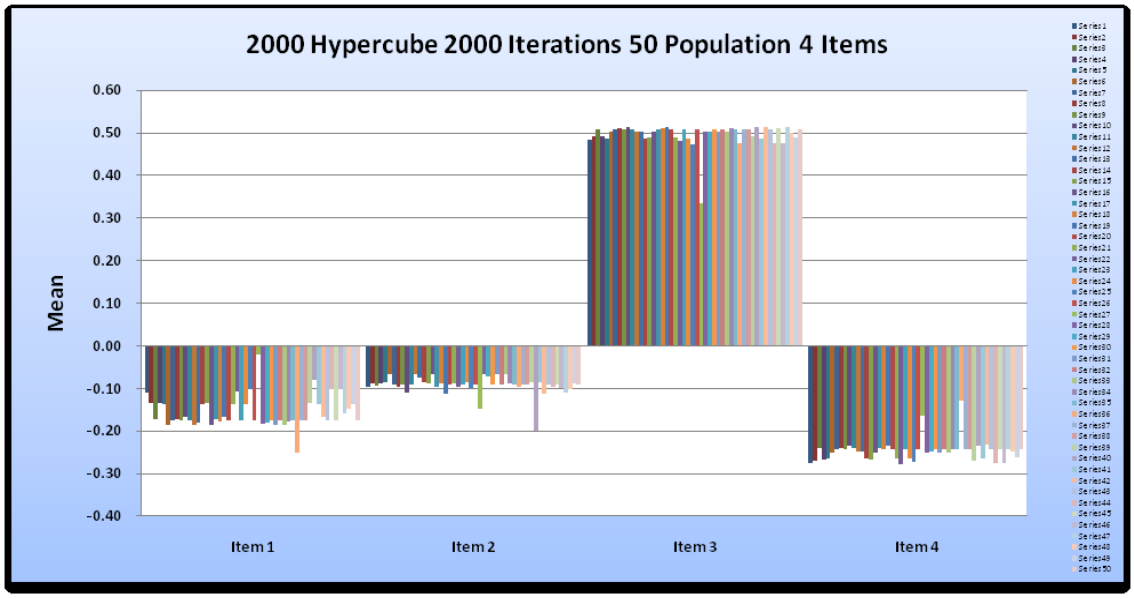


Figure 3 - 2000 Hypercube 2000 Iterations 50 Population 4 Items

The above graph visually shows quite a large degree of variation in the item means between each of the 50 sets of means and so a larger LHS of 10000 rows per item and an increased iteration count of 10000 generated the next set of data below.

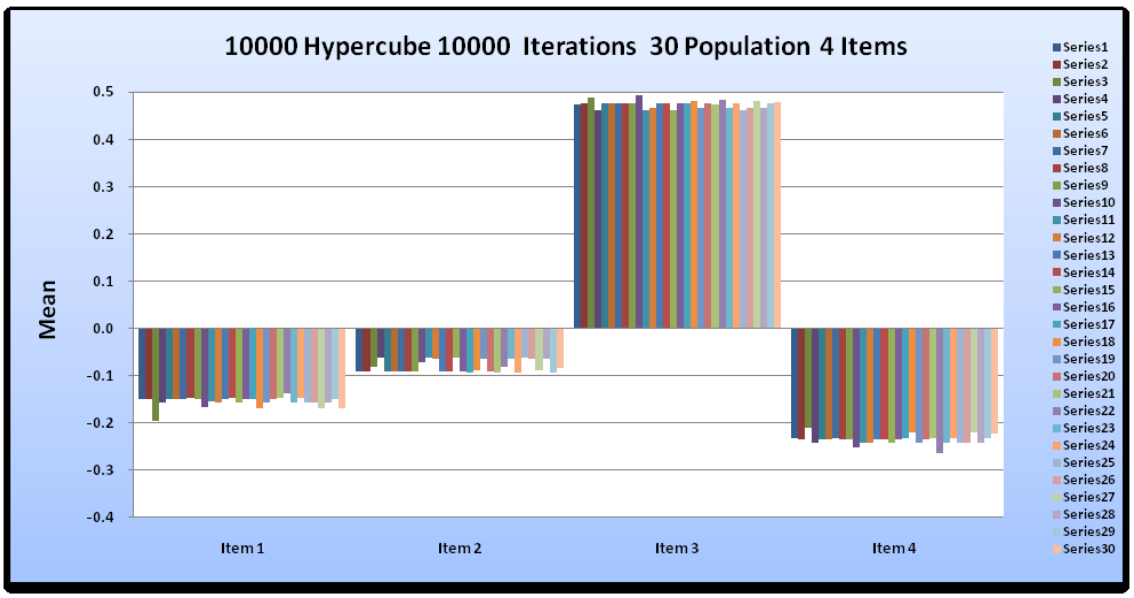


Figure 4 - 10000 Hypercube 10000 Iterations 30 Population 4 Items

The above graph showed improvement on the previous one and a larger LHS of 18000 rows per item with the recommended 1000 iterations generated the next adaptation in search of greater consistency in the results.

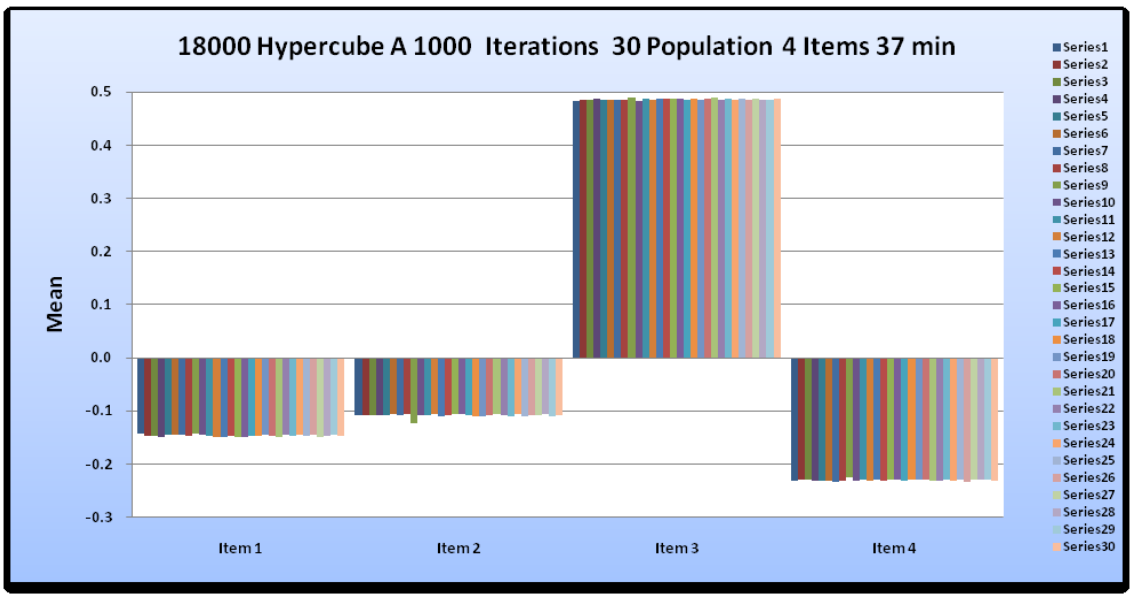


Figure 5 - 18000 Hypercube A 1000 Iterations 30 Population 4 Items 37 min

The above graph showed very good results and that the correct LHS size was very close to the 18000 mark, this led to the next adaptation of a 20000 rows per item LHS. The graph below showed that consistency had in fact gotten worse which is contrary to the expected outcome predicted by Cheng (2000) that the larger the sample the more precise the result.

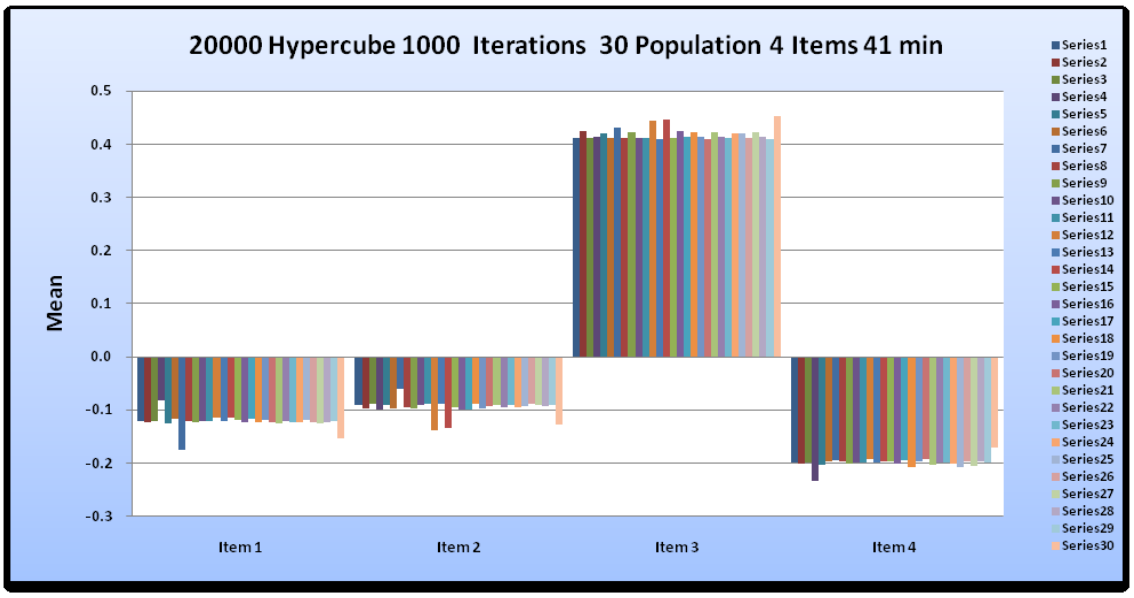


Figure 6 - 20000 Hypercube 1000 Iterations 30 Population 4 Items 41 min

The LHS relies on random sorting and it is likely that a random sorting could generate a more calculation friendly version, to check this required a re-sorting of the same LHS for 18000 rows per item, now labelled C, to generate the next set of data.

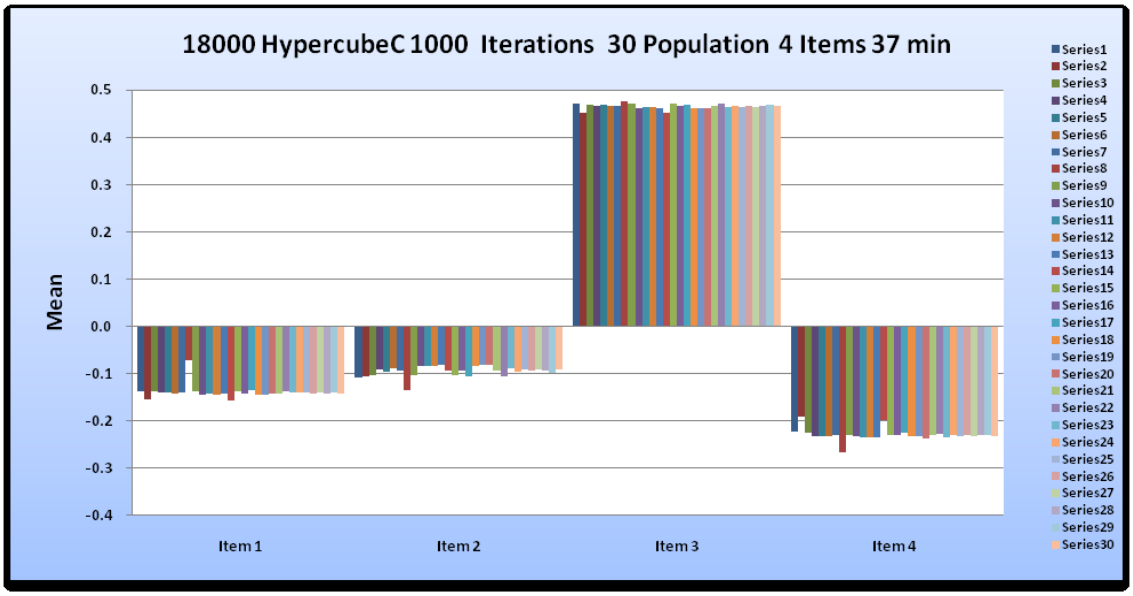


Figure 7 - 18000 Hypercube C 1000 Iterations 30 Population 4 Items 37 min

The above graph shows that the consistency is very similar to that of the 20000 rows per item LHS and another resorting produced the graph below based on a new LHS, labelled D, which showed yet another variation in consistency.

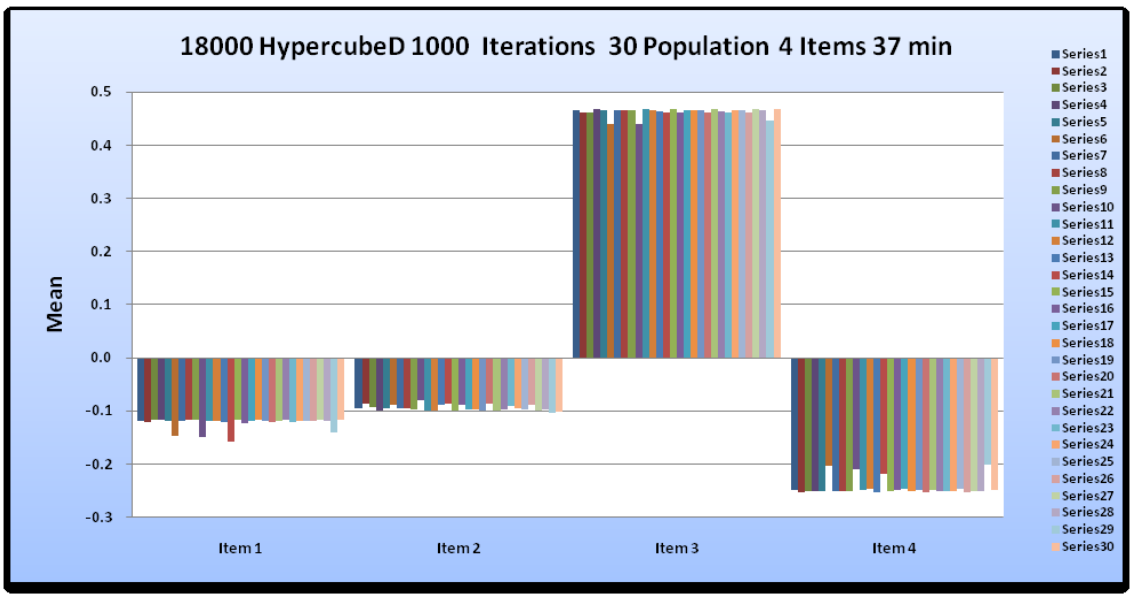


Figure 8 - 18000 Hypercube D 1000 Iterations 30 Population 4 Items 37 min

At this point, the question of how big should the range of LHS sizes be in order to get some level of consistency is raised; a LHS of 100000 rows per item generated the next set of data to try and answer that question.

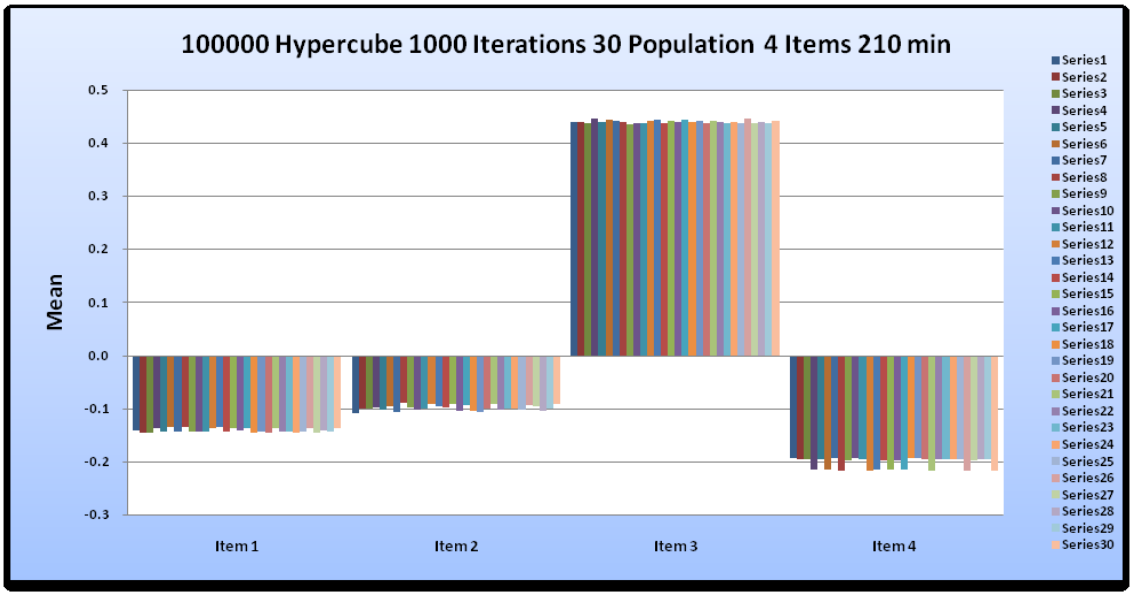


Figure 9 - 100000 Hypercube 1000 Iterations 30 Population 4 Items 210 min

Again, perfect consistency was questionable and LHS with a radical size of 1010000 rows per item generated the next set of data in the hopes of shooting past the required LHS size required for consistent results.

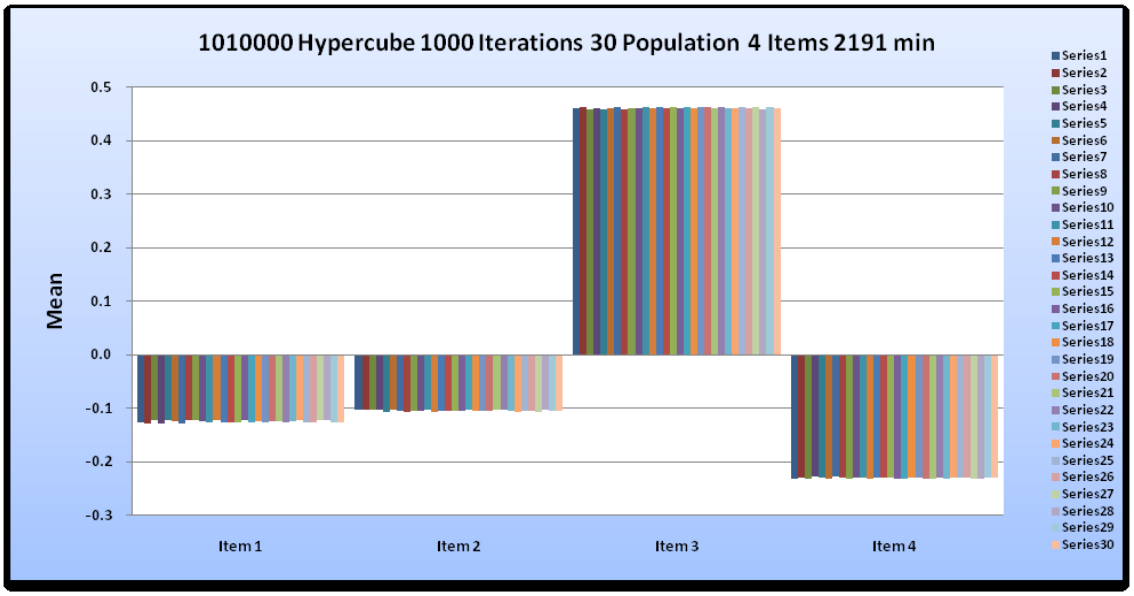


Figure 10 - 1010000 Hypercube 1000 Iterations 30 Population 4 Items 2191 min

The results showed a lot smoother consistent outcome however, this came at the high cost of a run time of 2191 minutes or approximately 1 ½ days and this only for 4 items.

4.3 Conclusion of initial model, Final Adopted Model

Using a maximum LHS of 1010000 proved to be the possible ideal however, the time proved to be especially prohibitive given the need to have re-sorted variations of the LHS for each of the picked sizes and across each variation of ranked items. There is some level of improvement between the 10000 and the 100000 LHSs this then decides the range of sizes for the LHS in the final experiment.

Taking into consideration, the re-sorting required to ensure a better spectrum of data based on the variations that could occur as well as the exponential impact on run times, it was decided to have four variations of each LHS labelled A,B,C and D.

The final experiment design therefore consisted of simulations run as follows:

- LHS range between 10000 and 90000, stepped in 20000 increases.
- Items for ranking will be 4, 6 and 8
- Four variations for each LHS

Therefore the total number of simulations:

= 3 (sets of ranked items) x 4 (variations of LHS) x 5 (LHS stepped increases)

= 60 simulations in total

A Dell D620 laptop with 1.8GHz Intel Duo Core CPU and 2 Gig Ram using Vista 32 Bit Operating System and the C# executable Algorithm program will run all the simulations. All data gathered regarding simulation times utilised the above combination of hardware and software and therefore are specific to that combination.

4.4 Results pertaining to changing the number of ranked items

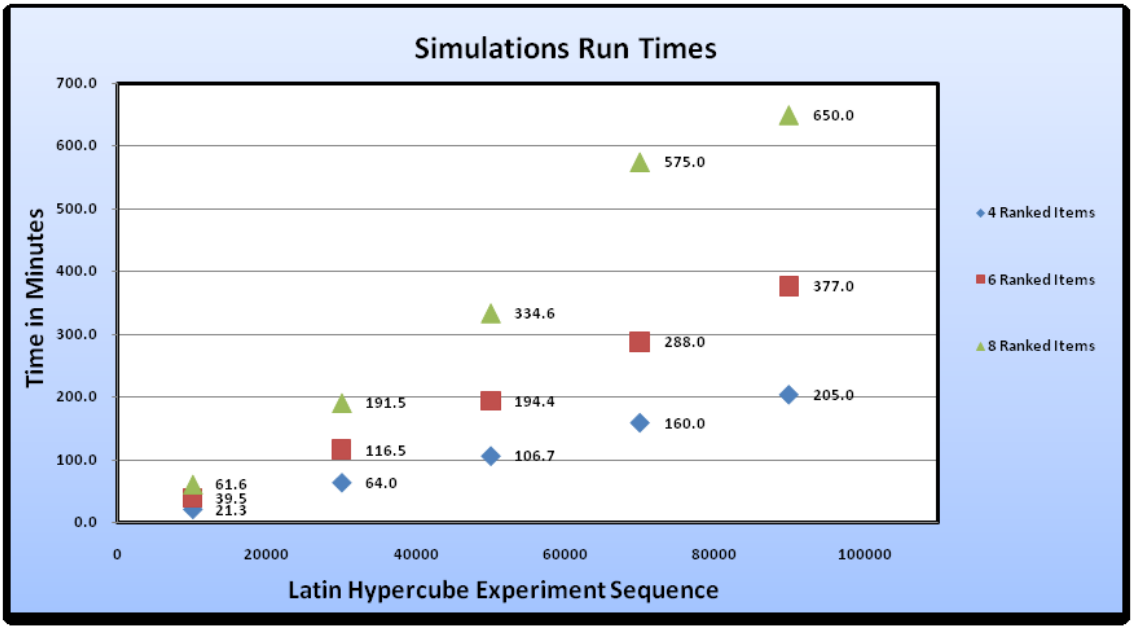


Figure 11 - Simulations Run Times

The above scatter plot chart is a plot of the approximate run times on the y-axis for each simulation with a certain LHS size on the x-axis, across each of the quantities of ranked items. The total time for all the simulations is calculated by adding all the times together and then multiplying it by 4 to account for the re-sorted LHS's.

= 4 x (4 ranks (21.3 + 64 + 106.7 + 160 + 205) + 6 ranks (39.5 + 116.5 + 194.4 + 288 + 377) + 8 ranks (61.6 + 191.5 + 334.6 + 575 + 650))

= 13540.4 minutes

= 9.4 days of continuous computer run time

Latin Hypercube Size & Sample		Standard Error of Sample mean			
		Four Ranks Item 1	Four Ranks Item 2	Four Ranks Item 3	Four Ranks Item 4
	10000 A	0.0022	0.0016	0.0013	0.0019
	10000 B	0.0084	0.0043	0.0018	0.0057
	10000 C	0.0012	0.0012	0.0006	0.0008
	10000 D	0.0037	0.0029	0.0030	0.0090
	30000 A	0.0012	0.0007	0.0009	0.0009
	30000 B	0.0020	0.0037	0.0011	0.0012
	30000 C	0.0019	0.0017	0.0018	0.0013
	30000 D	0.0024	0.0012	0.0010	0.0013
	50000 A	0.0040	0.0013	0.0013	0.0025
	50000 B	0.0009	0.0020	0.0007	0.0009
	50000 C	0.0029	0.0013	0.0011	0.0015
	50000 D	0.0003	0.0012	0.0005	0.0006
	70000 A	0.0005	0.0010	0.0005	0.0005
	70000 B	0.0035	0.0016	0.0008	0.0013
	70000 C	0.0008	0.0007	0.0012	0.0010
	70000 D	0.0017	0.0014	0.0012	0.0032
	90000 A	0.0004	0.0009	0.0008	0.0008
	90000 B	0.0012	0.0012	0.0011	0.0030
	90000 C	0.0013	0.0014	0.0010	0.0010
	90000 D	0.0005	0.0009	0.0007	0.0007

Table 6 - Standard Error of Sample Mean Four Ranks

The above table is the complete set of SE($\overline{\mathbf{X}}$)'s for 4 ranked items across each simulation variation and for each item.

Latin Hypercube Size & Sample		Standard Error of Sample mean					
		Six Ranks Item 1	Six Ranks Item 2	Six Ranks Item 3	Six Ranks Item 4	Six Ranks Item 5	Six Ranks Item 6
	10000 A	0.0031	0.0020	0.0037	0.0093	0.0029	0.0043
	10000 B	0.0020	0.0034	0.0063	0.0021	0.0045	0.0046
	10000 C	0.0040	0.0024	0.0049	0.0037	0.0032	0.0049
	10000 D	0.0038	0.0030	0.0047	0.0042	0.0029	0.0076
	30000 A	0.0009	0.0007	0.0007	0.0013	0.0010	0.0013
	30000 B	0.0013	0.0013	0.0016	0.0006	0.0007	0.0007
	30000 C	0.0023	0.0038	0.0023	0.0031	0.0027	0.0017
	30000 D	0.0035	0.0009	0.0015	0.0014	0.0018	0.0012
	50000 A	0.0012	0.0009	0.0014	0.0008	0.0007	0.0015
	50000 B	0.0013	0.0009	0.0021	0.0013	0.0011	0.0012
	50000 C	0.0013	0.0020	0.0017	0.0017	0.0009	0.0012
	50000 D	0.0009	0.0014	0.0012	0.0016	0.0009	0.0016
	70000 A	0.0015	0.0014	0.0008	0.0026	0.0021	0.0017
	70000 B	0.0007	0.0006	0.0009	0.0009	0.0006	0.0014
	70000 C	0.0011	0.0015	0.0009	0.0012	0.0012	0.0011
	70000 D	0.0014	0.0009	0.0014	0.0009	0.0007	0.0008
	90000 A	0.0017	0.0010	0.0012	0.0035	0.0011	0.0012
	90000 B	0.0010	0.0011	0.0010	0.0007	0.0014	0.0007
	90000 C	0.0010	0.0007	0.0015	0.0015	0.0010	0.0008
	90000 D	0.0013	0.0013	0.0011	0.0008	0.0024	0.0008

Table 7 - Standard Error of Sample Mean Six Ranks

The above table is the complete set of SE($\bar{X}$)'s for 6 ranked items across each simulation variation and for each item.

The table below is the complete set of SE($\bar{X}$)'s for 8 ranked items across each simulation variation and for each item.

Latin Hypercube Size & Sample		Standard Error of Sample mean							
		Eight Ranks Item 1	Eight Ranks Item 2	Eight Ranks Item 3	Eight Ranks Item 4	Eight Ranks Item 5	Eight Ranks Item 6	Eight Ranks Item 7	Eight Ranks Item 8
10000	A	0.0062	0.0041	0.0054	0.0020	0.0044	0.0053	0.0056	0.0039
10000	B	0.0080	0.0061	0.0061	0.0058	0.0098	0.0093	0.0135	0.0051
10000	C	0.0055	0.0054	0.0057	0.0066	0.0047	0.0076	0.0064	0.0054
10000	D	0.0048	0.0069	0.0040	0.0031	0.0050	0.0042	0.0031	0.0054
30000	A	0.0018	0.0027	0.0041	0.0031	0.0041	0.0064	0.0030	0.0043
30000	B	0.0074	0.0038	0.0036	0.0028	0.0037	0.0038	0.0022	0.0026
30000	C	0.0022	0.0036	0.0023	0.0052	0.0026	0.0040	0.0031	0.0021
30000	D	0.0027	0.0038	0.0041	0.0033	0.0030	0.0074	0.0038	0.0029
50000	A	0.0042	0.0021	0.0027	0.0040	0.0037	0.0034	0.0025	0.0020
50000	B	0.0021	0.0033	0.0019	0.0022	0.0031	0.0030	0.0038	0.0028
50000	C	0.0022	0.0053	0.0040	0.0029	0.0032	0.0043	0.0045	0.0021
50000	D	0.0052	0.0022	0.0030	0.0031	0.0035	0.0036	0.0032	0.0030
70000	A	0.0011	0.0013	0.0020	0.0011	0.0016	0.0026	0.0026	0.0027
70000	B	0.0027	0.0014	0.0012	0.0019	0.0037	0.0012	0.0011	0.0017
70000	C	0.0036	0.0041	0.0023	0.0026	0.0021	0.0038	0.0022	0.0022
70000	D	0.0030	0.0014	0.0018	0.0031	0.0029	0.0026	0.0023	0.0012
90000	A	0.0028	0.0033	0.0014	0.0021	0.0018	0.0026	0.0026	0.0015
90000	B	0.0025	0.0022	0.0032	0.0024	0.0028	0.0024	0.0028	0.0017
90000	C	0.0028	0.0029	0.0037	0.0033	0.0029	0.0025	0.0041	0.0026
90000	D	0.0021	0.0011	0.0028	0.0019	0.0054	0.0048	0.0019	0.0030

Table 8 - Standard Error of Sample Mean Eight Ranks

Latin Hypercube Size & Sample	Mean of the Standard Error of Sample Means		
	Four Ranks	Six Ranks	Eight Ranks
10000	0.0031	0.0041	0.0058
30000	0.0015	0.0016	0.0036
50000	0.0014	0.0013	0.0032
70000	0.0013	0.0012	0.0022
90000	0.0011	0.0012	0.0027
1010000	0.0003		

Table 9 - Mean of the Standard Error of Sample Mean for All Ranks

In the above table are the means taken from the SE($\overline{\mathbf{X}}$)'s, aggregated across all the simulation variations and individual items, it also includes a single SE($\overline{\mathbf{X}}$) value from the initial model run on 4 ranks for a LHS size of 1010000.

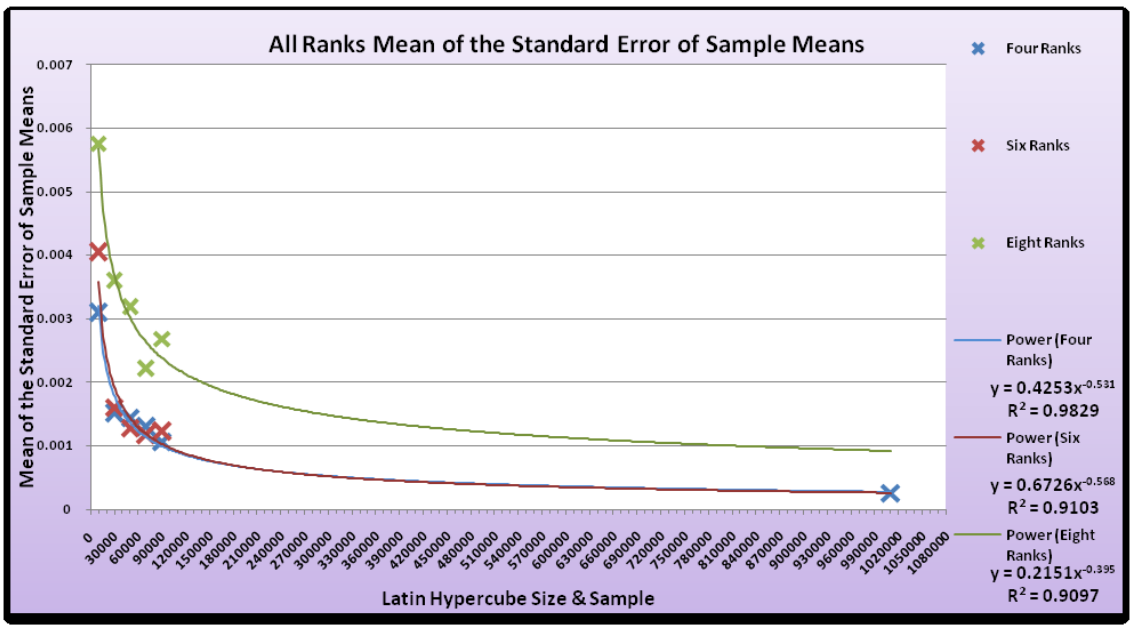


Figure 12 - All Ranks Mean of Standard Error of Sample Means

The above chart is the graphic representation of the previous table using a scatter plot chart and including trend lines for each quantity of ranked items. Also included is the Excel® formula for the trend lines as well as its associated R^2 value in the legend.

		Ranked Items Ratios		
		Four Ranks	Six Ranks	Eight Ranks
Latin Hypercube Size & Sample	10000	1	1.309	1.855
	30000	1	1.050	2.371
	50000	1	0.888	2.221
	70000	1	0.905	1.703
	90000	1	1.159	2.514
	Average	1	1.062	2.133

Table 10 - All Ranked Items Ratios of Comparison

The above table is a ratio of comparison between the quantities of ranked items across all LHS sizes and summarised in the last line by their mean value. The values used to derive the above came from **Table 9**.

4.5 Results pertaining to changing the Simulated Sample size

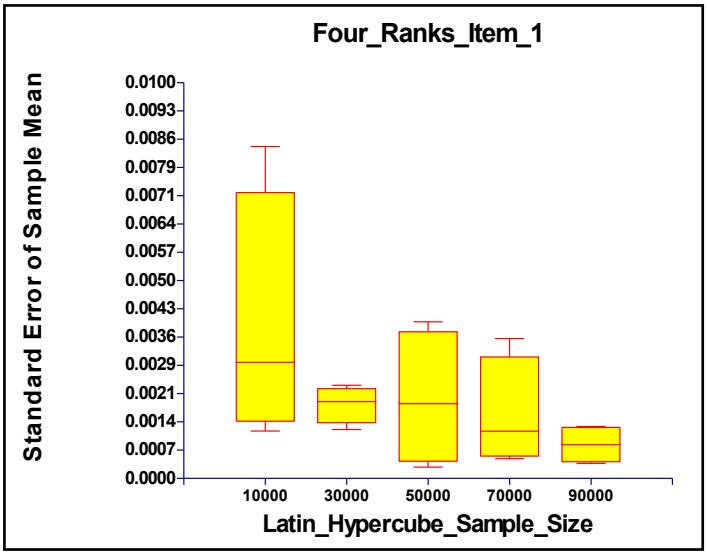


Figure 13 - Box Plot Four Ranks Item 1

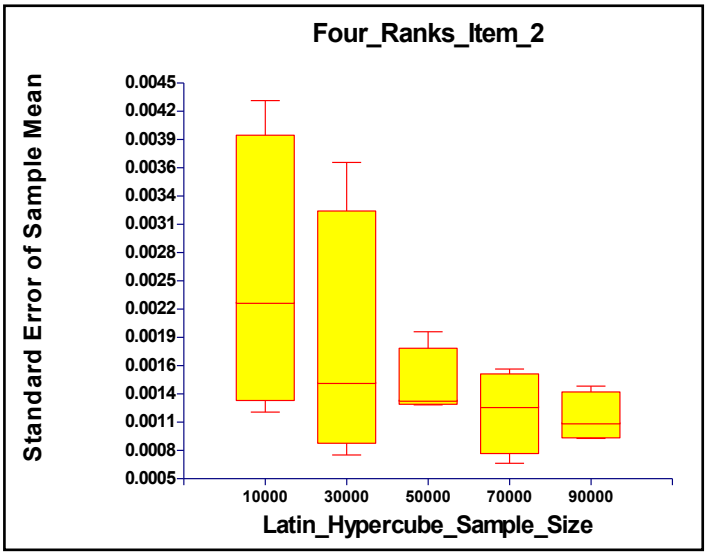


Figure 14 - Box Plot Four Ranks Item 2

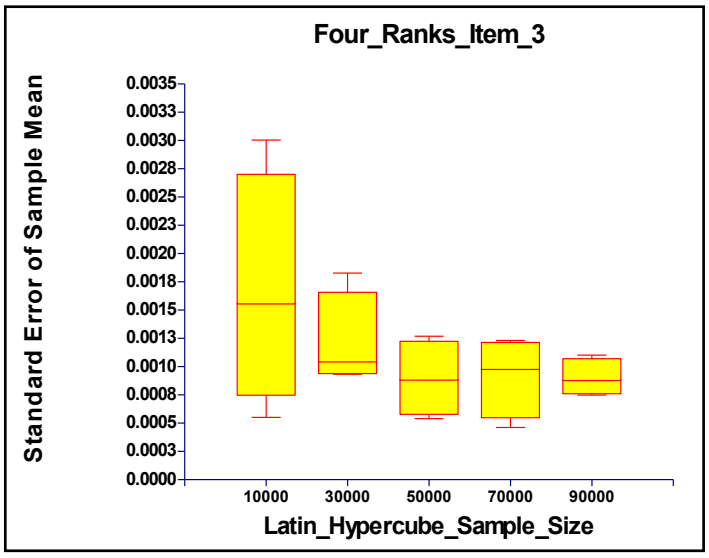


Figure 15 - Box Plots Four Ranks Item 3

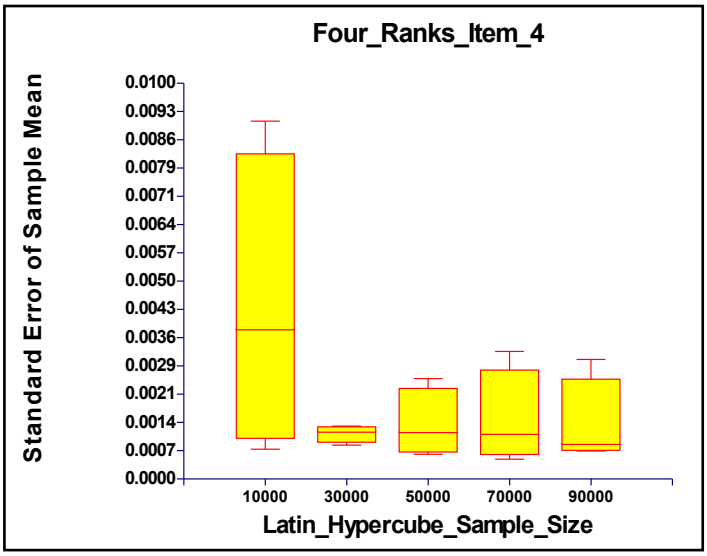


Figure 16 - Box Plot Four Ranks Item 4

The above four figures are the box plot graphs of all the available SE($\overline{\mathbf{X}}$) values for each of the four ranked items. The box plot graph will give indication of the spread of data for each stepped increase of the LHS across all variations.

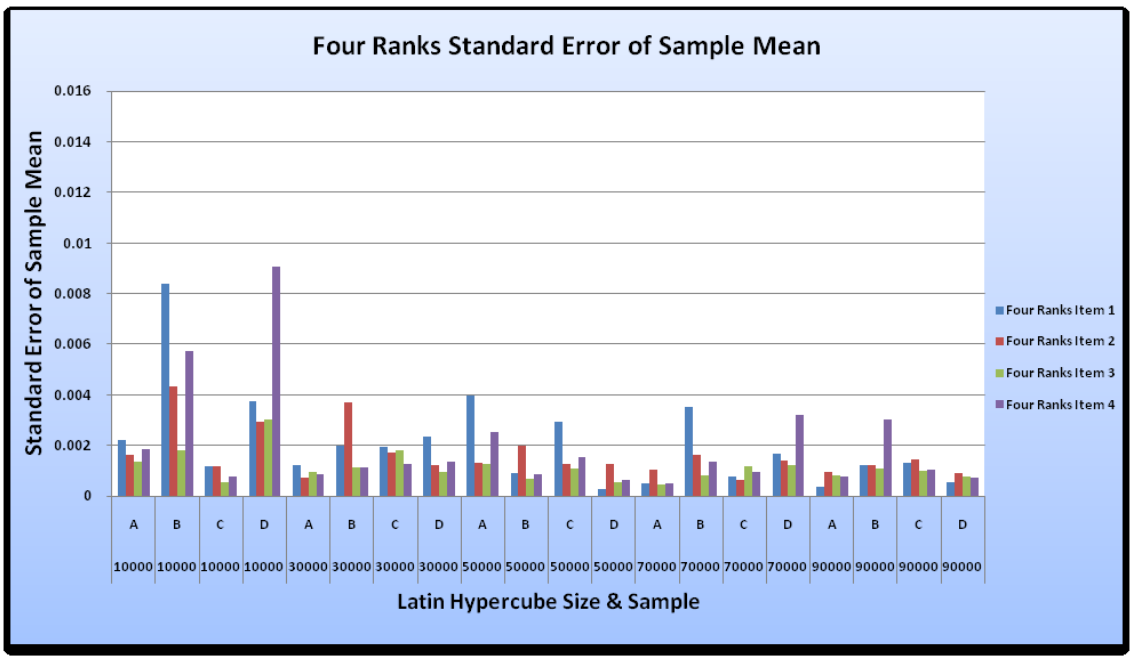


Figure 17 - Four Ranks Standard Error of Sample Mean

The above bar graph represents the same data used in the box plots but has separated the data sets of the various re-sorted LHS's for individual comparative analysis of the data surrounding four ranked items.

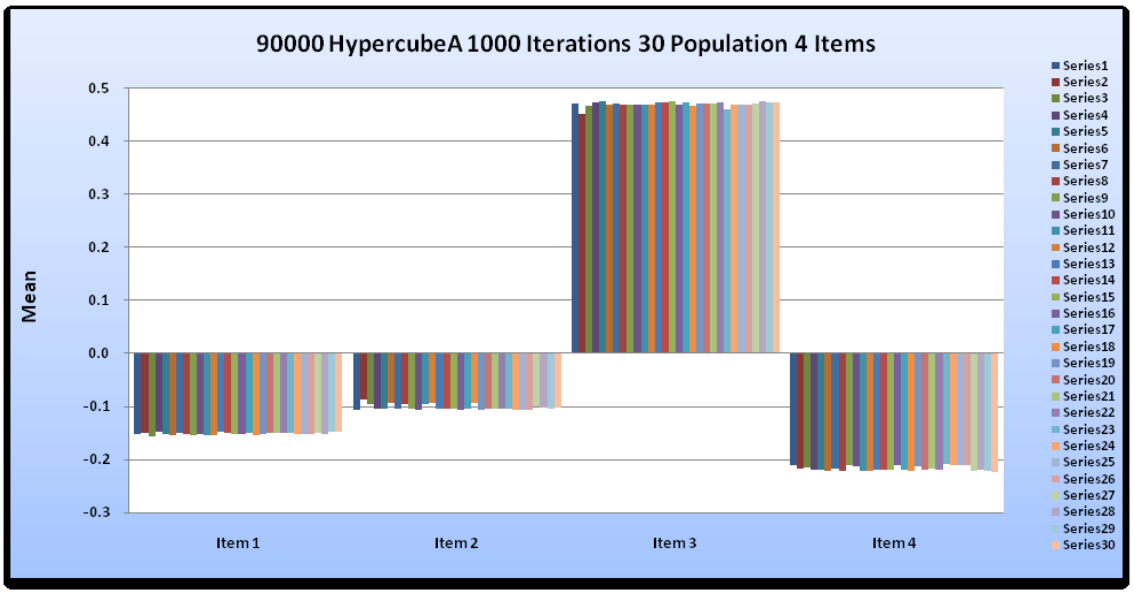


Figure 18 - 90000 Hypercube A 1000 Iterations 30 Population 4 Items

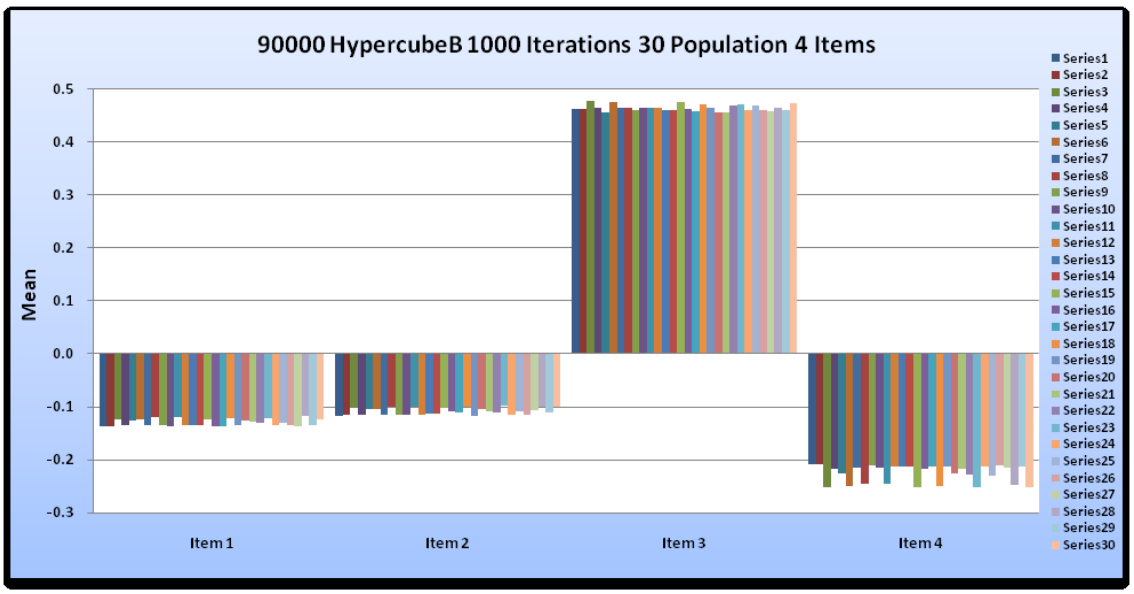


Figure 19 - 90000 Hypercube B 1000 Iterations 30 Population 4 Items

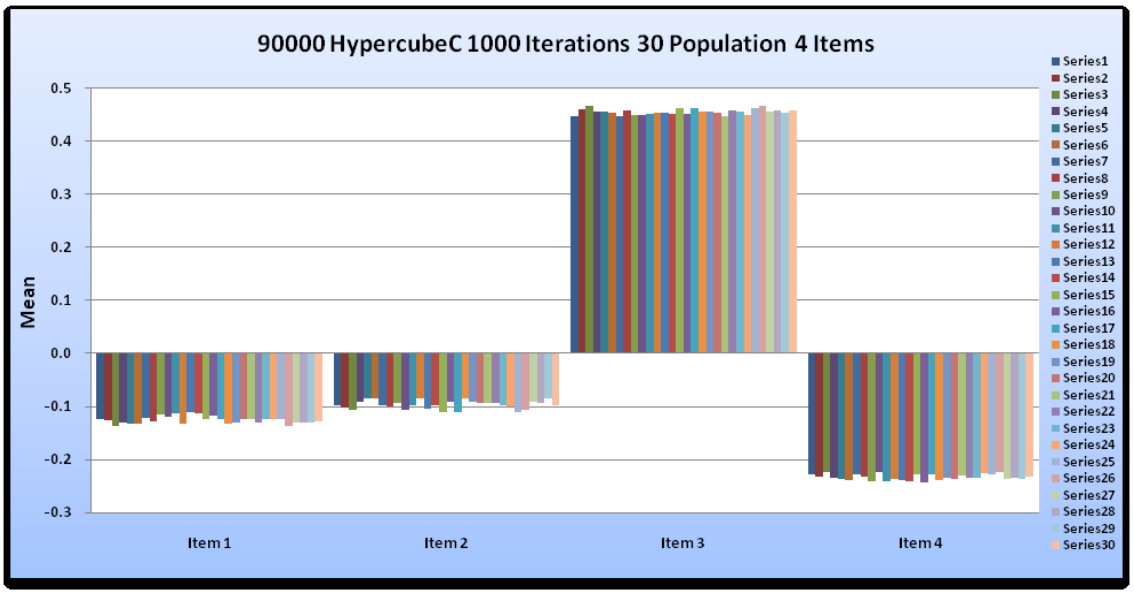


Figure 20 - 90000 Hypercube C 1000 Iterations 30 Population 4 Items

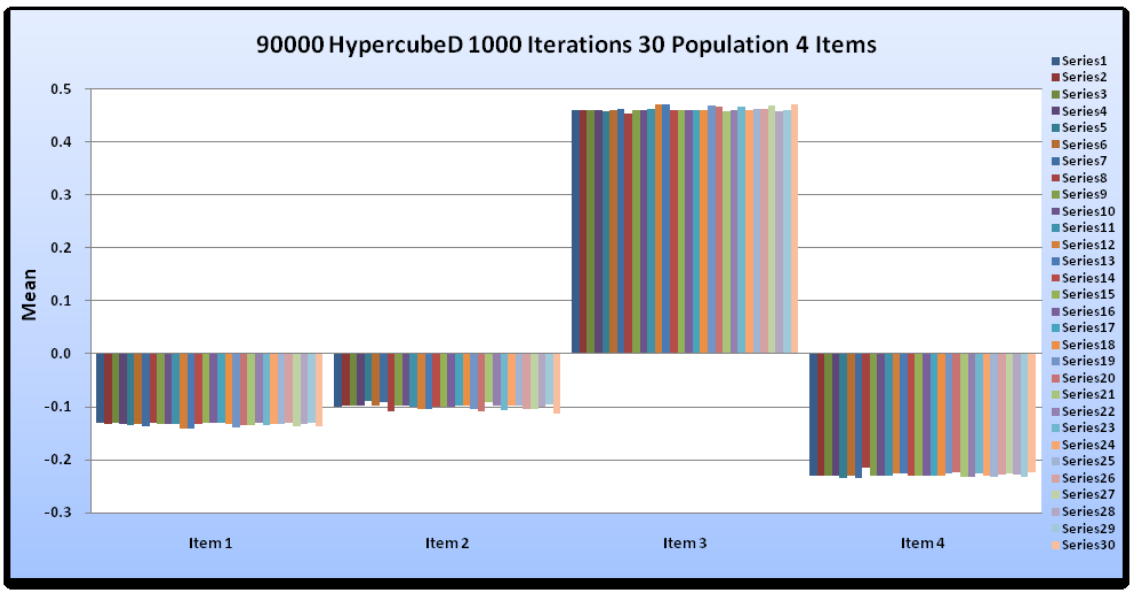


Figure 21 - 90000 Hypercube D 1000 Iterations 30 Population 4 Items

The above 4 bar graphs are the result of the outputs for the means for the four variations of a LHS with 90000 rows per item for 4 ranks. The graphs display all 30 population points for each item.

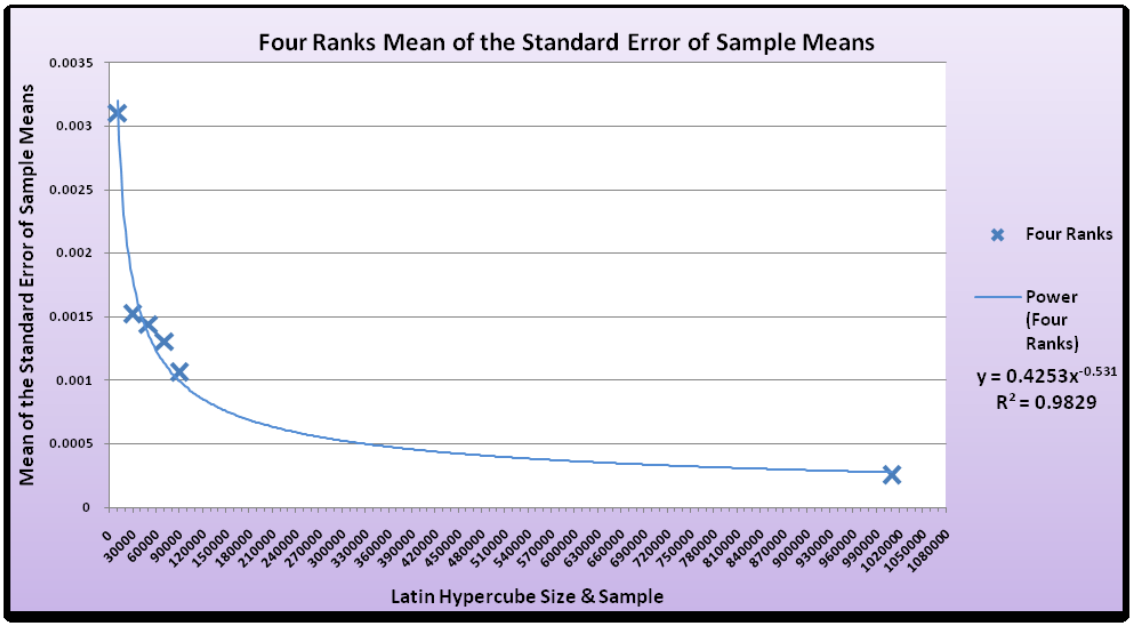


Figure 22 - Four Ranks Mean of the Standard Error of Sample Means

The above graph is a duplication of Figure 12 but this time only reflecting the data surrounding 4 ranks.

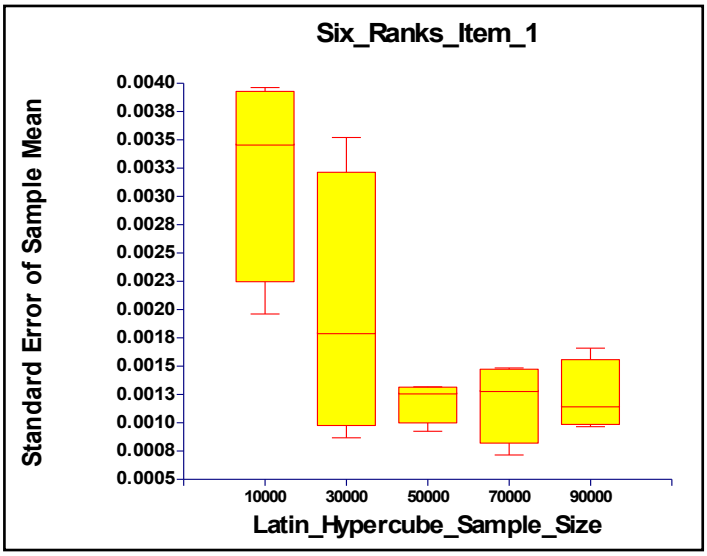


Figure 23 - Box Plot Six Ranks Item 1

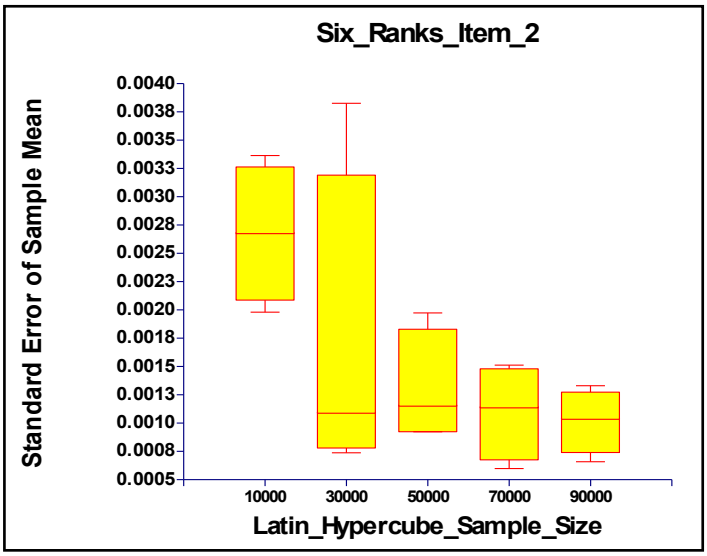


Figure 24 - Box Plot Six Ranks Item 2

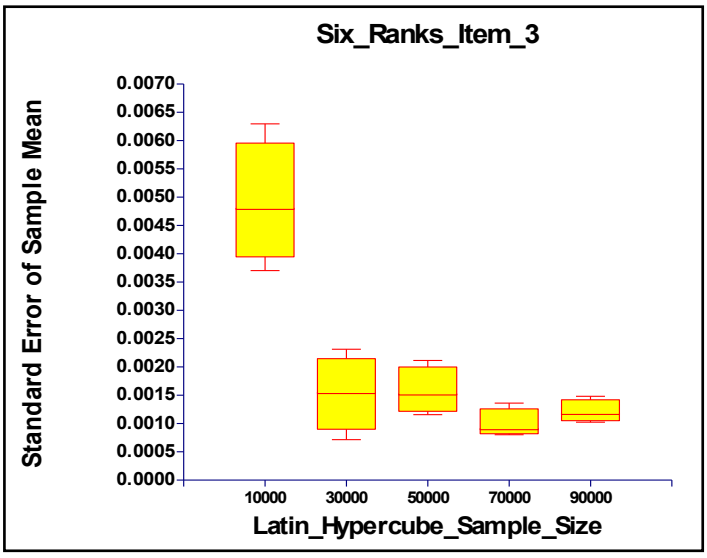


Figure 25 - Box Plot Six Ranks Item 3

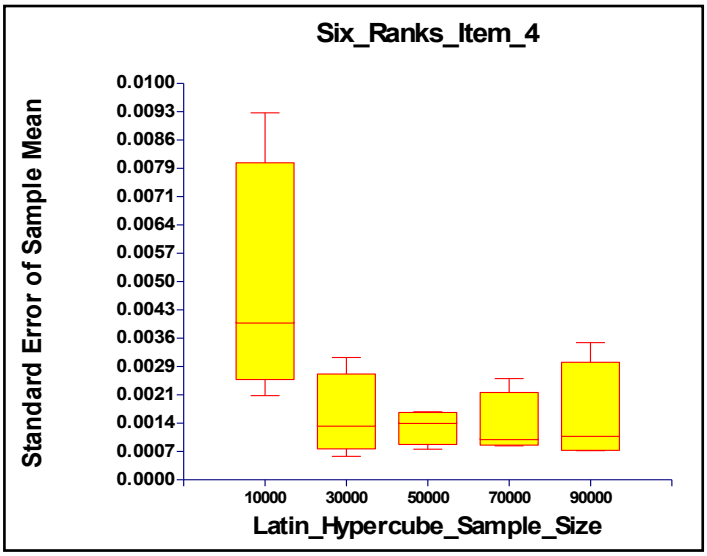


Figure 26 - Box Plot Six Ranks Item 4

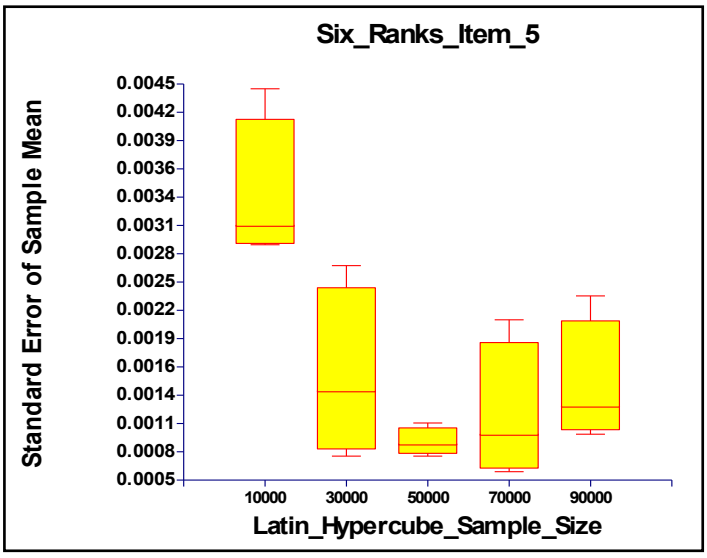


Figure 27 - Box Plot Six Ranks Item 5

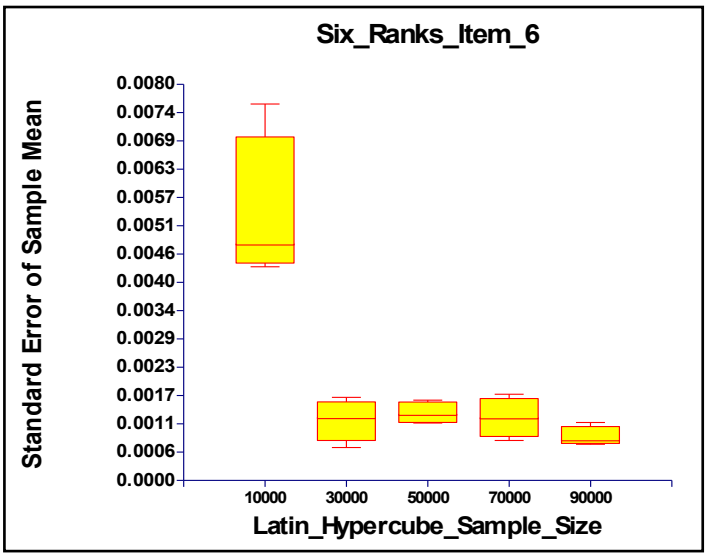


Figure 28 – Box Plot Six Ranks Item 6

The above six figures are the box plot graphs of all the available SE($\overline{\mathbf{X}}$) values for each of the six ranked items. The box plot graphs will give an indication of the spread of data for each stepped increase of the LHS across all variations.

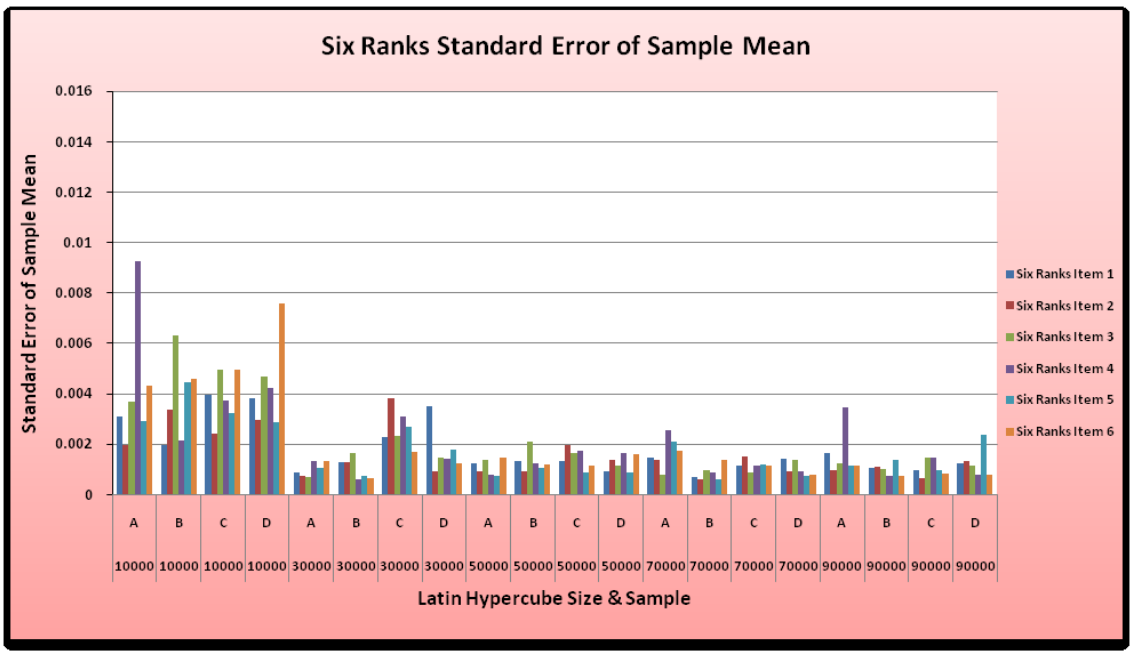


Figure 29 - Six Ranks Standard Error of Sample Mean

The above bar graph represents the same data used in the box plots but has separated the data sets of the various re-sorted LHS's for individual comparative analysis of the data surrounding six ranked items.

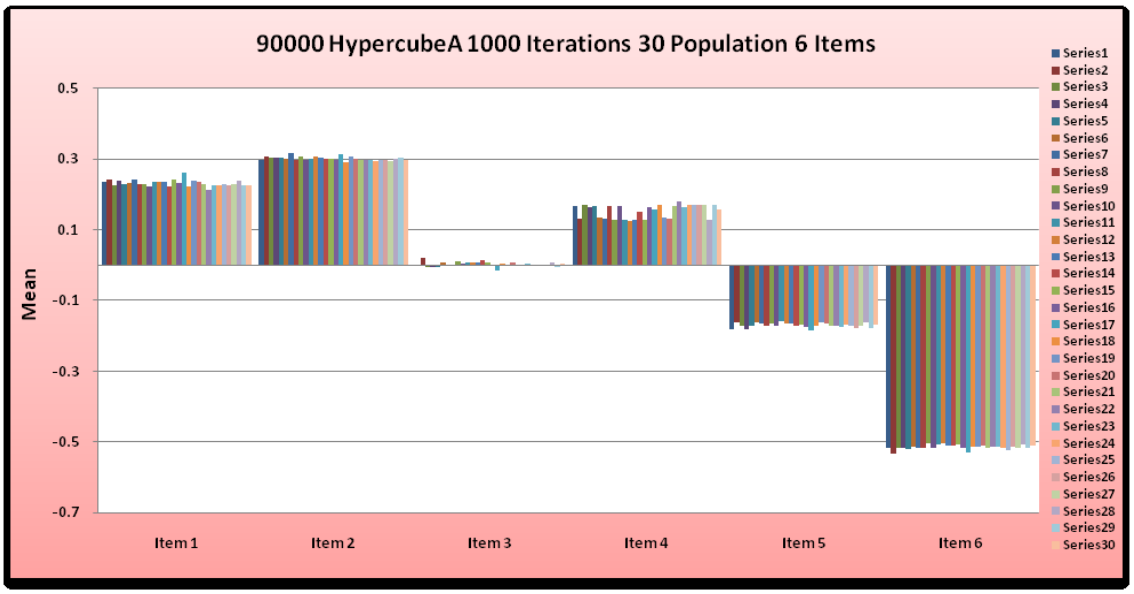


Figure 30 - 90000 Hypercube A 1000 Iterations 30 Population 6 Items

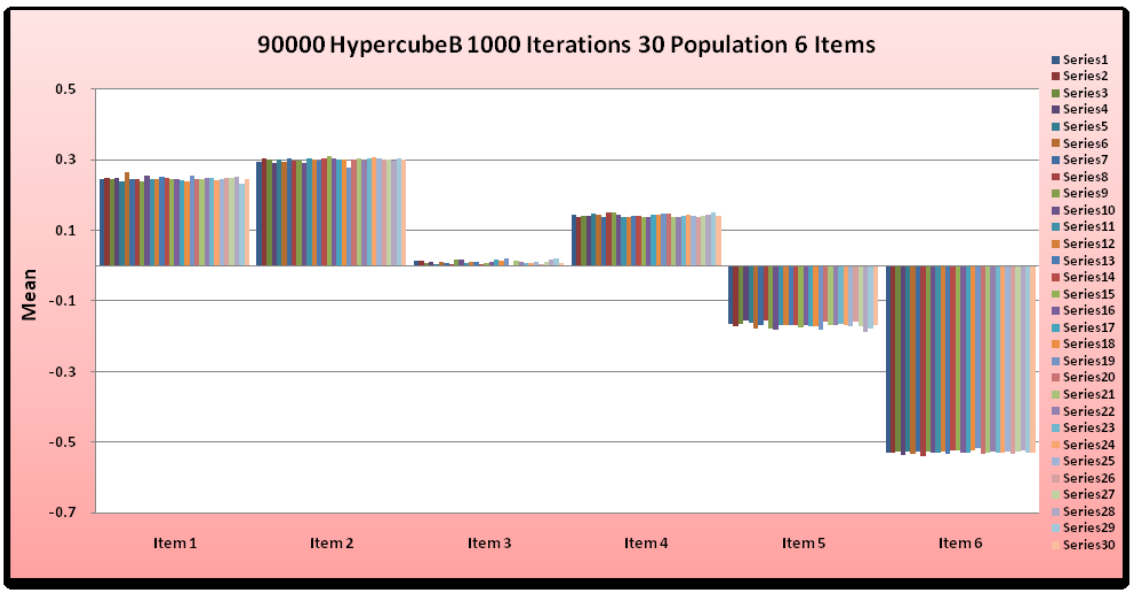


Figure 31 - 90000 Hypercube B 1000 Iterations 30 Population 6 Items

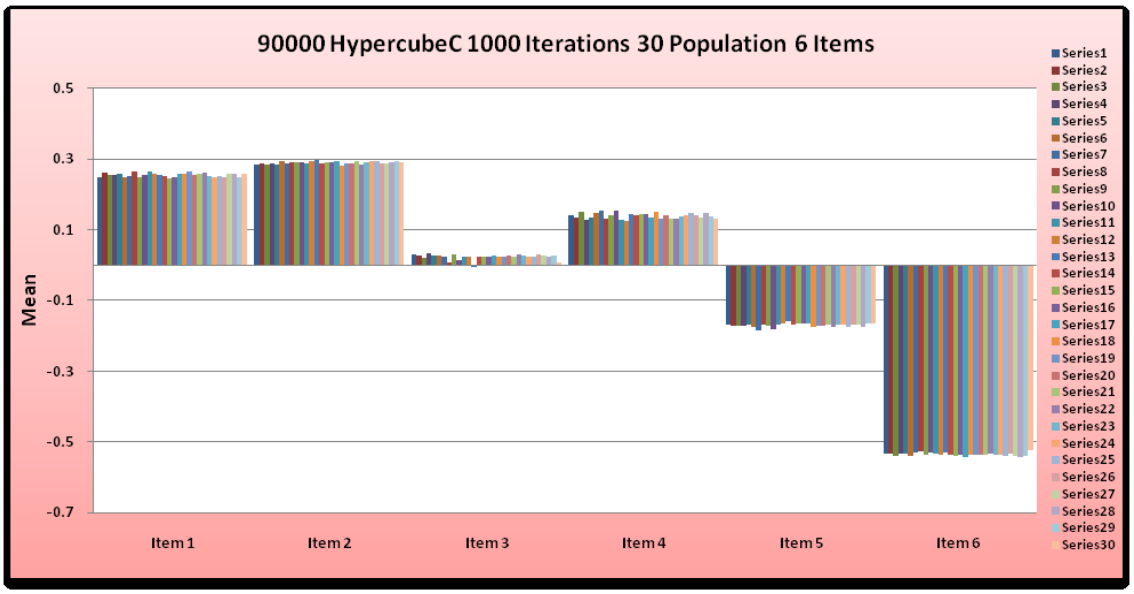


Figure 32 - 90000 Hypercube C 1000 Iterations 30 Population 6 Items

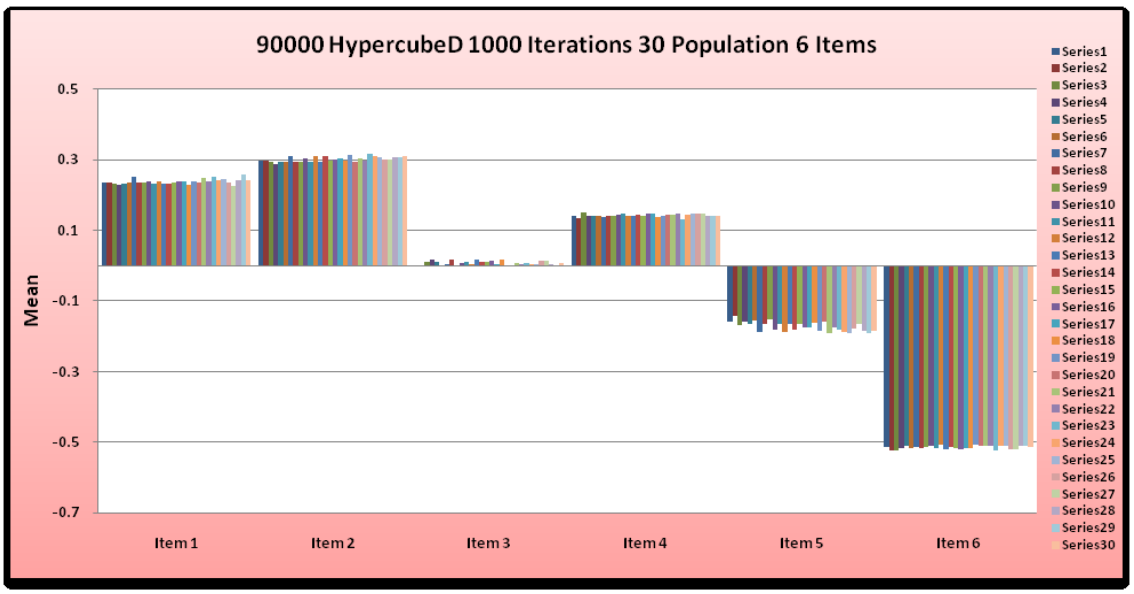


Figure 33 - 90000 Hypercube D 1000 Iterations 30 Population 6 Items

The above 4 bar graphs are the result of the outputs for the means for the four variations of a LHS with 90000 rows per item for 6 ranks. The graphs display all 30 population points for each item.

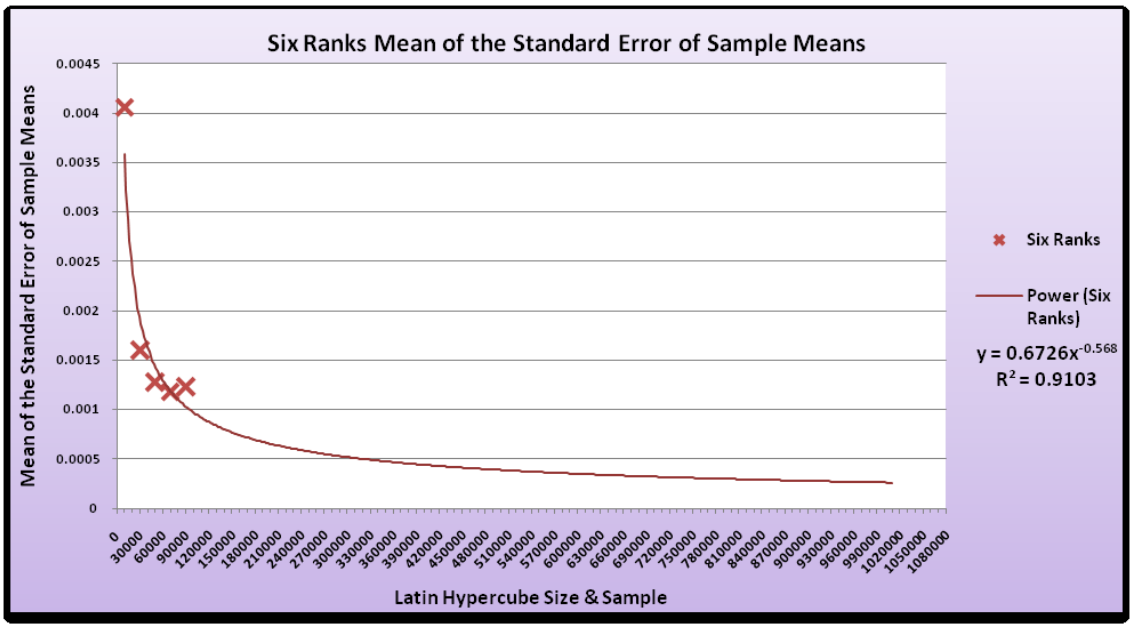


Figure 34 - Six Ranks Mean of the Standard Error of Sample Means

The above graph is a duplication of Figure 12 but this time only reflecting the data surrounding 6 ranks.

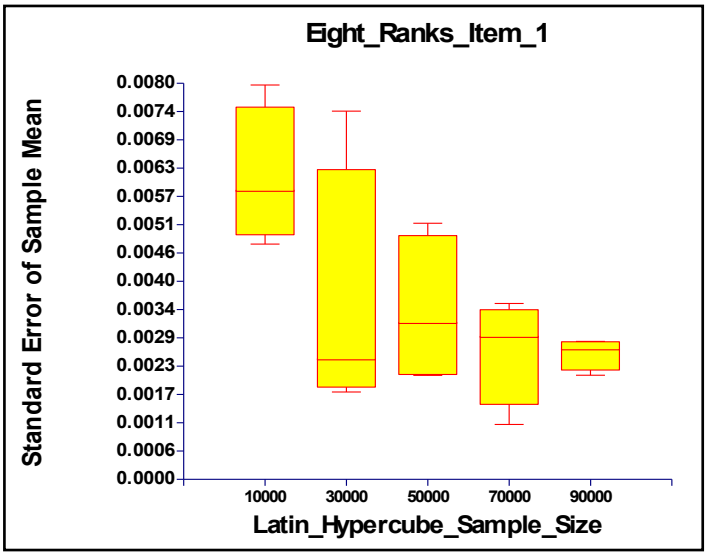


Figure 35 - Box Plot Eight Ranks Item 1

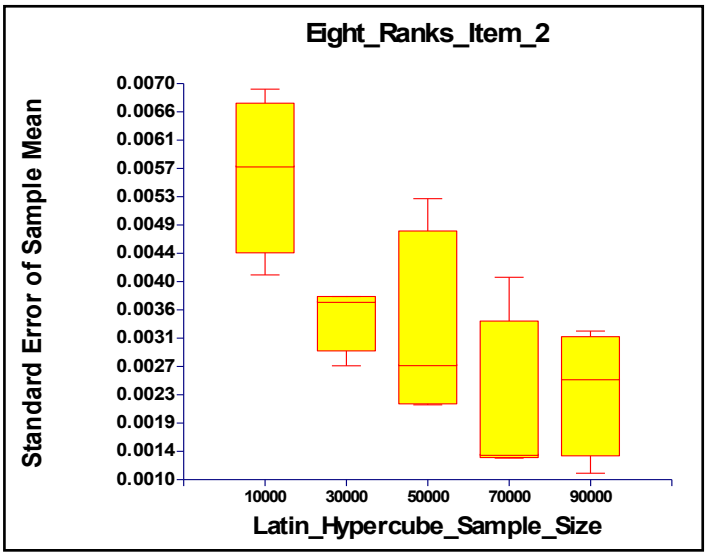


Figure 36 - Box Plot Eight Ranks Item 2

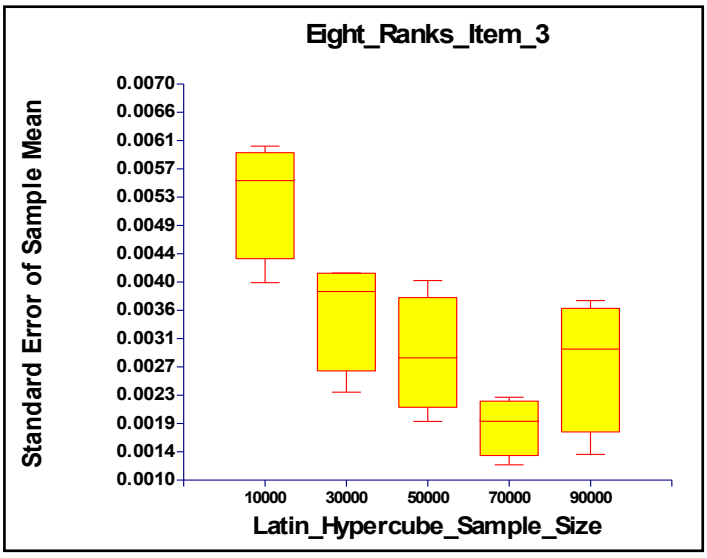


Figure 37 - Box Plot Eight Ranks Item 3

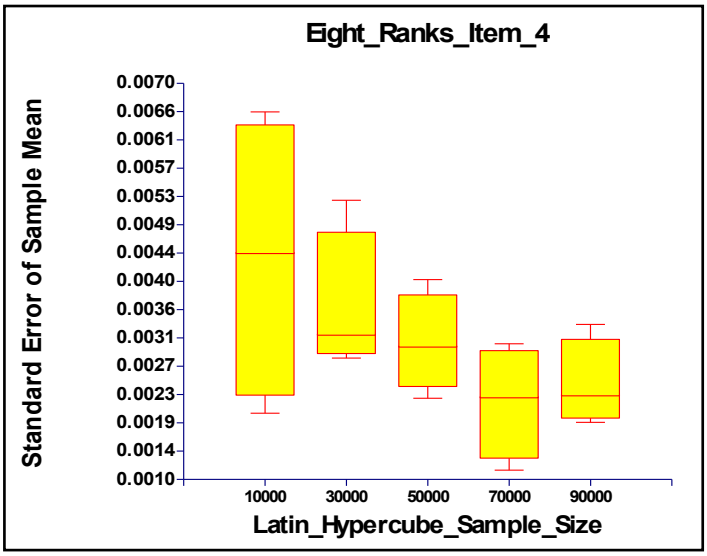


Figure 38 - Box Plot Eight Ranks Item 4

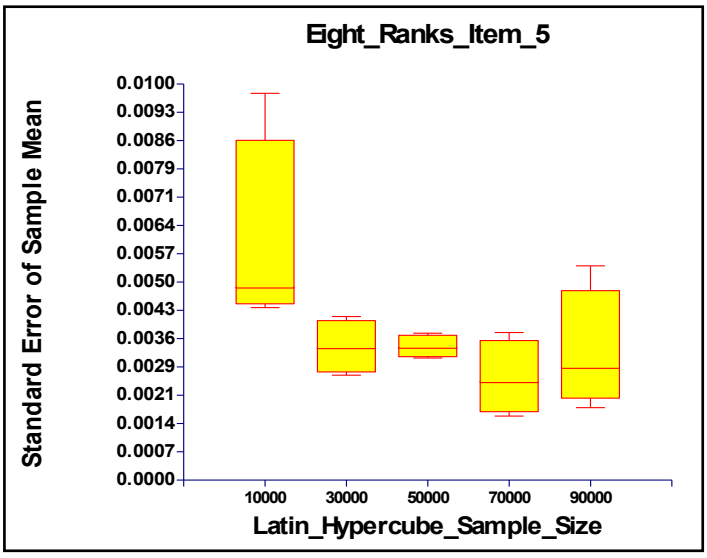


Figure 39 - Box Plot Eight Ranks Item 5

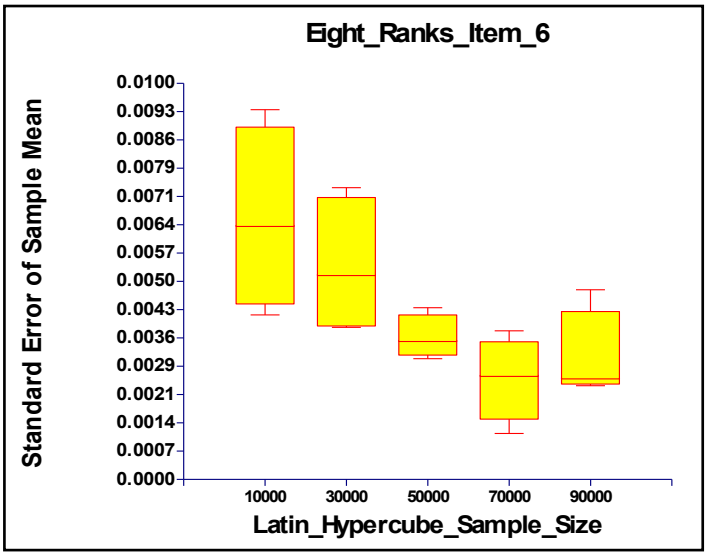


Figure 40 – Box Plot Eight Ranks Item 6

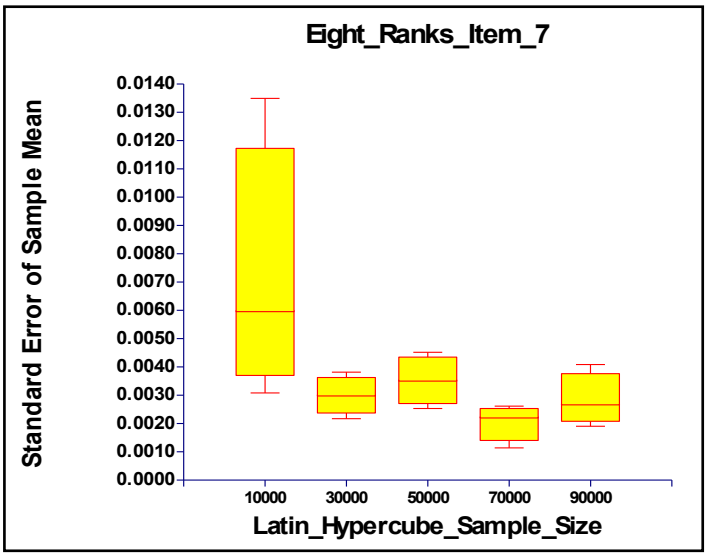


Figure 41 - Box Plot Eight Ranks Item 7

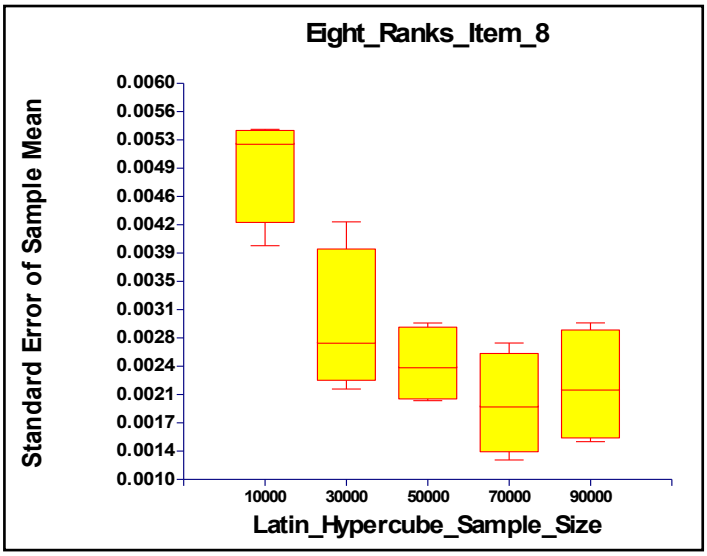


Figure 42 - Box Plot Eight Ranks Item 8

The above eight figures are the box plot graphs of all the available SE($\bar{\mathbf{X}}$) values for each of the eight ranked items. The box plot graph will give indication of the spread of data for each stepped increase of the LHS across all variations.

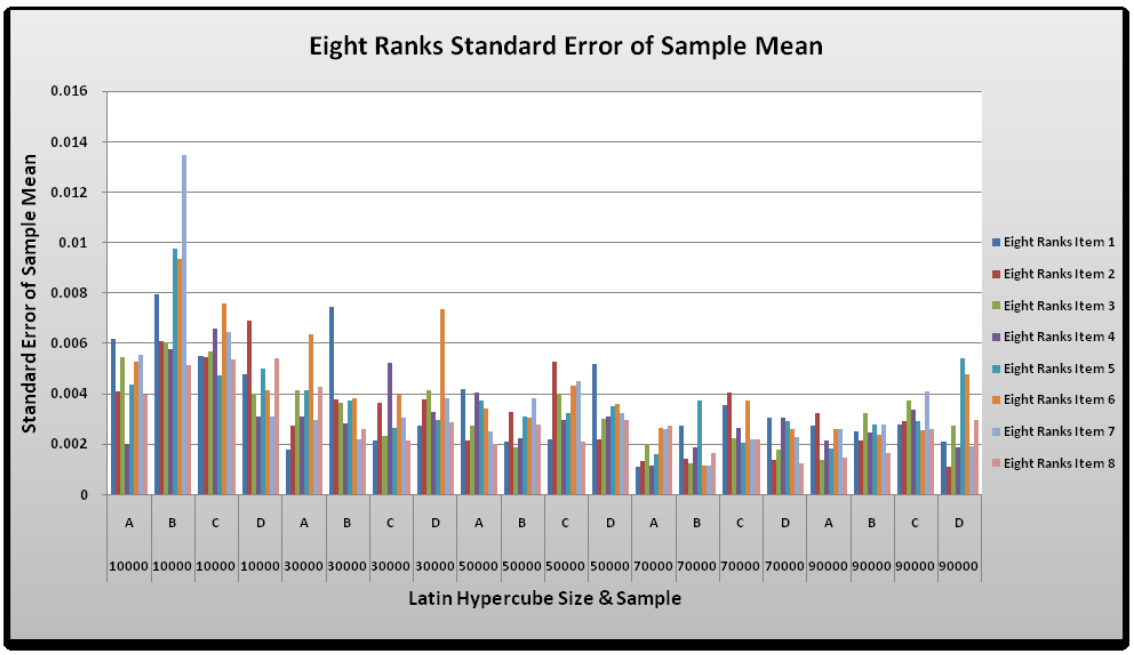


Figure 43 - Eight Ranks Standard Error of Sample Mean

The above bar graph represents the same data used in the box plots but has separated the data sets of the various re-sorted LHS's for individual comparative analysis of the data surrounding eight ranked items.

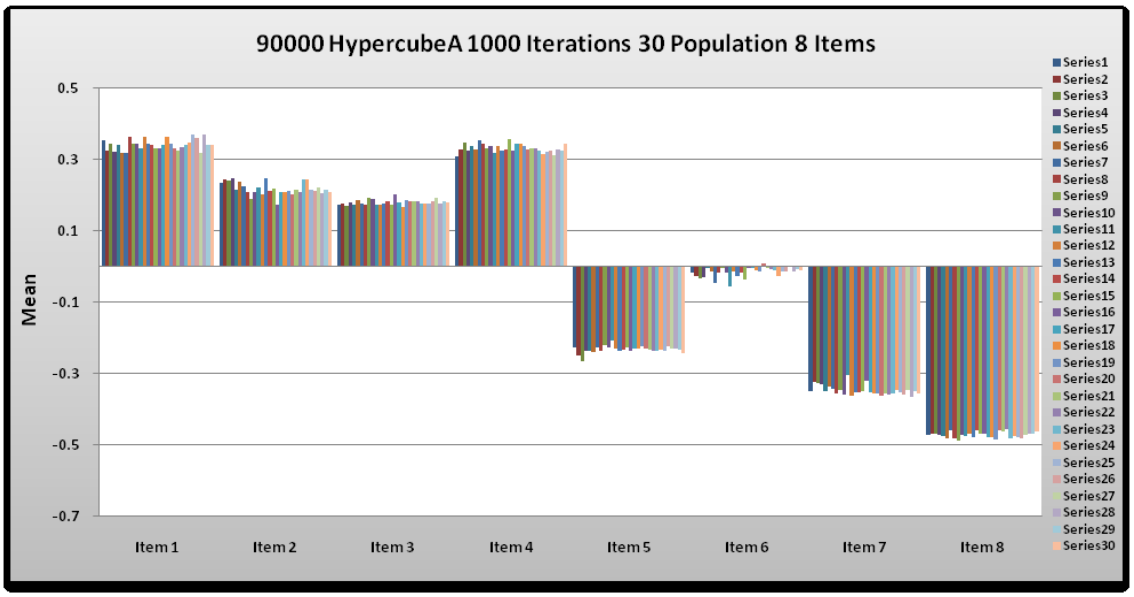


Figure 44 - 90000 Hypercube A 1000 Iterations 30 Population 8 Items

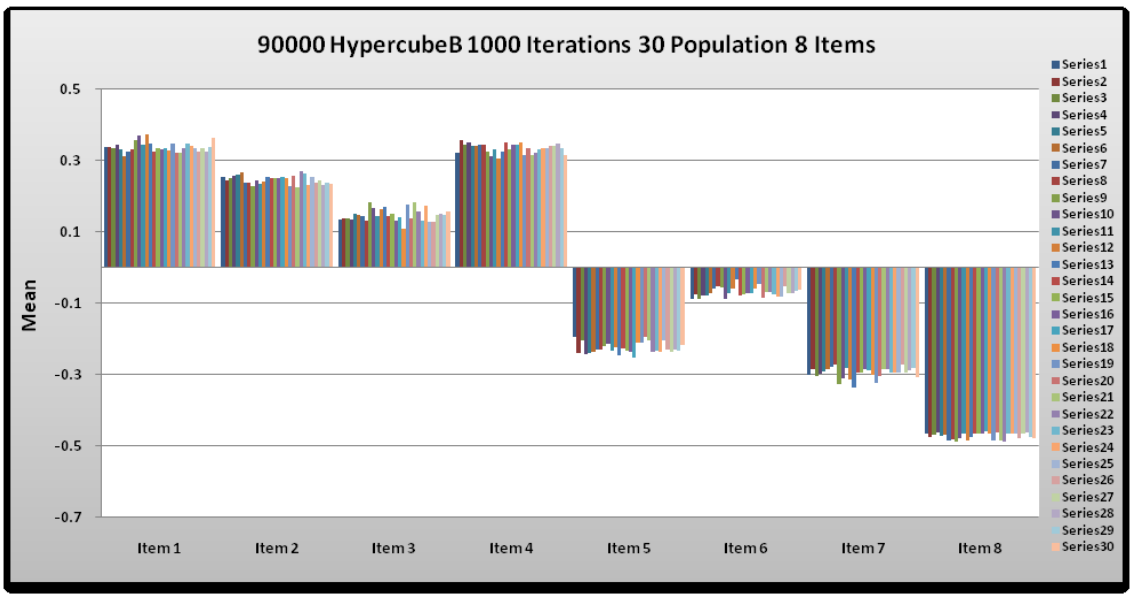


Figure 45 - 90000 Hypercube A 1000 Iterations 30 Population 8 Items

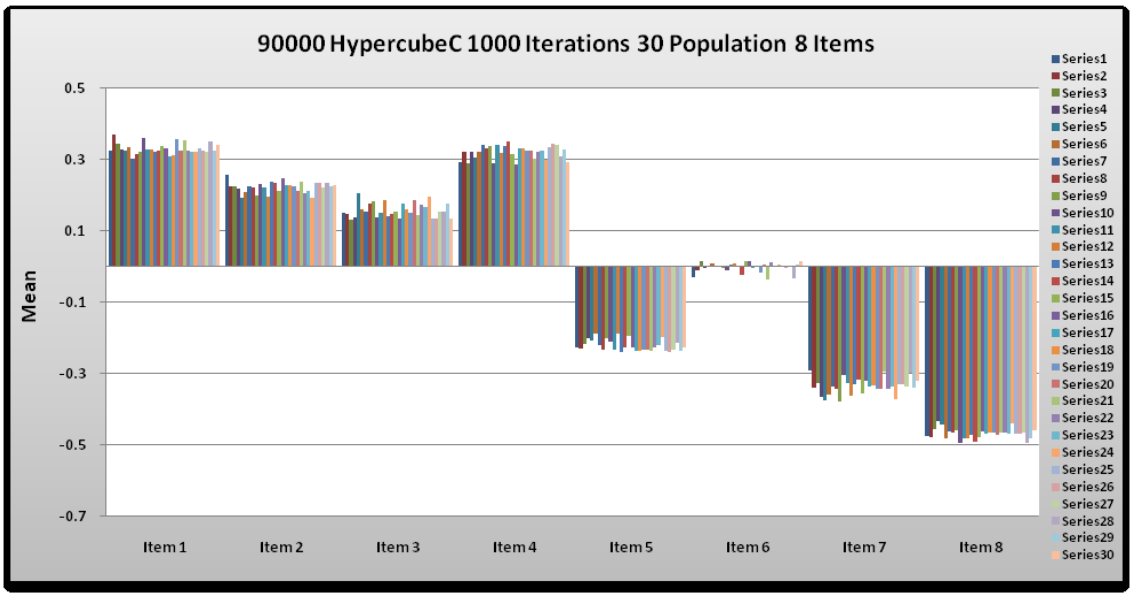


Figure 46 - 90000 Hypercube C 1000 Iterations 30 Population 8 Items

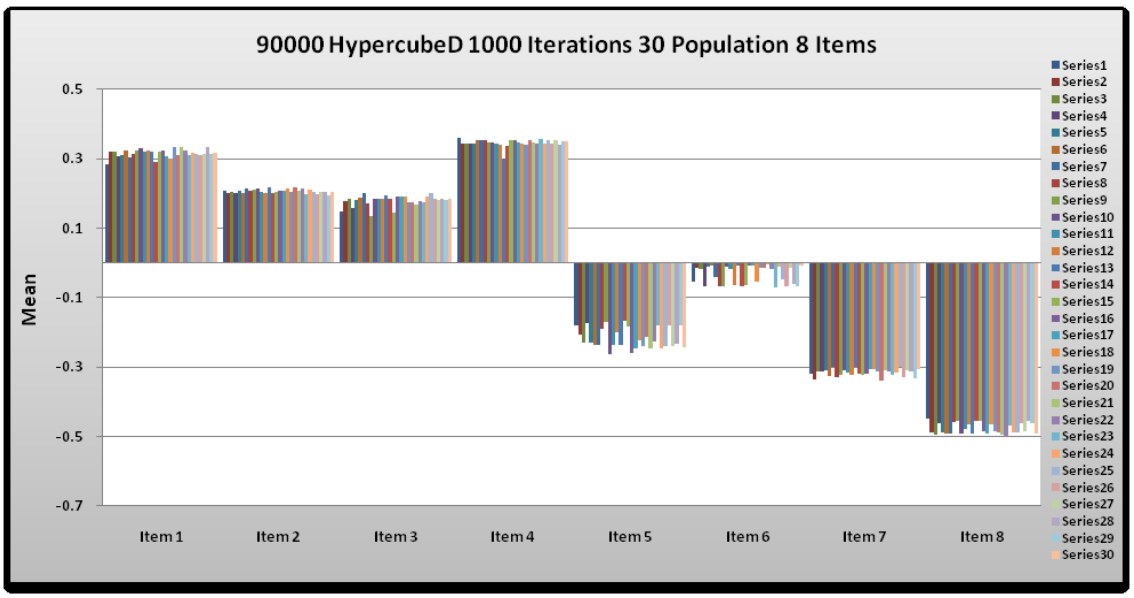


Figure 47 - 90000 Hypercube D 1000 Iterations 30 Population 8 Items

The above 4 bar graphs are the result of the outputs for the means for the four variations of a LHS with 90000 rows per item for 8 ranks. The graphs display all 30 population points for each item.

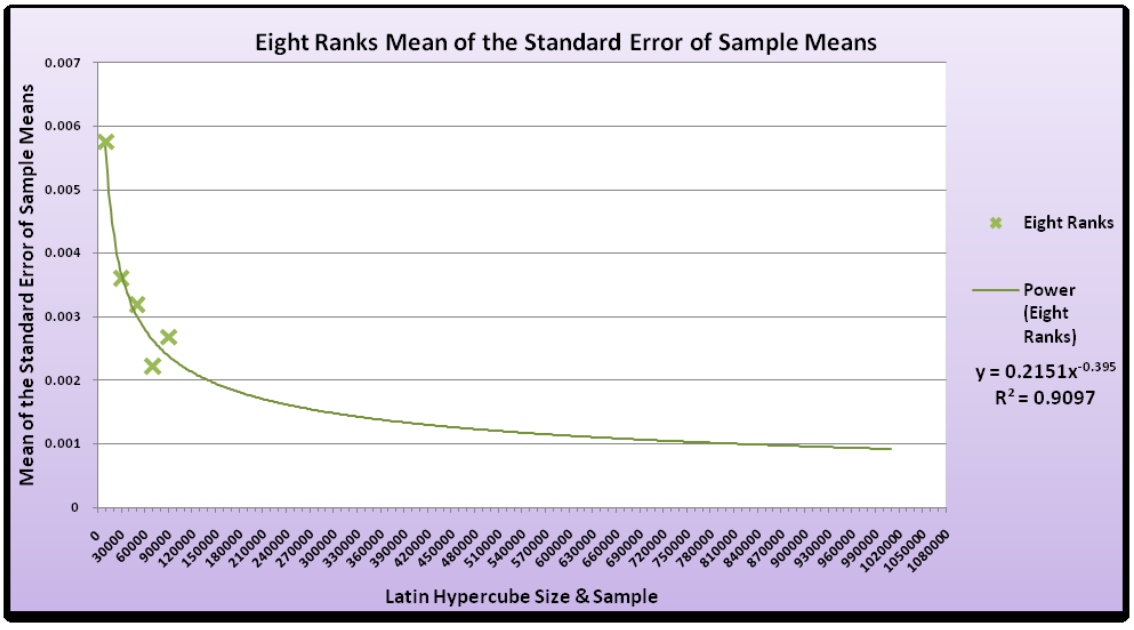


Figure 48 - Eight Ranks Mean of the Standard Error of Sample Means

The above graph is a duplication of Figure 12 but this time only reflecting the data surrounding 8 ranks.

4.6 Summary of the results

The representation of the data in this chapter followed a systematic and progressive design that allows for a logical argument regarding all the discussion points in the next chapter to appear in the same sequence. The mediums used are appropriate and will enable clear analysis from easy to understand graphical data representation, which effectively summarises a large volume of numeric data.

CHAPTER 5: DISCUSSION OF THE RESULTS

5.1 Introduction

In this chapter, the analysis of the presented data in the previous chapter will follow the same systematically laid out pattern of events that will lead to a sequentially improving set of conclusions. New variations of graphs will also be labelled using the same reference number as the original from which it was taken but including, in sequence, a letter of the alphabet to differentiate it.

Proposition One's analysis will take place under a single heading with a conclusion at the end of it, Proposition Two however, will be dealt with under sub headings for each different section of rankings with a final conclusion summarising the sections together.

The analysis begins with a validation for the final model design based on recorded run times using the available computer resources mentioned previously.

5.2 Simulation Duration

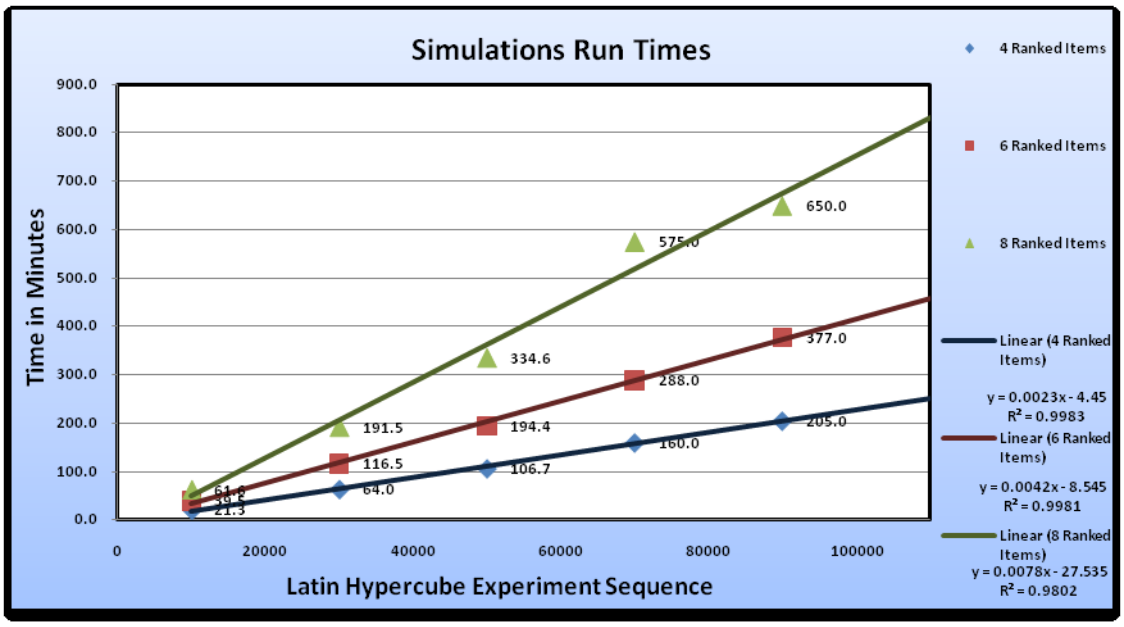


Figure 12A - Simulation Run Times with Linear Trend Lines

The reason behind the inclusion of this graph is to validate the decision to build the final model using LHS's with much less than 1 010 000 rows per item, which as previously stated would have been the ideal. Trend lines are now included in the original **Figure 12** with the following accompanying equations and R^2 values as calculated by Microsoft® Excel®:

- 4 Ranked items y = 0.0023x - 4.45 **R²** = 0.9983
- 6 Ranked items y = 0.0042x - 8.545 **R²** = 0.9981
- 8 Ranked items y = 0.0078x - 27.535 **R²** = 0.9802

These equations can be used to predict the likely run time of one complete set of simulations at the ideal 1 010 000 LHS size by substituting that value into the x, as can be seen by the high R^2 values they appear to be a good fit to the available data and are therefore good indicators for approximate values.

y = 0.0023(1 010 000) - 4.45

y = 2 318.55 minutes *(this figure is fairly close to the one confirmed in **Figure 10**, 2191 min)*

y = 0.0042(1 010 000) - 8.545

y = 4 233.45 minutes

y = 0.0078(1 010 000) - 27.535

y = 7 850.47 minutes

Total = 14 402.47 minutes

Taking the 4 variations of LHSs for each set of ranked items requires the above total to be multiplied by 4 to give the total time required to run the complete set of simulations at LHS size = 1 010 000 rows per item.

57 609.88 minutes which equates to approximately 40 days of computer run time.

The above finding confirms that a compromise was in fact necessary for research purposes as the accumulated time for all the incremental steps to get to the LHS size of 1 010 000 would indeed prove to be excessive.

**As stated before the above conclusion is based on the combination of hardware and software available.*

5.3 Discussion pertaining to changing the number of ranked items

The literature review suggests that as the number of ranked items increases the consistency of the Algorithm's outputs will decrease. As seen in **Figure 12** this is exactly what occurs particularly with eight ranked items however, between six and four ranked items there appears to be almost no difference.

Only allowing for 4 variations of the LHS is the suggested reason behind the lack of differentiation between four and six ranked items, as the graph, based on means could change as the sample size of variations is increased. In an attempt to try to rectify the reversal, a set of ratios as per **Table 10** with a summarised mean has been created to quantify the relationship sample size has between the numbers of ranked items.

	Ranked Items Ratios		
Latin Hypercube Size & Sample	Four Ranks	Six Ranks	Eight Ranks
Average	1	1.062	2.133

Table 10A - All Ranked Items Average Ratio of Comparison

The above table shows that the literature review was correct and that Proposition One is in fact true. The ratios show that for every LHS size, six ranked items SE($\overline{\mathbf{X}}$) will increase 1.062 times that of 4 ranked items and eight ranked items SE($\overline{\mathbf{X}}$) will increase 2.133 times as much as those of 4 ranked items. It is important to realise that these figures could change as stated before with an increase in sample size of variations for each LHS.

5.4 Discussion pertaining to changing the Simulated Sample size

5.4.1 *Four Ranked Items*

Figures 13 to 16 show individually the spread of SE($\overline{\mathbf{X}}$) per item, from this a trend is identified, an increase in sample size results in an increase in consistency (the size of each box reduces) as well as a decrease in the average SE($\overline{\mathbf{X}}$) (the line across the inside of each box). Unfortunately, each figure is not exactly the same as the other and so the above finding is not conclusive.

Figure 17 shows that the trend identified in the box plot graphs is true but also true is the inconclusive rate of decrease of the SE($\overline{\mathbf{X}}$).

Figures 18 to 21 show the actual results of the Algorithm output means at LHS size 90 000, LHSs A and D seem very consistent but B and C seem to become inconsistent raising concern over the appropriate distance between Items 1 and 2.

Figure 22 utilises the average value of the SE($\overline{\mathbf{X}}$)s for each set of four variations of the stepped LHSs to provide a clearer picture of the trend by also quantifying it into an equation.

$$\mathbf{y = 0.4253x^{-0.531}}$$

The inclusion of the one point at LHS size 1 010 000 taken from the initial model's simulation run assisted in confirming this as a reasonably accurate,

given the **R²** number of 0.9829, predictor of the unknown given either the LHS size ($\mathbf{x}$) or a required SE($\overline{\mathbf{X}}$) value ($\mathbf{y}$).

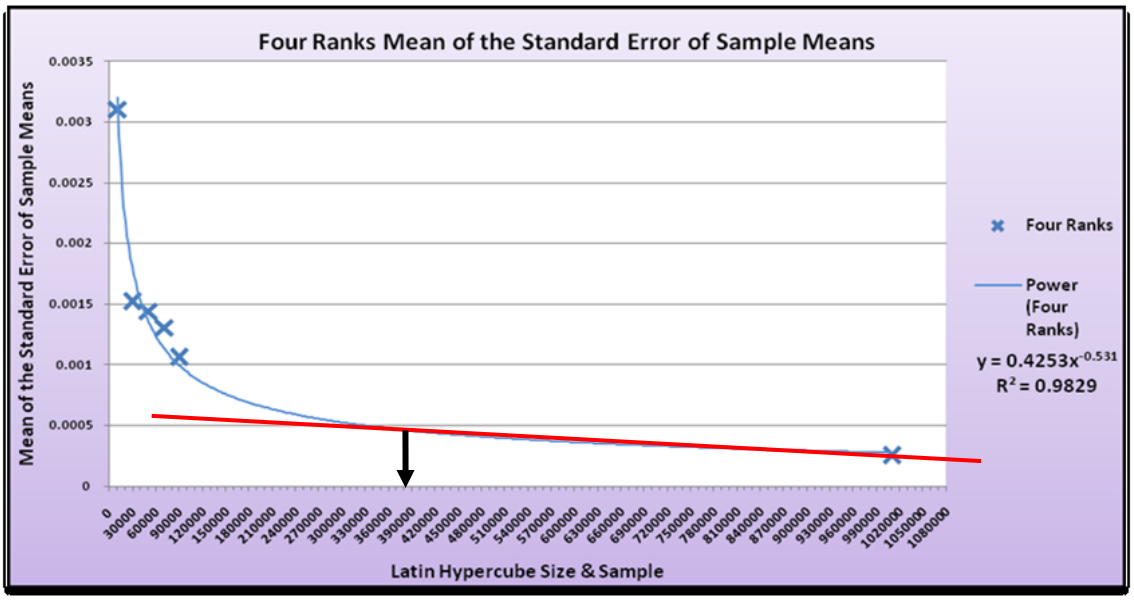


Figure 22A - Four Ranks Mean of the Standard Error of Sample Means

In the above variation of **Figure 22,** a line drawn at the base of the trend line with a downward arrow marking the point after which improvements in SE($\overline{\mathbf{X}}$) values occur in much larger steps of LHS sizes is the suggested best combination for both LHS and SE($\overline{\mathbf{X}}$). At this Point the LHS is roughly at 380 000 rows per item large which, according to the equation, makes the SE($\overline{\mathbf{X}}$) 0.000463.

The simulation run time for the above suggested input is approximately 14 ½ hours per simulation.

5.4.2 *Six Ranked Items*

Figures 25, 26 and 28 are very similar to those for four ranked Items however, **Figures 23, 24 and 27** are much different in that there consistency appears to be much wider and the averages more unsettled.

Figure 29 shows that a trend is difficult to predict as in the box plot graphs but it does appear that the SE($\overline{\mathbf{X}}$)s are constantly reducing, again also true is the inconclusive rate of decrease of the SE($\overline{\mathbf{X}}$).

Figures 30 to 33 show the actual results of the Algorithm output means at LHS size 90 000, all seem to have a much greater variation in them as those of four ranked items.

Figure 34 utilises the average value of the SE($\overline{\mathbf{X}}$)s for each set of four variations of the stepped LHSs to provide a clearer picture of the trend by also quantifying it into an equation.

$$\mathbf{y = 0.6726x^{-0.568}}$$

The shape of the trend line is assumed to be the same as the confirmed four ranked items trend line, given the **R²** number of 0.9103, it is a reasonably good predictor of the unknown given either the LHS size ($\mathbf{x}$) or a required SE($\overline{\mathbf{X}}$) value ($\mathbf{y}$).

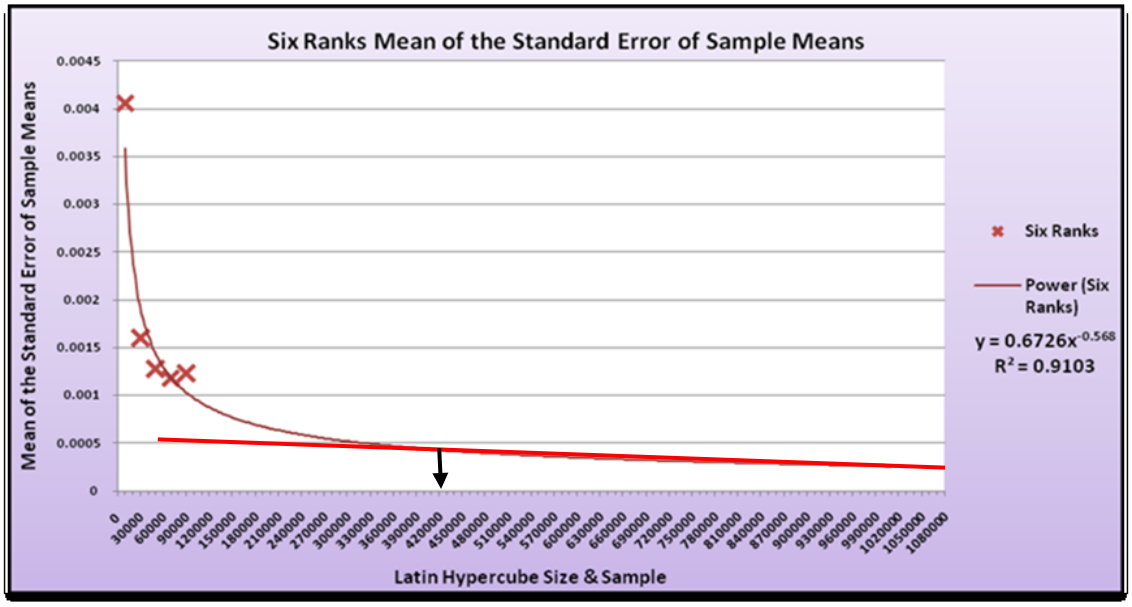


Figure 34A - Six Ranks Mean of the Standard Error of Sample Means

In the above variation of **Figure 34,** again the line drawn at the base of the trend line with a downward arrow marking the point after which improvements in SE($\overline{\mathbf{X}}$) values occur in much larger steps of LHS sizes is the suggested best combination for both LHS and SE($\overline{\mathbf{X}}$). At this Point the LHS is roughly at 420 000 rows per item large which, according to the equation, makes the SE($\overline{\mathbf{X}}$) 0.00043028.

The simulation run time for the above-suggested input is approximately 29.3 hours per simulation.

5.4.3 *Eight Ranked Items*

Figures 35 to 42 show a reducing trend in both consistency and the average SE($\overline{\mathbf{X}}$) value but they are inconclusive as to where that trend will lead.

Figure 43 shows that a trend is even more difficult to predict than that of six ranks as in the box plot graphs but it does appear to be reducing.

Figures 44 to 47 show the actual results of the Algorithm output means at LHS size 90 000, all seem to have a much greater variation in them as those of six ranked items and if looked closely in sample 25 of LHS D there appears to be a near reversal of ranking between items 3 and 2.

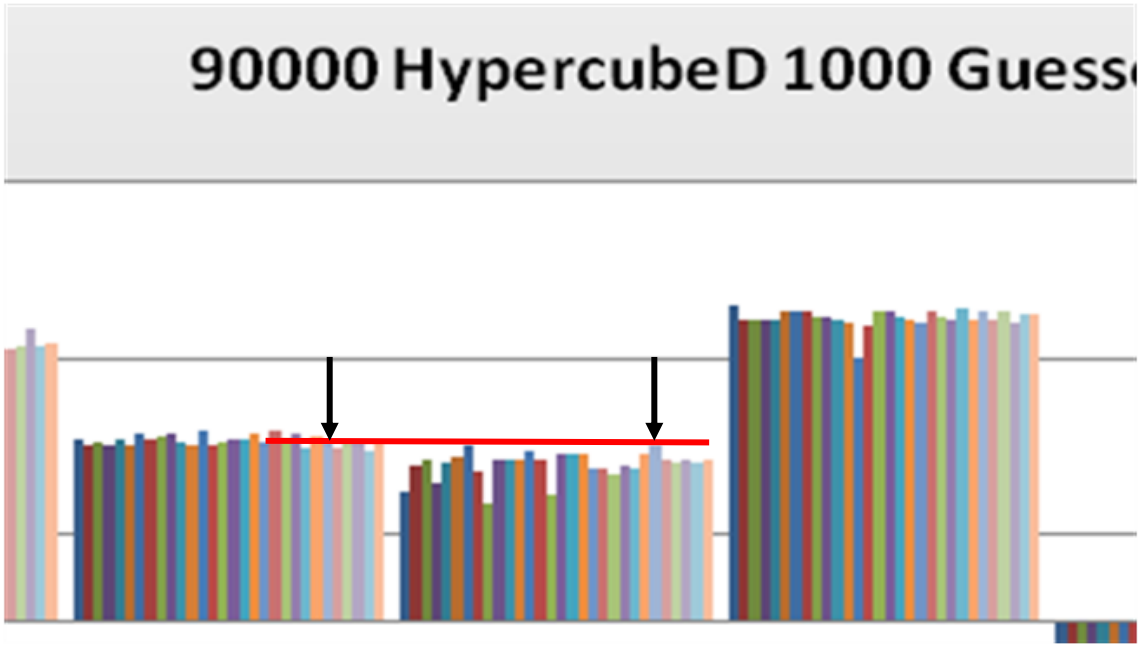


Figure 47B - 90000 Hypercube D 1000 Iterations 30 Population 8 Items

Figure 48 utilises the average value of the SE($\bar{\mathbf{X}}$)s for each set of four variations of the stepped LHSs to provide a clearer picture of the trend by also quantifying it into an equation.

$$\mathbf{y = 0.2151x^{-0.395}}$$

The shape of the trend line is assumed to be the same as the confirmed four ranked items trend line, given the **R²** number of 0.9097, it is a reasonably good predictor of the unknown given either the LHS size ($\mathbf{x}$) or a required SE($\overline{\mathbf{X}}$) value ($\mathbf{y}$).

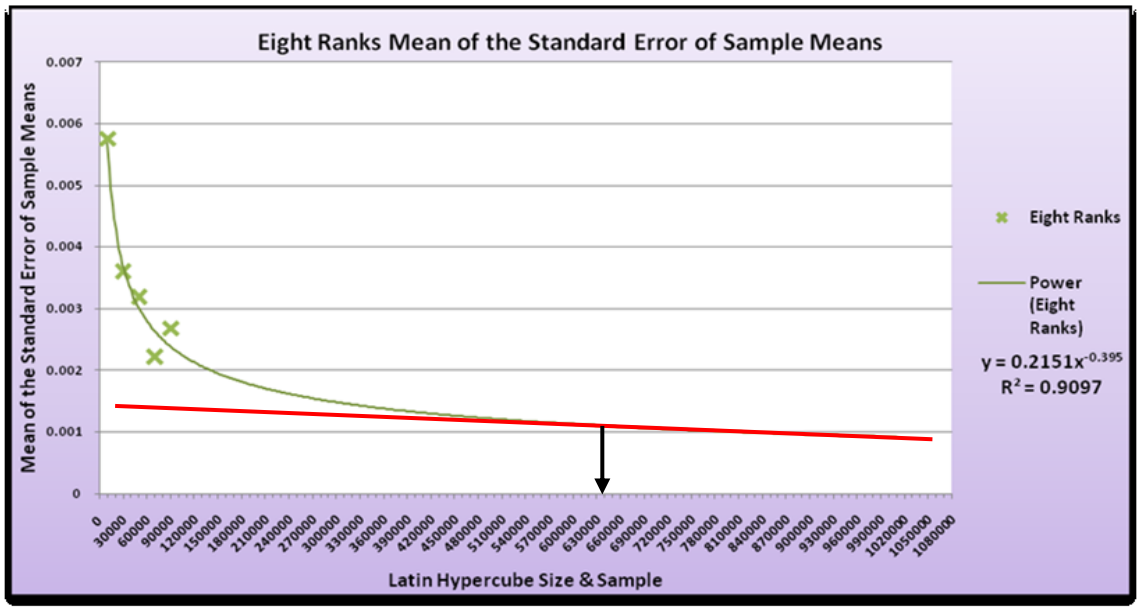


Figure 48A - Eight Ranks Mean of the Standard Error of Sample Means

In the above variation of **Figure 48,** again the line drawn at the base of the trend line with a downward arrow marks the point after which improvements in SE($\overline{\mathbf{X}}$) values occur in much larger steps of LHS sizes is the suggested best combination for both LHS and SE($\overline{\mathbf{X}}$). At this Point the LHS is roughly at 640 000 rows per item large which, according to the equation, makes the SE($\overline{\mathbf{X}}$) 0.0010945.

The simulation run time for the above suggested input is approximately 82.7 hours per simulation or 3 ½ days.

5.4.4 *Conclusion: Changing Simulated Sample size*

Proposition Two is proved to be true as can be seen by the evidence presented above, in all instances of items ranked 4,6 and 8 the SE($\overline{\mathbf{X}}$) reduces as the size of the LHS increases from which the simulated set of observations are derived. The reduction however, is initially quite rapid until a certain point is reached after which a large increase in LHS size creates a small decrease in the SE($\overline{\mathbf{X}}$).

5.5 Conclusion

Proposition One was proven to be true and a suggested average set of ratios was proposed to quantify the relationships between the increases of ranked items and the LHS size. The ratios show that for every LHS size, six ranked items SE($\overline{\mathbf{X}}$) will increase 1.062 times that of 4 ranked items and eight ranked items SE($\overline{\mathbf{X}}$) will increase 2.133 times as much as those of 4 ranked items. As it appears that increasing the sample size of LHS variations for each size could lead to different results, the research indicates that greater study into these relationships is required.

Proposition Two also proved true as each increase in LHS size, which leads to the creation of the simulated population sample, resulted in a decrease in the SE($\overline{\mathbf{X}}$) value. What was interesting was the shape that those reductions took for each quantity of ranked items. After a certain point, the decrease in SE($\overline{\mathbf{X}}$) values became a lot less in relation to the increases in LHS size. The summary of the suggestions below reflect the findings of this research for the best reasonable possible combinations of inputs as well as time (in hours) required to run the necessary simulations.

		LHS Size	SESM Value	Simulation Run Time
Ranked Items	4 Ranks	380 000	0.0005	14.5
	6 Ranks	420 000	0.0004	29.3
	8 Ranks	640 000	0.0011	82.7

Table 11 - Suggested Minimum Inputs

Although delimited, the research must still consider the application of the Algorithm relative to observed sample sets. Logically it would not make sense to apply a high degree of accuracy to an Observed sample set that itself lacks the same level of accuracy, this would only lead to a false sense of security in the results. Below is an amended version of Table 9 but this time it includes the sample size of an observed sample given the standardised Standard Deviation of 0.95 and the respective Standard Error from the Algorithm given a certain LHS/Simulated Sample size.

		Mean of the Standard Error of Sample Means			Sample Size (given standard deviation = 0.95)		
		Four Ranks	Six Ranks	Eight Ranks	Observed Data 4 Ranks	Observed Data 6 Ranks	Observed Data 8 Ranks
Latin Hypercube Size & Sample	**10000**	0.0031	0.0041	0.0058	**93810**	**54732**	**27260**
	30000	0.0015	0.0016	0.0036	**389085**	**352731**	**69230**
	50000	0.0014	0.0013	0.0032	**435702**	**552266**	**88360**
	70000	0.0013	0.0012	0.0022	**530069**	**647197**	**182708**
	90000	0.0011	0.0012	0.0027	**793606**	**590970**	**125565**
	380000	0.0005			**4204580**		
	420000		0.0004			**4874666**	
	620000			0.0011			**753383**

Table 9A – Suggested Latin Hypercube/Simulated Sample Size for Observed Sample size

This table shows that the methodological standard error that exists within the Algorithm is much less than the sample error that would occur from observing a sample of an actual population. For an observed sample to have the same

sample error as the Algorithm has at its simulated sample size of 10000 it would have to incorporate approximately 93810 observations for four ranked items.

CHAPTER 6: CONCLUSIONS AND RECOMMENDATIONS

6.1 Introduction

This chapter briefly summarises the entire outcome of the research report and makes suggestions for further research.

6.2 Conclusions of the study

The purpose of this research was to be the first step in setting up a set of best practises for all users of the Algorithm and in that vein below is a summary of the conclusions drawn.

- Input data should default to those tabulated in **Table 12** only when the observed sample sizes are greater than those in the Table 9A or when it is known that the entire population is the sample.
- If the means of two items are very close to each other in magnitude, additional running of the Algorithm a second and possibly third time with a re-sorting of the LHS in each one should take place to check whether they change.
- The use of settings other than those recommended in Table 12, because of the Observed sample's sample error, should utilise a similar method to that used in the final model i.e. multiple populations per LHS and multiple re-sorting of the LHS. This should alert the researcher to any possible variations that will occur. The LHS size should follow the

recommendations made in Table 9A relative to the observed sample size.

- Increasing the number of ranked items will lead to a greater level of inconsistency given the same LHS size used to calculate the simulated sample set.
- Greater computing resources such as a larger CPU could improve run times.

The Algorithm appears to be a reliable and, from experience, a relatively simple method to use in comparison to those cited in the literature review. The fact that it can easily be translated into as widely spread used program such as Microsoft Excel indicates the accessibility of the Algorithm coupled together with its effectiveness makes it a powerful tool to access deeper understanding of ranked ordered data for any interested party.

6.3 Recommendations

The purpose of this study was not to justify the use of the Algorithm in disciplines such as Marketing or Sociology but instead to offer guidance in its general use so that those disciplines can start the research to validate its use.

6.4 Suggestions for further research

6.4.1 *Re-sorting the LHS*

The differences between the results obtained between four ranked items and six ranked items indicates that the four sets of re-sorted LHS were not enough. In some instances, six ranked items gave better average SE($\overline{\mathbf{X}}$) values than

four ranks at the same LHS size. Therefore, it is suggested that future research should be directed at increasing the number of re-sorted LHS sets with the recommendation that greater computational power (perhaps a server) be utilised to do so to decrease the prohibitive simulation running times.

6.4.2 *LHS Characteristic*

In Figure 5, the Algorithm produced estimates that had very small SE($\overline{\mathbf{X}}$)s it can only be concluded that a certain characteristic of the LHS used caused this. It is suggested then as a possible area of research to discover what this characteristic is as it would assist in eliminating the need for larger simulated sample sizes to increase the degree of accuracy as well as reduce simulation-running times.

6.4.3 *Application*

The research looked at the Algorithm from a purely methodological viewpoint and therefore its applicability to real situations remains an un-investigated area of research. A particularly interesting area would be some sort of link with the research carried out by Mason, Dyer and Norton(2009) whereby they actually measured brain activity when people who had been exposed to social influence viewed symbols. These peoples brain activity indicated that preference based on social interaction can be noticeably measured and it would be interesting to see if it could be correlated with a physical rank ordering exercise.

6.4.4 *Multiple Observed population sets*

The research made use of a single set of contrived observed data with unknown parameters. Further research could include observed sample sets with known parameters as well as different kinds of controlled biases to see if the results differ from those in this study as well comparing the known parameters with the estimated ones.

6.4.5 *Iterations*

The impact of changing the number of iterations (which have an effect on convergence rate for the Algorithm) which the Algorithm performs was not within the scope of this research and is an area that should be considered for future research.

6.4.6 *Sampling Method*

This research made use of Latin Hypercube Sampling to generate the simulated population as the literature indicated that it would deliver the best results however, the impact of different methods on the effectiveness of the Algorithm remains un-researched.

REFERENCES

Abelson, R. P. and Tukey, J. W. (1963) Efficient Utilization of Non-Numerical Information in Quantitative Analysis General Theory and the Case of Simple Order, *The Annals of Mathematical Statistics,* 34(4), pp. 1347-1369.

Albright, S., Winston, W. and Zappe, C. (2008) *Data analysis and decision making with Microsoft Excel*, South-Western Pub.

Bell, J. (2005) *Doing your research project: a guide for first-time researchers in education, health and social science,* 4th ed., Open Univ Pr, Maidenhead.

Cheng, J. and Druzdzel, M. (2000) 'Latin Hypercube Sampling in Bayesian Networks', AAAI Press, pp. 287-292.

Cresswell, J. (2003) *Research Design, Qualitative, Quantitative, and Mixed Methods Approaches,* Second ed., Sage Publications, Thousand Oaks.

Diaconis, P. (1989) A Generalization of Spectral Analysis with Application to Ranked Data, *The Annals of Statistics,* 17(3), pp. 949-979.

Efron, B. and Tibshirani, R. (1986) Bootstrap Methods for Standard Errors, Confidence Intervals, and Other Measures of Statistical Accuracy, *Statistical Science,* 1(1), pp. 54-75.

Greatorex, M. (2007) "multidimensional scaling." The Blackwell Encyclopedia of Management. , Cooper, Cary L. Blackwell Publishing

Hensler, C. and Stipak, B. (1979) Estimating Interval Scale Values for Survey Item Response Categories, *American Journal of Political Science,* 23(3), pp. 627-649.

Hollingshead, A. B. (1996) The Rank-Order Effect in Group Decision Making, *Organizational Behavior and Human Decision Processes,* 68(3), pp. 181-193.

Kalof, L., Dan, A. and Dietz, T. (2008) *Essentials of social research*, Open University Press, Maidenshead.

Knapp, T. (1990) Treating ordinal scales as interval scales: an attempt to resolve the controversy, *Nursing Research,* 39(2), pp. 121-123.

Labovitz, S. (1970) The Assignment of Numbers to Rank Order Categories, *American Sociological Review,* 35(3), pp. 515-524.

Mason, M., Dyer, R. and Norton, M. (2009) Neural mechanisms of social influence, *Organizational Behavior and Human Decision Processes,* 110, pp. 152-159.

Mayer, L. S. (1971) A Note on Treating Ordinal Data as Interval Data, *American Sociological Review,* 36(3), pp. 519-520.

McKay, M. D., Beckman, R. J. and Conover, W. J. (1979) A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code, *Technometrics,* 21(2), pp. 239-245.

Siegel, S. (1957) Nonparametric Statistics, *The American Statistician,* 11(3), pp. 13-19.

Spearman, C. (1904) The Proof and Measurement of Association between Two Things, *The American Journal of Psychology,* 15(1), pp. 72-101.

Stacey, A. G. (2006) Estimating the means and standard deviations of rank ordered survey items, *Management Dynamics : Journal of the Southern African Institute for Management Scientists,* 15(3), pp. pp. 26-35.

Steinberg, D. M. and Hunter, W. G. (1984) Experimental Design: Review and Comment, *Technometrics,* 26(2), pp. 71-97.

Stevens, S. S. (1946) On the Theory of Scales of Measurement, *Science,* 103(2684), pp. 677-680.

Thurstone, L. (1927) A law of comparative judgment, *Psychological review,* 34(4), pp. 273–286.

Unknown (2009) *Survey Question and Answer Types,* (last accessed 2010/04/22, 2010), from http://www.questionpro.com/tutorial/2.html.

Velleman, P. F. and Leland, W. (1993) Nominal, Ordinal, Interval, and Ratio Typologies Are Misleading, *The American Statistician,* 47(1), pp. 65-72.

Willig, C. (2008) *Introducing qualitative research in psychology,* Second ed., Open University Press Maidenshead.

Yao, G. and Bockenholt, U. (1999) Bayesian estimation of Thurstonian ranking models based on the Gibbs sampler, *British Journal of Mathematical and Statistical Psychology,* 52(1), pp. 79-92.

Young, F. (1981) Quantitative analysis of qualitative data, *Psychometrika,* 46(4), pp. 357-388.

Yu (2000) Bayesian analysis of order - statistics models for ranking data, *Psychometrika,* 65(3), p. 281.

APPENDIX A

The Shotgun Stochastic Parameter Estimation Algorithm

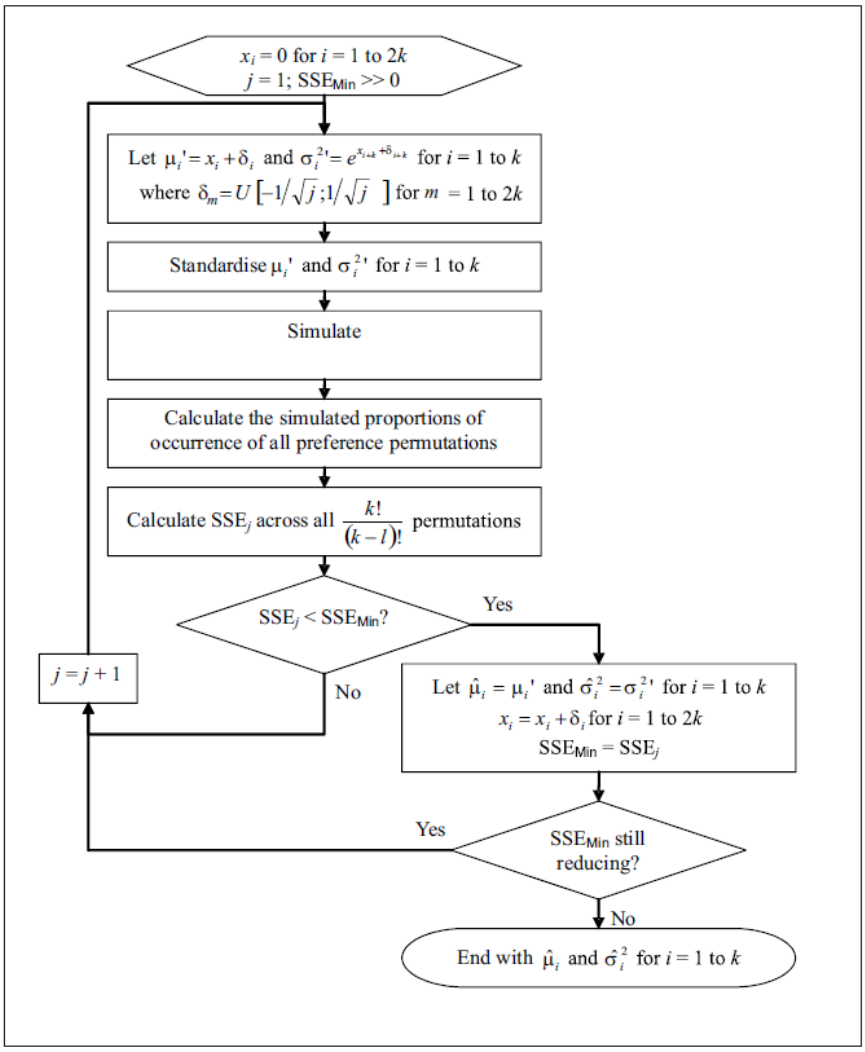


Figure 49 - The Shotgun Stochastic Parameter Estimation Algorithmic Flow

Stacey (2006: 29) explains the process as follows:

1. **Start with initial guesses for the item means, $\mu_i = 0$, and item variances, $\sigma_i^2 = 1$ for $i = 1$ to k.**

	A	B	C	D	E	F	G	H
1	**Rank ordered data**							
2	**Stochastic search**			Ctrl-Shift-C to Clear iterations				
3	Current test parameters			Ctrl-Shift-S to run the Search algorithm				
4	***SSE***		**Item 1**	**Item 2**	**Item 3**	**Item 4**	**Item 5**	**Item 6**
5	**2.03779**	=C13+(RAND()-0.5)/($A$9^0.5)			-0.2482626	-0.1181966	-0.4734414	-0.2100497
6								
7	Best case calculation							
8	***n***	***SSE***	**1**	**2**	**3**	**4**	**5**	**6**
9	1	**2.037789**	-0.1755336	0.2469316	-0.2482626	-0.1181966	-0.4734414	-0.2100497
10								
11	Best case yet							
12	***n***	***SSE***	**1**	**2**	**3**	**4**	**5**	**6**
13	0	**100.00000**	0	0	0	0	0	0
14								
15								
16								

On a Worksheet labelled “Stochastic Search” start by entering the equation “=C13+(RAND()-0.5)/($A$9^0.5)” into cell C5 and then copy it across to H5. Cell A9 contains the formula “=A13+1” and simply type the numbers in row 13 into the cells as per above.

	A	B	C	D	E	F	G
10		**Items**					
11		**Item 1**	**Item 2**	**Item 3**	**Item 4**	**Item 5**	**Item 6**
12	Mean	='Stochastic Search'!C5		-0.24826	-0.1182	-0.47344	-0.21005
13	Variance	1	1	1	1	1	1

On a new worksheet labelled “Search Calcs” type “=’Stochastic Search’!C5” into cell B12 and copy across to G12. Simply enter the number 1 into the cells B13 to G13.

2. **Standardise the means and variances such that the total variance is equal to 1; i.e. $\sigma^2_{Total} = \sigma^2_{\mu} + \mu_{\sigma^2} = 1$**

	A	B	C	D	E	F	G	H	I
10		**Items**							
11		**Item 1**	**Item 2**	**Item 3**	**Item 4**	**Item 5**	**Item 6**	Mean	Vaiance
12	Mean	-0.17553	0.24693	-0.24826	-0.1182	-0.47344	-0.21005	-0.16309	0.04601
13	Variance	1	1	1	1	1	1	1	
14								Overall	1.04601
15	Std Mean	=(B12-$H$12)/SQRT($I$14)	-0.08328	0.0439	-0.30345	-0.04591	-8.1E-18	0.04399	
16	Std Var	0.95601	0.95601	0.95601	0.95601	0.95601	0.95601	0.95601	
17	Std Stdev	0.97776	0.97776	0.97776	0.97776	0.97776	0.97776	Overall	1

	A	B	C	D	E	F	G	H	I
10		Items							
11		Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Mean	Vaiance
12	Mean	-0.17553	0.24693	-0.24826	-0.1182	-0.47344	-0.21005	-0.16309	0.04601
13	Variance	1	1	1	1	1	1	1	
14								Overall	1.04601
15	Std Mean	-0.01216	0.4009	-0.08328	0.0439	-0.30345	-0.04591	-8.1E-18	0.04399
16	Std Var	=B13/$I$14		0.95601	0.95601	0.95601	0.95601	0.95601	
17	Std Stdev	0.97776	0.97776	0.97776	0.97776	0.97776	0.97776	Overall	1

	A	B	C	D	E	F	G	H	I
10		Items							
11		Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Mean	Vaiance
12	Mean	-0.17553	0.24693	-0.24826	-0.1182	-0.47344	-0.21005	-0.16309	0.04601
13	Variance	1	1	1	1	1	1	1	
14								Overall	1.04601
15	Std Mean	-0.01216	0.4009	-0.08328	0.0439	-0.30345	-0.04591	-8.1E-18	0.04399
16	Std Var	0.95601	0.95601	0.95601	0.95601	0.95601	0.95601	0.95601	
17	Std Stdev	=SQRT(B16)		0.97776	0.97776	0.97776	0.97776	Overall	1

	A	B	C	D	E	F	G	H	I
10		Items							
11		Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Mean	Vaiance
12	Mean	-0.17553	0.24693	-0.24826	-0.1182	-0.47344	-0.21005	-0.16309	0.04601
13	Variance	1	1	1	1	1	1	1	
14								Overall	1.04601
15	Std Mean	-0.01216	0.4009	-0.08328	0.0439	-0.30345	-0.04591	-8.1E-18	0.04399
16	Std Var	0.95601	0.95601	0.95601	0.95601	0.95601	0.95601	0.95601	
17	Std Stdev	0.97776	0.97776	0.97776	0.97776	0.97776	0.97776	Overall	=I15+H16

In the work sheet, “Search Calcs” enter the formulas as per the above figures. (Copy formulas occurring in column B across to column G)

3. **Numerically simulate the population rank ordering responses by taking a very large simulated sample.**

	A	B	C	D	E	F	G	H	I
2	**Latin hypercube sample generation**								
3	Specifications								
4	Number of survey items			6					
5	Sample size			10000					
6	=NORMSINV(($A9-RAND())/$D$5)								
7		Items							
8		1	2	3	4	5	6		
9	1	0.592	-1.46	-0.72	-0.84	2.397	-0.31		=RAND()
10	2	0.373	0.828	1.792	-0.81	-0.9	0.573		0.74771
11	3	0.392	0.624	0.408	0.02	0.556	-0.97		0.24904
12	4	-0.96	1.43	0.605	2.095	0.546	0.174		0.97561

In a new work sheet, labelled “Latin Hypercube Sample Generati” enter sequential numbers starting at “1” from cell A9 down to A10008.

Then enter the formula “=NORMSINV(($a(-RAND())/$D$5)” into cell B9, copy that formula across to cell G9 and then copy B9 to G9 all the way down to row 10008.

Press the function key “F9” to recalculate the worksheet. Then copy from cell B9 to G10008 and paste special just the values into the same space.

Enter the formula “=RAND()” into cell H9 and copy down to H10008 the press “F9”.

Sort each column starting at B according to column H pressing”F9” after each sort and then exclude it from the next column sort. Delete column H after all the columns have been re-sorted.

	A	B	C	D	E	F	G
10				Items			
11		Item 1	Item 2	Item 3	Item 4	Item 5	Item 6
12	Mean	-0.20414	-0.39094	-0.01403	-0.32464	0.01854	-0.06343
13	Variance	1	1	1	1	1	1
14							
15	Std Mean	-0.04055	-0.22513	0.14731	-0.15962	0.17949	0.0985
16	Std Var	0.97643	0.97643	0.97643	0.97643	0.97643	0.97643
17	Std Stdev	0.98814	0.98814	0.98814	0.98814	0.98814	0.98814
18				Items			
19	Respondent	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6
20	1	=B$15+B$17*'Latin Hypercube Sample Generati'!B9					-0.21039
21	2	0.32833	0.59331	1.91836	-0.96101	-0.71113	0.66443
22	3	0.34665	0.39172	0.5508	-0.13944	0.72928	-0.85528

Enter the formula as per the above figure on work sheet “Search Calcs” and copy from column B to G and then 10000 rows down with corresponding numbering in column A.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
10					Items										
11		Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Mean	Vaiance						
12	Mean	-0.17553	0.24693	-0.24826	-0.1182	-0.47344	-0.21005	-0.16309	0.04601						
13	Variance	1	1	1	1	1	1	1							
14								Overall	1.04601						
15	Std Mean	-0.01216	0.4009	-0.08328	0.0439	-0.30345	-0.04591	-8.1E-18	0.04399						
16	Std Var	0.95601	0.95601	0.95601	0.95601	0.95601	0.95601	0.95601							
17	Std Stdev	0.97776	0.97776	0.97776	0.97776	0.97776	0.97776	Overall	1						
18					Items								Items		
19	Respondent	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6		Ranking	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6
20	1	0.56654	-1.02484	-0.79098	-0.77274	2.04005	-0.35155			=RANK(B20,$B20:$G20)		4	4	1	3
21	2	0.35283	1.21074	1.66916	-0.74906	-1.18471	0.51407			4	2	1	4	4	3

Then enter the formula as per the above figure and copy from column J to O and then 10000 rows down.

4. **Calculate the sum of squared errors, (SSE) between the observed proportions and the simulated population proportions across all preference permutations.**

In work sheet, “Search Calcs” the observed population set of percentages of instances where a rank was assigned to an item should be entered in cells R21 to W26.

	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
19	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6		Rank position	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6
20	2	4	4	4	1	3			Observed					
21	4	3	1	4	4	2		1	0.215	0.045	0.104	0.189	0.045	0.402
22	4	3	2	4	1	4		2	0.114	0.229	0.104	0.141	0.080	0.114
23	4	2	3	1	4	4		3	0.074	0.104	0.293	0.077	0.043	0.069
24	1	3	4	4	4	2		4	0.593	0.620	0.497	0.590	0.830	0.412
25	1	4	4	4	2	3		5						
26	3	4	2	1	4	4		6						
27	4	1	4	3	2	4								
28	4	3	4	4	1	2			Bootstrapped					
29	1	4	2	4	3	4		=COUNTIF(J:J,$Q29)/COUNT($A:$A)	.202	0.129	0.217	0.190		
30	1	4	2	4	3	4		2	0.165	0.134	0.188	0.149	0.188	0.176
31	4	4	3	4	2	1		3	0.165	0.161	0.180	0.151	0.172	0.173
32	4	4	1	3	4	2		4	0.522	0.591	0.430	0.572	0.424	0.462
33	4	4	3	1	4	2		5	0.000	0.000	0.000	0.000	0.000	0.000
34	1	4	4	2	3	4		6	0.000	0.000	0.000	0.000	0.000	0.000

Enter the above formula into cell R29 and copy across and down to cell W34. The same format for the above figures is recommended but important is the numbering next to each population set regarding rank position.

	P	Q	R	S	T	U	V	W	X
19		Rank position	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	
20			Observed						
21		1	0.215	0.045	0.104	0.189	0.045	0.402	
22		2	0.114	0.229	0.104	0.141	0.080	0.114	
23		3	0.074	0.104	0.293	0.077	0.043	0.069	
24		4	0.593	0.620	0.497	0.590	0.830	0.412	
25		5							
26		6							
27									
28			Bootstrapped						
29		1	0.149	0.114	0.202	0.129	0.217	0.190	
30		2	0.165	0.134	0.188	0.149	0.188	0.176	
31		3	0.165	0.161	0.180	0.151	0.172	0.173	
32		4	0.522	0.591	0.430	0.572	0.424	0.462	
33		5	0.000	0.000	0.000	0.000	0.000	0.000	
34		6	0.000	0.000	0.000	0.000	0.000	0.000	
35									
36			Errors						
37		=(R21-R29)^2/((R29+0.001)*(1-R29+			)060	0.032	0.173	0.291	
38			0.001))		)046	0.001	0.076	0.026	
39		3	0.059	0.024	0.086	0.042	0.116	0.075	
40		4	0.020	0.003	0.018	0.001	0.673	0.010	
41		5	0.000	0.000	0.000	0.000	0.000	0.000	
42		6	0.000	0.000	0.000	0.000	0.000	0.000	

Enter the above formula into cell R37 and copy across and down to cell W42.

	A	B	C	D	E	F	G	H
1	Partial rank ordered data							
2	Stochastic search			Ctrl-Shift-C to Clear iterations				
3	Current test parameters			Ctrl-Shift-S to run the Search algorithm				
4	*SSE*		Item 1	Item 2	Item 3	Item 4	Item 5	Item 6
5	=SUM(		-0.2041361	-0.3909355	-0.014028	-0.3246439	0.0185442	-0.0634268
6	'Search							
7	Calcs'!	alculation						
8	R37:	*SSE*	1	2	3	4	5	6
9	W42)	2.007282	-0.2041361	-0.3909355	-0.014028	-0.3246439	0.0185442	-0.0634268
10								
11	Best case yet							
12	*n*	*SSE*	1	2	3	4	5	6
13	0	100.00000	0	0	0	0	0	0
14								
15								
16								

In work sheet Stochastic Search" copy the above formula into cell A5.

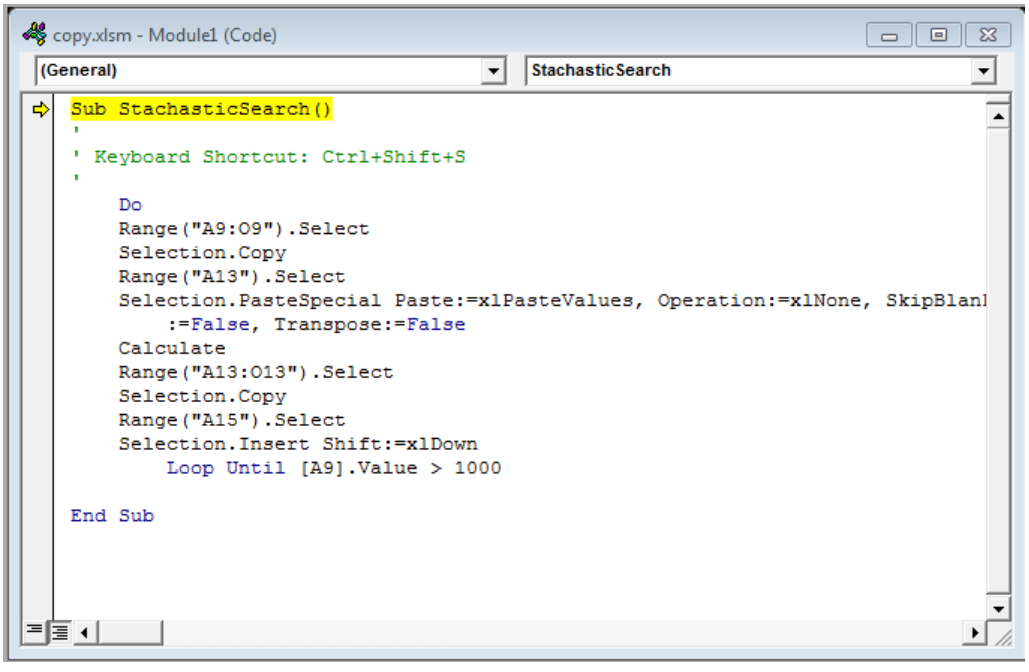
5. **Compare the SSE to that obtained using previous estimates of item means and variances. If the SSE is less than the previously**

recorded minimum SSE, then the current guesses become the best estimates of the items' means and variances.

	A	B	C	D	E	F	G	H
1	**Partial rank ordered data**							
2	**Stochastic search**			Ctrl-Shift-C to Clear iterations				
3	Current test parameters			Ctrl-Shift-S to run the Search algorithm				
4	***SSE***		**Item 1**	**Item 2**	**Item 3**	**Item 4**	**Item 5**	**Item 6**
5	**2.00728**		-0.2041361	-0.3909355	-0.014028	-0.3246439	0.0185442	-0.0634268
6								
7	Best case calculation							
8	***n***	***SSE***	**1**	**2**	**3**	**4**	**5**	**6**
9	1		=IF($A$5<$B$13,C5,C13)		-0.014028	-0.3246439	0.0185442	-0.0634268
10								
11	Best case yet							
12	***n***	***SSE***	**1**	**2**	**3**	**4**	**5**	**6**
13	0	**100.00000**	0	0	0	0	0	0
14								
15								
16								

Enter the above formula into cell C9 and copy across to H9

6. **Introducing uniformly distributed random perturbations to the guesses of the item means and variances, which decrease in magnitude with the square root of the number of iterations, carried out, repeat steps 2 to 6 until the SSE is no longer reducing with multiple successive iterations.**



Create the above macro and run it.

On a new work sheet, labelled “Final Results” enter the following formulas as directed using the same headings as below.

	A	B	C	D	E	F	G	H	I	J	K
1	**Rank ordered data - Results**										
2					**Item means**						
3	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Std Stdev	Raw Mean	Vaiance of means	Mean Variance	Overall Variance
4	-0.1755	0.2469	-0.2483	-0.1182	-0.4734	-0.2100		-0.1631	0.0460	1	1.0433
5	**-0.012**	**0.401**	**-0.083**	**0.044**	**-0.304**	**-0.046**	**0.979**	0.0000	0.0441	0.9585	1.0026

Cell A4 “=’Stochastic Search’!C9” copied across to cell F4

Cell H4 “=AVERAGE(A4:F4)”

Cell I4 “=VARP(A4:F4)”

Cell J4 “1”

Cell K4 “=I4+J4*(1-1/’Search Calcs’!C9)”

Cell A5 “=(A4-$H4)/SQRT|($K$4)” Copied across to cell F5

Cell G5 “=SQRT(J5)”

Cell H5 “=AVERAGE(A5:F5)”

Cell I5 “=VARP(A5:F5)”

Cell J5 “=1/$K$4”

Cell K5 “=I5+J5”

Cells A5 to F5 will give you the results of the means for the ranked items.