BAYESIAN RANKING OF CRICKET TEAMS USING PAIRED COMPARISON MODELS

30

By

Bilal Ahmed

Department of Statistics

Quaid-i-Azam University, Islamabad

2007

In the name of Allah, The compassionate, The merciful.

BAYESIAN RANKING OF CRICKET TEAMS USING PAIRED COMPARISON MODELS

By

Bilal Ahmed

A thesis submitted in the partial fulfillment of the requirements for the degree of

THE MASTER OF PHILOSOPHY

IN

STATISTICS

Department of Statistics
Quaid-i-Azam University, Islamabad
2007

Certificate

Bayesian Ranking of Cricket Teams Using Paired Comparison Models

By

Bilal Ahmed

A thesis submitted in the partial fulfillment of the requirements for the degree of

THE MASTER OF PHILOSOPHY

IN

STATISTICS

We accept this thesis as conforming to the required standard.

1. ______________
Prof. Dr. Muhmmad Aslam
(Chairman)

2. ______________
Prof. Dr. Muhmmad Aslam.
(Supervisor)

3. ______________
(External Examiner)

Date: 25-08-2007

Department of Statistics
Quaid-i-Azam University, Islamabad
2007

DECLARATION

I hereby declare that this thesis is the result of my individual research and that it has not been submitted concurrently to any other university for any other degree.

Bilal Ahmed

ACKNOWLEDGMENT

IN THE NAME OF ALLAH, THE MOST MERCIFUL AND BENEFICIENT

Praise is to ALMIGHTY ALLAH, creator of heavens and earth and Lord of lords, who enable and give me courage to complete this Research. All my respect goes to HOLY PROPHET MUHAMMAD (PBUH) who emphasized the significance of knowledge and research.

First of all, I would like to pay tributes and thanks to my parents whose constant, unending and countless prayers buoyed me up and moral support from my brothers and sisters.

I am deeply indebted to Professor Dr. Muhammad Aslam, Chairman, Department of Statistics Quaid-e-Azam University, my supervisor, for providing all the facilities for this research work. I thanks to entire faculty member for their valuable advices and assistance in completing my thesis.

I also pay tributes to my seniors, especially Ms Saima Altaf and Mr Nasir Abbas for their guidance.

BILAL AHMED

DEDICATED TO

My Parents and Teachers especially **Prof. Dr. Muhammad Aslam** *and my special friends whose prayers have always been a source of great inspiration for me and whose continued hope in me has led me to where I stand today.*

Chapter No#4

RANKING OF TOP SEVEN CRICKET TEAMS

Chapter 1

INTRODUCTION

Statistics is the science of collecting, organizing and interpreting numerical facts, which we call data. Data bombards us in everyday life. Each month, for example, government statistical offices release the latest numerical information on unemployment and inflation. Economist and financial advisors as well as policy makers in government and business study these data to make informed decisions. Doctors must understand the origin and trustworthiness of the data appear in medical journals if they are offer their patients the most effective treatment. Politicians rely on data from polls of public opinion. Market research data that reveal consumer tastes influence business decisions. Farmers study data from field trials of new crop varieties. Engineers gather data on the quality and reliability of manufactured products. Most areas of academic study make use of numbers, and therefore also make use of the method of statistics. Statistical methods not only describe important features of the data but also allow us to proceed beyond the collected data into the area of decision-making through generalization and predictions. Statistics assists in a sound and effective planning in any field of inquiry. Statistics teaches us how to gather, organize and analyze data, and then to infer the underlying reality from these data. Statistical tools are frequently applied for research purpose in almost all areas of study. Statistical techniques being powerful tools for analyzing numerical data; are used in almost every branch of learning. Biological and Physical Sciences, Genetics, Agronomy, Astronomy, Physics, Geology, etc. are the main areas where statistical techniques have been developed and are increasingly used.

The application of statistical tools in research work can be viewed in any research journal of any subject. If there is data, there must be statistical tools to analyze and interpret this data for the significant use.

H.G Wells anticipated that statistical thinking (numerical literacy) would one day be as necessary for efficient citizenship as the ability to read and write.

In this study, we are mainly concern with the Bayesian Statistics, the modern branch of Statistics. Prior to the modern branch of the statistics, the former is called Classical Statistics. So we may say that there are two schools of thoughts, the former is called Classical statistics and the latter is called Bayesian statistics. The approach to statistics that formally seeks to utilize prior information is called Bayesian statistics, names after Bayes (1763). In Bayesian statistics, we reflect on prior information, arising from sources other than the statistical investigation. Prior information about parameter generally comes from the past experience about similar situations involving similar parameter. Bayesian analyses are needed to solve real decision problem

The role of Bayesian statistics in the updating of beliefs about observables in the light of new information is of importance. This development, in combination with an operational approach to the basic concepts, has led us to view the problem of statistical modeling as that of identifying or selecting particular forms of representation of beliefs about observables.

In this study, we have presented an analysis for paired comparison data through Bayesian approach. In the method of paired comparisons, objects are presented in pairs to one or more judges. Sometimes it may be difficult to a panelist to rank or compare more than two objects or treatments at the same time especially when differences between objects are trivial. For this reason paired comparisons data is sometime regarded as more reliable. This method is broadly used in industry for assessing customer preference and designing products using trained panelists. Suppose we have 't' treatments or objects $T_1, T_2 ... T_t$. There are $\frac{t(t-1)}{2}$ different comparisons are possible with a single judge. For t objects and n judges the number of paired comparisons will be $n\binom{t}{2}$.

This method was indeed introduced in embryonic form by Fechner (1860) and after considerable extensions, made popular by Thurstone (1927). Several extensive reviews are available in the research paper of Bock and Jones (1968), Coombs (1964), Kendall (1940), and Torgerson (1958). Davidson and Farquhar (1976) have presented an extensive bibliography on paired comparisons. David's monograph (1988) has a detailed survey of the literature and references concerning the method and models of paired comparisons. The detailed discussion on paired comparison is focused in chapter-3.

The Bayesian Ranking of the top Seven Cricket teams by using the Rao-Kupper and Davidson models for paired comparison is presented in this study.

Chapter#2, In this chapter introduction to Bayesian statistics is presented briefly. Its advantages and greater scope over classical statistics is revealed. Basic terminologies of Bayesian statistics, prior distribution, formation of posterior distribution, Bayesian hypotheses testing along with its advantages over classical hypotheses testing are enlightened. As Bayesian Statisticians need computational tools to calculate a variety of summaries from posterior distributions that are mathematically complex, so Gibbs sampling for complex and multidimensional integration is also clarified.

Chapter#3, detail discussion about paired comparison methods is presented. Its application in various fields is demonstrated. Formulation of linear model and modifications of basic model into other paired comparisons models are detailed. A brief review of the existing literature on the methods and models of paired comparisons is also included in this chapter.

Chapter#4, comprises the Rao-Kupper and Davidson models with the notations and the likelihood for the parameters of the model. The ranking of Top Seven Cricket Teams (Australia, South Africa, Pakistan, India, New Zealand, Sri Lanka and England) using Posterior Mode and Posterior Mean of the parameters of these models is presented.

The predictive probabilities that one team would be better to another team in a future single comparison are included in this chapter. We use the technique of Gibbs sampling for obtaining the posterior means. We also determine the preference probabilities for the teams to verify the ranking of the teams. Preference probabilities are calculated by just calculating the probability of preferring team T_i to team T_j.The posterior probabilities of the hypotheses for the comparison of two parameters are calculated. Appropriateness of the model is tested using chi-square goodness of fit test. Posterior (Marginal) densities for the parameters of the Rao-Kupper model are sketched. The Comparison of ICC Ranking and Bayesian Ranking also presented here.

Chapter Appendix: In appendix, a set of programs designed in SAS package and C^{++} Language for the numerical solution is scheduled.

Chapter 2

2.1 Introduction

The section-wise scheme of this chapter is given below:

In sections 2 and 3, the Bayesian statistics with its ideas and features has been discussed while Section 4 deals with the Classical or Frequentist statistics. In sections 5 and 6, the Bayes' theorem and posterior distribution, the concept of conditional probability and the likelihood function are described. In section 7, subjective determination of prior density through different approaches and prior distribution is illustrated. The Informative prior has been explained in section 8. Sections 9 and 10 cover the concept of the Non-informative prior and the Uniform prior. The Prior Predictive and Posterior Predictive Distributions are defined in sections 11 and 12. Bayesian hypothesis testing is described in section 13. Section 14 covers Gibbs sampling with its mechanism and benefits. The last section 15 holds the historical overview and the scope of Bayesian statistics.

2.2 Bayesian Statistics

The particular advantages offered by Bayesian Statistics make it very useful in situations where uncertainty is unavoidable – Bayesian methods provide a mechanism to model the uncertainty. Such methods can also be used where normal optimization and decision-making techniques are difficult to apply. Sometimes we want to formal quantitative coherence in the context of decision making in situations of uncertainty. In these situations, the role of Bayes theorem in the updating of beliefs about observables in the light of new information is of importance. This development, in combination with an operational approach to the basic concepts, has led us to view the problem of statistical modeling as that of identifying or selecting particular forms of representation of beliefs about observables.

This approach makes use of not only the sample information but also the prior information. This is named so after Bayes (1763) who first introduced this approach. The Bolstad (2004) discusses the key features of this approach as:

- The unknown parameter is considered as a random variable.
- Prior information is the probability statement about parameter, which is interpreted as "degree of belief" or the relative weights that expert, gives to every possible value of the parameter. It measures how "plausible" the expert considers each parameter value before observing the data.
- We revise our beliefs about parameters using Bayes' theorem after getting the data. This is posterior distribution, which gives the relative weights to each parameter value after analyzing the data. The posterior distribution comprises the prior distribution and the observed data.
- Parameter estimates, along with confidence intervals (known as credibility intervals) or higher density region (HDR), are calculated directly from the posterior distribution. Credibility intervals are legitimate probability statements about the unknown parameters, since these parameters now are considered random, not fixed.
- It makes a great deal of practical sense to use all the information available, old and/or new, objective or subjective, when making decisions under uncertainty. This is especially true when the consequences of the decisions can have a significant impact, financial or otherwise.

- The rules of probability are used directly to make inferences about the parameters.

2.3 Features of Bayesian Statistics

Berger (1985) and Bolstad (2004) discuss the following features of Bayesian statistics:

- Classical approach cannot take into account the given prior information. So, in such situation where prior information is available, Bayesian viewpoint is more appropriate for statistical analysis.
- In Statistical analysis, the Bayesian approach is the only way which consistently uses probability to directly address uncertainty as discussed by Jeffreys (1961), Edwards et. al. (1963) and deFinetti (1972, 1974, 1975). On the other hand, Classical approach addresses the probability (which is the language of uncertainty), indirectly related to the probability of the hypothesis.
- Berger and Wolpert (1984) argue that conditional analysis of the observed data, as opposed to Frequentist averaging over is potential data, supporting to use of the Bayesian analysis.
- In decision theory it is natural to consider the only admissible decision rule which reduce consideration to the class of Statistical procedures. In a simple versus simple hypothesis situation, it has been repeatedly shown that the class of acceptable (Classical) decision rules corresponds to the class of Bayes decision rules.

- In Bayesian analysis, realistic model can more easily be chosen for analysis since there is less need to have models which allow special classical calculations following Rubin (1984).
- Robustness (violation of usual assumption) can be dealt with more easily by using Bayesian analysis as compared to Classical analysis.
- Berger and Wolpert (1984) define an important advantage of Bayesian analysis that various kinds of censoring of data cause no essential problem, on the other hand serious problems can be seen for Classical analysis.
- Bayesian analysis gives a final distribution for the unknown parameter (which is called posterior distribution) and from this a large number of questions can be answered simultaneously.
- Bayesian procedures almost always equal to the Classical sample procedures when sample size is very large and are relatively improved for moderated small sample sizes cases.
- Bayesian Statistics has a general way of dealing with a nuisance parameter. A nuisance parameter is one, which we don't want to make inference about, but we don't want them to interfere with the inferences we are making about the main parameters. Frequentist statistics does not have a general procedure for dealing with them.
- Bayesian Statistics is predictive, unlike conventional Frequentist Statistics. This means that we can easily find the conditional probability distribution of the next observation given the sample data.

- Bayesian Statistics has a single tool, Bayes theorem, which is used in all situations. This contrasts to Frequentist procedures, which require many different tools.
- Bayesian methods often outperform Frequentist methods, even when judged by Frequentist criteria.
- Bayesian approach can compare multiple hypotheses simultaneously unlike Classical approach.

2.4 Classical or Frequentist Statistics

It is the most commonly used Statistical approach. According to Bolstad (2004) it is based on the following ideas:

- The use of sample information only in making inferences about the population parameter.
- It considers the unknown parameters as fixed and its estimate as a random variable.
- Here Statistical procedures are judged by how well they perform in the long run over an infinite number of hypothetical repetitions of the experiment.
- A confidence interval for an unknown parameter is really a frequency statement about the likelihood that numbers calculated from a sample capture the true parameter. Strictly speaking, one cannot make probability statements about the true parameter since it is fixed, not random. Mostly, these inferences are made without regards to the use to which they are to be put.
- Probabilities are always interpreted as long run relative frequency.

2.5 Bayes' Theorem and Posterior Distribution

An English clergyman Thomas Bayes (1702-1761) derives a rule based on conditionally reduced sample space and first use in a paper that was published posthumously in 1763. After this, this Rule is known as Bayes' theorem. The typical phrasing of Bayes's theorem is in terms of disjoint events $A_1, A_2, ..., A_n$, whose union has probability one. Prior probabilities $P(A_i)$, for the events, are assumed known. An event B occurs, for which $P(B|A_i)$ (the conditional probability of B given that A_i) is known for each A_i. Bayes theorem then states that

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{\sum_{j=1}^{n} P(B|A_i)P(A_i)} \tag{2.1}$$

It should be known that the original probabilities are known as the *a priori* probabilities and the conditional probabilities $P(A_i|B)$ are called the *a posteriori* probabilities, so probabilities are revised after some additional information has been obtained. It is also called the formula for *probabilities of hypotheses* on account of the reason that the events $A_1, A_2, ..., A_n$ may be thought of as hypotheses to account for occurrence of the event B.

Now expressing the posterior density for unknown parameter θ, given $\mathbf{x}$ = ($X_1, X_2, ..., X_n$), the observed data, in term of the parametric model for $\mathbf{x}$ given θ, and the prior density for θ, the equation (2.1) may be written as:

$$p(\theta|\mathrm{x}) = \frac{p(\theta)p(\mathrm{x}|\theta)}{\int_{\theta} p(\theta)p(\mathrm{x}|\theta)d\theta} \quad \text{, for continuous case} \tag{2.2}$$

Here $p(\theta)$ is a prior distribution for θ; $p(\mathbf{x}|\theta)$ is a likelihood function of the observed data. Furthermore $\int_{\theta} p(\theta)p(\mathbf{x}|\theta)d\theta = p(\mathbf{x})$ i.e. the marginal distribution of $\mathbf{x}$. It can be observed from (2.2) that the posterior distribution $p(\theta|\mathbf{x})$ can be written as

Posterior distribution $\propto$ (Prior distribution) (Likelihood function),

or

$$p(\theta|\mathbf{x}) \propto p(\theta)p(\mathbf{x}|\theta) \tag{2.3}$$

So the posterior distribution consists of two sources of information: the prior information through the prior distribution and the information via the likelihood function. We can say that the posterior distribution 'updates' the information. All inferences and decisions about parameter are made through the posterior distribution.

In equation (2.2) and (2.3), $p(\theta|\mathbf{x})$ is known as posterior distribution which is proportional to the product of prior information and likelihood function. It is noted that prior information are acquired before the data is observed.

There are situation in real life when important prior information is available then it is a wise decision to take into account of this available information in order to make reliable estimates. The posterior distribution is the combination of prior information and the sample information, so it reflects the updated beliefs about θ after observing the sample $\mathbf{x}$. In other words, the posterior distribution combines the prior beliefs about θ with the information about θ contained in the sample $\mathbf{x}$, to give a composite picture of the final beliefs about θ.

2.6 Likelihood Function

The term Likelihood may be defined as for a random sample $X_1, X_2, ..., X_n$ following a probability density function $f(\mathbf{x}, \theta)$, the joint probability density function or the product of independent probability density functions is called the likelihood function, which is a function of θ

Symbolically, $$\mathrm{L}(\mathbf{x}, \theta) = f(x_1|\theta)\, f(x_2|\theta)\, \ldots\, f(x_n|\theta) \quad (2.4)$$

Where $\mathbf{x} = (X_1, X_2, ..., X_n)$

The likelihood function is used to find the set of parameter values that gives the highest possible likelihood, but in Bayesian statistics, the purpose is to obtain a complete probability distribution over all possible parameters values. The likelihood principal makes explicit the natural conditional idea that only the actual observed x should be relevant to conclusions or evidence about θ.

The likelihood is developed as a separate principle by Bernard (1949), and become a focus of interest when Birnbaum (1962) shows that it followed from the widely accepted sufficiency and principle. The arguments of Birnbaum for the likelihood principle is a proof its equivalence with other almost universally accepted natural principles. Stein (1962) also presents some remark on likelihood principle. Edwards (1974) presents the history of likelihood. Berger and Wolpert (1984) provide an extensive discussion on the likelihood principle.

2.7 The Subjective Probability the Prior Distribution

In equation (2.2), $p(\theta)$ is called prior distribution. Prior distribution is a convenient way to quantify prior information in term of probability distribution on θ.

According to Berger (1985), the main idea of subjective probability is to let the probability of an event reflects the personal belief in the "chance" of the occurrence of the event. For example, one may have a personal feeling as to the chance that θ will be between 3% and 4% in some particular situation, even though no frequency probability can be assigned to the event. The simplest way of determining subjective probabilities is to compare events, determining relative likelihoods. For example, if it is desired to find P(A). Simply compare A with A^c (the complement of A). If A is felt to be twice as likely to occur as A^c, then clearly $P(A)$ will be $\frac{2}{3}$ according to sample space. French (1980) describes in detail updating of belief in the light of someone else's opinion.

Subjective determination of the prior density can be made under any of the following approaches.

I. Histogram Approach

II. The Relative Likelihood Approach

III. Matching a Given Functional Form

IV. CDF Determination

Among the four approaches described above for subjectively determining a prior distribution, approaches I and II are most in use.[For Detail see Berger (1985)]. To gain the full benefits of the Bayesian approach, prior information should not be ignored. Formulating a prior distribution required expert judgment, which is known as elicitation.

Lindley (1961) and Blum (1967) discuss the importance of prior information in statistical inference and decision-making. Dalal and Hall (1983) present approximating priors by mixtures of natural conjugate priors. Dickey (1980) presents a theory of stochastic assessment of subjective probabilities.

One of the main differences between Classical Statistics and Bayesian Statistics is that the latter can utilize prior information in a formal way. This information can be quantified in terms of a probability distribution, which is known as the prior distribution, and this represents the knowledge about the parameter prior to observing the data. If there is no relevant prior information available then there are ways to drive a non-informative prior distribution. The prior distribution further depends on parameters, which are called hyper parameters.

2.8 The Informative Prior

An informative prior expresses specific, definite information about a variable parameter. An example is a prior distribution for the temperature at noon tomorrow. A reasonable approach is to make the prior a normal distribution with expected value equal to today's noontime temperature, with variance equal to the day-to-day variance of atmospheric temperature. This example has a property in common with many priors, namely, that the posterior from one problem (today's temperature) becomes the prior for the other problem (tomorrow's temperature); pre-existing evidence which has already been taken into account is part of the prior and as more evidence accumulates the prior is determined largely by the evidence rather than any original assumption, provided that the original assumption admitted the possibility of what the evidence is suggesting. The term prior and posterior is generally relative to specific datum or observation.

2.9 The Non - Informative Prior

One of the advantages of Bayesian approach is that it can be applied even when no prior information is available. What is needed in such situations is a *noninformative prior,* by which is meant a prior which contains no information about θ. For example, in throwing a die, the prior who gives probability 1/6 to each outcome is clearly noninformative.

Berger (1985) explains that due to the compelling reasons to perform a conditional analysis and the attractiveness of using Bayesian analysis to do so, there have been attempts to use the Bayesian approach even when no (or minimal) prior information is available. In such situation we require only a non-informative prior, by which mean a prior which contains no information about parameter θ. For example, in testing between two hypotheses, the prior who gives probability $\frac{1}{2}$ to each of the hypotheses is clearly non-informative. Box and Tiao (1973) define a non-informative prior as a prior, which provides little information relative to the experiment. Bernardo and Smith (1994) use a similar definition; they say that non-informative priors have minimal effect relative to the data, on the final inference. They regard the non-informative prior as a mathematical tool; it is not a uniquely non-informative prior.

2.10 The Uniform Prior

The uniform density used by Bayes (1763) and Laplace (1812), and supported by Geisser (1984). A Uniform distribution is used as a noninformative prior. Uniform priors are particularly easy to specify in the case of a parameter with bounded support. The simplest situation to consider is when θ is a finite, consisting of say n observation. The obvious prior is to give each θ probability $1/n$. It is noted that Uniform prior have been applied to many problems and mostly results are entirely satisfactory.

A uniform prior for the probability parameter (say λ) in a Bernoulli, binomial, or negative binomial model can be specified by: $p(\lambda) = 1,\ 0 \le \lambda \le 1$, or if there is some reason to specify a non-normalized uniform: $p(\lambda)=1,\ 0\le\lambda\le \mathrm{k}$. The second is non-normalized because it does not integrate to one although it provides no problem whatsoever in Bayesian analysis. Both of these forms are referred to as proper since they integrate to a finite quantity. Proper uniform priors can be specified for parameters defined over unbounded space if we are willing to impose prior restrictions.

It is also possible to specify improper uniform priors that do not possess bounded integrals; these can results in fully proper posteriors under some circumstances. Consider the common case of a noninformative uniform prior for the mean of a normal distribution.

2.11 The Prior Predictive Distribution

If a random variable X has probability density $f(x|\lambda)$ and λ has probability density $p(\lambda)$, then the joint density of X and λ is defined as:

$$h(x,\lambda)=f(x|\lambda)\,p(\lambda). \tag{2.5}$$

The marginal $p(x)$ density of X is defined as:

$$p(x)=\begin{cases}\int_{\Theta} f(x|\lambda)\,p(\lambda)\,d\lambda; & \text{for continuous case}\\ \sum_{\Theta} f(x|\lambda)\,p(\lambda); & \text{for discrete case}\end{cases} \tag{2.6}$$

which is known as prior predictive distribution of X .

2.12 The Posterior Predictive Distribution

The joint posterior density for a random variable $Y=x_{n+1}$ and parameter λ given data $(x=x_1,x_2,\ldots,x_n)$ is defined as:

$$h(y,\lambda|x)=f(y|\lambda)\,p(\lambda|x) \tag{2.7}$$

Also we assume that $\boldsymbol{x}$ and y are independent. The marginal density $p(y|x)$ of $Y = x_{n+1}$ given data $(\boldsymbol{x} = x_1, x_2, \ldots, x_n)$ is defined as:

$$p(y|\boldsymbol{x}) = \begin{cases} \int_{\Theta} f(y|\lambda)\, p(\lambda|\boldsymbol{x})\, d\lambda; & \text{for continuous case} \\ \sum_{\Theta} f(y|\lambda)\, p(\lambda|\boldsymbol{x}); & \text{for discrete case} \end{cases} \tag{2.8}$$

which is the posterior predictive distribution of $Y = x_{n+1}$ given data $(\boldsymbol{x} = x_1, \ldots, x_n)$.

2.13 Bayesian Hypothesis Testing

Directly probability statements about uncertainty essentially require Bayesian analysis. The hypotheses are uncertain, and the result of a Bayesian analysis will be simply the statement of the believed probabilities of the hypothesis. Suppose we have two hypotheses, the null hypothesis H_0 that the unknown parameter θ belongs to some interval $\Theta_0; (\theta \in \Theta_0)$, verses the alternative hypothesis that H_1 that θ belongs to the alternative set $\Theta_1; (\theta \in \Theta_1)$ providing that $(\Theta_0 \cap \Theta_1) = \varphi$ (i-e there is no common points between the two sets). Then the task of deciding between the two hypotheses is conceptually more straightforward. Here we have to find posterior probabilities for both hypotheses. That is,

$$p\,(H_0) = p[(\theta \,|\, x) \leq \theta_0)] = \int_{0} p(\theta \,|\, x)\, d\theta, \tag{2.9}$$

and $\quad p(H_1) = (1 - p(H_0))$

Here the choice of deciding between H_0 and H_1 is clear-cut than classical testing. The hypothesis with greater probability will be accepted. The decision rule used here, for accepting or rejecting the above hypotheses is;

let s =min $(p(H_0), p(H_1))$, if $p(H_0)$ is small then H_1 is accepted and if $p(H_1)$ is small, H_0 is accepted. And if s>0.1, the decision is inconclusive (Aslam 2002).

2.14 Gibbs Sampling

Gibbs sampling is well suited to coping with incomplete information and is often suggested for such applications. However, generality comes at some computational cost, and for many applications including those involving missing information there are often alternative methods that have been shown to be more efficient in practice. Geman and Geman [1984] place the idea of Gibbs sampling in a general setting in which the collection of variables is structured in a graphical model and each variable has a neighborhood corresponding to a local region of the graphical structure. Geman and Geman use the Gibbs distribution to define the joint distribution on this structured set of variables. In the case of Bayesian networks, the neighborhoods correspond to the Markov blanket of a variable and the joint distribution is defined by the factorization of the network. Gibbs sampling is a general inference algorithm. Gibbs sampling can be used to learn Bayesian networks with missing data. The first step is to represent the learning problem itself as a Bayesian network.

The Bayesian statistician needs computational tools to calculate a variety of summaries from posterior distribution that are mathematically complex and also often high dimensional. Geman and Geman (1984) introduce the Gibbs sampler via simulation from the high-dimensional distributions arising in image restoration. The method, actually, based on the work of Metropolis, Rosonbluth, Rosonbluth, Teller and Teller (1953) for studying Boltzmann distributions from statistical mechanics and further development by Hastings (1970). More development can be seen in Gelfand and Smith (1990).

Aslam (2007) discusses in detail about the mechanism under Gibbs sampling. He shows that the Gibbs sampler is a method for sampling from a multivariate probability density function. It operates by simulating a random variate from univariate distribution only. Let us consider a k-dimensional density function $f(\psi) = f(\psi_1, \psi_2, ..., \psi_k)$ and we are interested to draw a sample from it. The 'k' one-dimensional conditional densities $f(\psi_s \mid \psi_t, s \neq t, s = 1,...,k)$ up to proportionality are required. We specify an arbitrary value of $\psi^{(0)} = (\psi_1^{(0)}, \psi_2^{(0)}, ..., \psi_k^{(0)})$ and simulate a random variate $\psi_1^{(1)}$ from the conditional distribution $f(\psi_1 \mid \psi_2^{(0)}, \psi_3^{(0)} ..., \psi_k^{(0)})$, a random variable $\psi_2^{(1)}$ from the $f(\psi_2 \mid \psi_1^{(1)}, \psi_3^{(0)} ..., \psi_k^{(0)})$ using the simulated random variate $\psi_1^{(1)}$ and similarly $\psi_k^{(1)}$ from the conditional distribution $f(\psi_k \mid \psi_1^{(1)}, \psi_3^{(1)} ..., \psi_{k-1}^{(1)})$. This completes single iteration. After 'm' such iterations we have $\psi^{(m)} = f(\psi_1^{(m)}, \psi_2^{(m)} ..., \psi_k^{(m)})$ and the Gibbs sequence of random variate is $\psi^{(1)}, \psi^{(2)}, ..., \psi^{(m)}$. In this way, the distribution of sequence (which is a Markov chain of random variables simulated by the Gibbs sampler) converges to the distribution of interest. This sequence can be used to estimate some characteristics of distribution. He also presents two sampling methodologies, *one run sequence* and *Parallel sequence*. The ratio-of-uniforms method is illustrated. A program is designed in SAS package for the application of the Gibbs Sampler. [For more detail see Aslam (2007)]

The Gibbs sampling, one of the several MCMC method, is called Monte Carlo because they involve drawing random numbers from specified distribution and Morkov chain because each sample depends upon the previous sample.

2.15 Foundation of Bayesian Statistics

This branch of statistics comes into being after the Bayes's Rule (1763), which is developed by Thomas Bayes (1702-1761). The Bayesian school of thought is almost in dominating position since its beginning over classical or frequentist statistics. Its application is present in almost all areas of research. The general outlook of Bayesian probability is promoted by Laplace (1812). Amster (1963) presents further modification and calls it Bayes' stopping rule. Other well-known proponents of Bayesian probability have included Savage (1962) and Ramsey (1926). Sacks (1963) and Robbins (1964) also present solutions in estimation problem under Bayesian approach. Tiao and Tan (1966) present Bayesian analysis of random effect models in the analysis of variance. A comprehensive review is presented by Lindley (1971). Smith (1973) modifies Bayes estimates in the one-way and two way models. Lord and Cressie (1975) make use of this approach in interval estimation. Akaike (1978) enlightens Bayes' procedure with a different look. Further comments on Bayesian methods can also be viewed by Lindley (1980). Press (1982) presents comparison of Bayesian approach with Frequentist approach in multivariate analysis. Susarla (1982) applies Bayes theory in Encyclopedia of Statistical Science. Smith (1983) also discusses Bayesian approaches to outliers and robustness in specifying statistical models. Louis (1984) presents some comments on estimating a population of parameter values using Bayes and empirical Bayes methods.

2.15.1 Scope of Bayesian Statistics

Bayesian analysis is an essentially self-contained paradigm for statistics. Consider, for instance, when important prior information is available then failure to take prior information into account can lead to conclusions ranging from merely inferior to illogical. Of course, most non-Bayesians would agree to the use of reliable and significant prior information, so the impact of this consideration for general adoption of the Bayesian is unclear. So when significant prior information is available, the Bayesian approach shows how to sensibility utilize it, in contrast with the most non-Bayesian approaches.

It is being used in Basic hypothesis testing and estimation, Design and sample-size computations, Linear and non-linear regression, Non-parametric statistics, Econometrics, genetics and spatial. In statistical hypotheses testing, Bayesian analysis will be simply the statement of the believed probabilities of the hypothesis in the light of the data and the prior information.

Bayesian methods are gaining popularity in many areas such as clinical trials, genomics, marketing, environmental science, and other fields where prediction and decision making must follow from statistical analysis. Since Bayesian methods are highly computational, they are also gaining wider acceptance as technology makes analyses possible that were not feasible in the recent past.

Bayesian techniques have wide-ranging uses within the financial sector. By its intrinsic nature, Bayes' Theorem lends itself perfectly to use within risk management. It is therefore commonly exploited within hedging and within quality and group management. Banks and credit card companies are using Bayesian techniques.

Bayesian methods have also changed the face of computer network security and the detection of credit card fraud. Companies would monitor the spending on the card, and anomalous behaviour can be seen as a possible sign of fraud. Bayes'Theorem is crucial in assessing the likelihood of fraud given the spending patterns. The Bayesian process means that patterns of behaviour are sought rather than individual anomalies - the sort that could lead to incorrect results.

An interesting example is related to American Telecommunication's (AT) detection of telephone fraud, a problem that costs the U.S. telephone industry around $4 billion per year. AT's system can be represented by a graph showing a "fraud score" over time, where a high score is recorded for unusual calls (calls to a previously uncalled country, calls of unusually long duration). Bayesian techniques are utilized to calculate the probability of a call being fraudulent. A record is kept of the *expected financial loss* on the account. This was measured as the call value multiplied by the probability that it was fraudulent. Once the loss had exceeded a certain threshold, the account could be deemed fraudulent and action taken.

This system can limit false alarm cases, because the Bayes approach will consider all calls in the region of a particular call when determining that call's fraud probability. Thus, over several calls, only probable fraud cases can be shown to exceed the financial loss threshold. This example serves to highlight the increasing awareness of using Bayesian Statistics as an everyday inference method.

Bayesian Statistics have been used in a diverse range of other software systems.

The US Navy have developed real-time software for determining the performance of various ship self-defense weapon systems against varying types and ranges of incoming attack weapons. Traditional techniques to solve the problem had been unsuccessful, but an approach that involved the use of Bayesian Networks led to a solution that was both effective and efficient.

The Vista system is a decision-theoretic system that has been used at NASA Mission Control Center in Houston for several years. The system uses Bayesian Networks to interpret live telemetry and provides advice on the likelihood of failures of the space shuttle's propulsion systems. It also considers time criticality and recommends actions of the highest expected utility. Intel is using Bayesian Networks for diagnosis of faults in processor chips. Given end-of-line tests on semi-conductor chips, their statistical process can be used to infer possible processing problems.

Nokia Networks uses the Hugin Decision Engine (a commercial tool making use of Bayesian Networks) in a prototype tool for efficient diagnosis of mobile networks. By having an automated tool that reads network performance data and from that estimates and monitors network problems ranked by probability, the network operator gets an efficient troubleshooting procedure saving both expensive expert resources and downtime of the network. Bayesian Networks have proven to be far better at expressing the belief of a set of potential problems than other techniques.

Bayesian systems are also being used in the continuing fight against spam, the unsolicited marketing and other junk email that deluges most companies email systems. An open source anti-spam email filter, called POP File, can be downloaded from the internet, which makes use of a simple Bayesian component that "learns" how to recognize spam and differentiate it from non-spam. This is achieved by training the software – telling it which of the emails you receive is acceptable, and which are spam.

es the words in the message, and builds up a model of the ord. After training, the system can then accurately predict which to be spam. In practice this system has been found capable of am emails.

Chapter 3

PAIRED COMPARISONS: METHOD AND MODEL

3.1 Introduction

This chapter covers the method of paired comparison in detail. Basic units, mechanism and derivation for the models of paired comparisons are elaborated in section 2. Applications of paired comparisons are detailed in section 3. Advantages of paired comparisons are highlighted in section 4. The existing inscription about paired comparison models is being considered in section 5.

3.2 Paired Comparison Method

The method of paired comparison is used basically in cases when the objects are compared subjectively. Paired comparison analysis helps us to set priorities where there are conflicting demands on resources. It has found increasingly used in applications. It is being used in epidemiology, food science, optics, sports and others. Paired comparisons are widely employed by psychometricians. Most frequent applications have been to sensory testing, especially taste testing, personnel rating and quite generally to the study of preference and choice behavior. Paired comparison data provide a rich source of information about individual differences and similarity relationship in the item evaluations.

Before going in other field of application of paired comparisons, we first introduce the method of paired comparisons, how it works and what are its major functions. In this method, objects are presented in pairs to one or more judges for the purpose of comparison. The "objects" may be "a person", "a treatment", "stimuli", and the like. The basic experimental unit is the comparison of two objects. The purpose of paired comparison experiment is to test the null hypothesis that every preference is equally likely against an unclearly defined alternative of consistency. The paired comparisons method imposes minimal constraints on the response behavior of a judge; internal consistency checks are available which allow for the identification of judges who are systematically inconsistent in their judgments.

Suppose we have two treatments T_1 and T_2, by a single judge, who must choose one of these treatments. It is the example of simplest situation in paired comparisons when no tie is allowed i.e. we shall say that the judge prefers this treatment although the choice will not necessarily represent a preference.

Some extent to the above stated simplest situation is that the judge may be allowed to declare a tie, or asked to record a preference on some finer scale. Also number of judges may be more than one. Then all the judges will make the comparison between two treatments and result can be obtained in following three categories.

1. $T_1 \rightarrow T_2$ *(T_1 is preferred on T_2)*
2. $T_2 \rightarrow T_1$ *(T_2 is preferred on T_1)*
3. No preference *(The judge is unable to decide between two treatments.)*

And the best treatment will be that contains maximum numbers of preference by the judges. It is also noted that these following situation also result in category of "No preference"

(i) The Judge may be strongly in favor of both treatments.

(ii) The Judge may be weakly in favor of both treatments.

(iii) The Judge may found both treatments at the same moderate level of preference.

For more than two objects are under observation then it is still easy to arrange that every judge performs every possible pair of these treatments. This situation may be called a "Balance paired comparisons experiment". Suppose we have three treatments T_1, T_2 and T_3, and then the possible numbers of pairs will be 3. And suppose there are 4 judges, then total number of paired comparisons will be 12.

Generally, for t objects and n judges the total number of paired comparison will be

$$n\binom{t}{2}, \qquad \text{(without regard of order and)} \qquad (3.1)$$

and

$$n\left({}^{t}P_{2}\right), \qquad \text{(with regard of order)} \qquad (3.2)$$

The method of paired comparison has some advantages when a fine judgment is needed. It is sometimes the obvious experimental procedure. Sometimes, the judge may be able to compare several objects at the same time. Then a simple ranking of all objects may well be preferable. However, when differences between objects are small, it is desirable to make the comparison between two of them as free as possible from any extraneous influence caused by the presence of other objects. Otherwise, process of ranking requires in practice many repeated pair wise comparisons of tentative neighbors before a reasonable ordering becomes established. It is also possible to score a paired comparison on some point scale. Number of point on that scale will depend upon the differences among the treatments under comparison. For example, let take 5-point scale

Table 3.1 **Preference Score**

Score	2	1	0	-1	-2
Preferences	Strong preference for T_1	Slight preference for T_1	No Preferenc e	Slight preference for T_2	Strong preference for T_2

The successfulness of this score depends on the existence of differences that should be clear enough.

For pair wise comparisons of three objects *A, B* and *C* are to be judged in pairs. Since this simple situation brings out many of the essential features of paired comparisons experiments. Consider the case in which ties are not permitted. Each of three comparisons *(AB), (AC), (BC)* has two possible outcomes, so there are eight distinguishable experiment results. Six of these are of the type

$$A\rightarrow B,\ A\rightarrow C,\ B\rightarrow C, \tag{3.3}$$

It is also noted that one object is preferred two times, another one time and the third none. The two remaining results are

$$A\rightarrow B,\ B\rightarrow C,\ C\rightarrow A, \tag{3.4}$$

$$B\rightarrow A,\ C\rightarrow B,\ A\rightarrow C, \tag{3.5}$$

Kendall and Babington Smith (1940) called them (3.4) & (3.5) circular triads. A circular triad denotes an inconsistency on the part of the judge and its simplest explanation is that the judge is at least partially guessing when declaring preferences.

But circularity can occur even with a well-defined preference criterion. For example, there are three cricket teams Pakistan, Australia and England and it is quite possible that

(i) *Australia wins England*

(ii) *England wins Pakistan*

(iii) *Pakistan wins Australia*

It is valuable feature of the method of paired comparisons that it allows such contradictions to show themselves. Every practical precaution must be taken to ensure that the individual comparisons are independent or nearly so. But in personnel rating, there is a real danger that, due to a good memory, the judge's paired comparisons will degenerate into a ranking unless the number of people to be compared is very large. For a single judge, one partial way out of this problem is to make only a fraction of all possible comparisons. If in a particular experiment, approximate independence has not been achieved, then the situation is intermediate between a straight ranking and independent paired comparisons. The analysis should therefore be made according to the both methods. Only when the result agrees can a conclusion be drawn with any comfort.

The forgoing discussion may be formalized in a number of possible models, which impose increasingly severe restriction on the preference probabilities. The method of paired comparisons has led to a surprising amount of model building to provide stochastic representation of the experimental process.

Suppose there are t objects to be compared in pairs by the judge. Let X_{ij} $i, j = 1, ..., t$ is an indicator random variable which can take value 0 or 1 according as judge prefers T_i or T_j. We assume throughout that all comparisons are statistically independent except that $X_{ij} + X_{ji} = 1$, then the preference probabilities are

$$Pr\,(X_{ij} = 1) = \pi_{ij} \text{ (Preference probability that } T_i \rightarrow T_j) \tag{3.6}$$

$$Pr\,(X_{ji} = 1) = \pi_{ji} \text{ (i-e Preference probability that } T_j \rightarrow T_i) \tag{3.7}$$

More generally

$\pi_{ij} = \pi_{ijkl}$ (With replication and judge effect)

That is, T_i is preferred on T_j at k'th Comparison from l'th judge. As π_{ij} and π_{ji} are probabilities so, $0 < \pi_{ij}$, $\pi_{ji} < 1$

In case of three objects different values of π_{ij}, π_{ji} and π_{jk} may cause Stochastic transitivity, strong stochastic transitivity and moderate stochastic transitivity. The linear models for paired comparisons are of importance. Suppose that the objects have "merit" V_i when judged on some characteristic and may be represented by the continuous r.v Y_i $(-\infty < Y_i < +\infty)$. In a paired comparisons T_i and T_j, T_i will be preferred if $Y_i > Y_j$ and T_j if $Y_j < Y_i$. Let $Z_i = y_i - V_i$, $i=1, 2,\ldots, n$ if every pair (Z_i, Z_j) has the same bivariate distribution then $(Z_i - Z_j)$ must have the same distribution as $(Z_j - Z_i)$,Now

$\Pr\{(Z_i - Z_j) < x\} = H(x)$,.It follows that

$$\pi_{ij} = \Pr\{(y_i - y_j) > 0\}$$

$$= \Pr\{(z_i - z_j) > -(V_i - V_j)\}$$

$$= H(V_i - V_j)$$

Whenever the preference probabilities can be expressed in terms of a symmetrical cdf, the y_i may be said to satisfy a linear model. This linear model is a generalization of the Thurstone-Mosteller model for which y_i are assumed to be normal $N(V_i, \sigma^2)$ variates, equi-correlated with common correlation co-efficient ρ.

$$\pi_{ij} = H(V_i - V_j) = \int_{-(v_i' - v_j')}^{\infty} z(x)dx \tag{3.8}$$

$$\text{Where } V_i' = \frac{V_i}{\left[2\sigma^2(1-\rho)\right]^{1/2}} \text{ and } z(x) = \frac{1}{\sqrt{2\pi}} e^{-1/2x^2}$$

another special case is provided by the Bradley-Terry model (Bradley, 1953) for which

$$H(V_i - V_j) = \frac{1}{4} \int_{-(\ln\theta_i - \ln\theta_j)}^{\infty} \sec h^2(y/2)dy \tag{3.9}$$

Here (3.9) is known as the Bradley-Terry model. It is the fundamental model of paired comparisons. Various authors make several modifications of this model. We shall discuss these extensions in section 4.

3.3 Application of Paired Comparison Method

Paired comparison is a practical technique for comparing items. These items may be ideas, options or criteria etc. Number of items may be "between" 10-15. The goal of paired comparison statistics is to deduce a ranking from an uneven matrix of observed results, from which the contestants can be sorted from best to worst. In the knowledge that crushing all the complexities of the situation into just one number is a large simplification, one wishes to have the best one-dimensional explanation of the data. Paired comparisons statistics are an open research area.

What flavors of Ice Cream do you prefer?

	Flavors	Comparison		
1.	Chocolate	1	1	1
		2	3	4
2.	Vanilla		2	2
			3	4
3.	Stawberry			3
4.	Black Walnut			4

This is an experiment of using the paired comparison technique to decide something simple like the flavor of ice cream we prefer. We start by writing down the flavors of ice cream: Chocolate, Vanilla, Stawberry and Black Walnut. Then we compare one flavor to the next and "bold" our preference. Here to the right of "Chocolate" we "bold" the first numerical one, 1, indicating that we prefer Chocolate (#1) over Vanilla (#2). We then "bold" (#3) indicating that we prefer Stawberry over (#1) Chocolate. In the last column we "bold" (#1), preferring Chocolate over (#4) Black Walnut.

We continue this selection process until we have made a choice of each pair of flavors. Next, we count up the number of times we "bold" each number. The number "bold" the most frequently, reflects our ice flavor preference. In this experiment, we "bold" Chocolate (#1) two times, Vanilla (#2) one time, Stawberry (#3) three times, and Black Walnut (#4) is not "bold". Using the paired comparison technique, Stawberry is our favorite flavor of ice cream.

Rosenberger *et al* (2002) discuss method of paired comparisons to measure economic values for multiple goods sets. A method of paired comparison is adapted for use in estimating economic measures of value. The method elicits multiple binary choices for paired items in a choice set. The method is applied in an experimental context with a choice set composed of four private goods and several sums of money. Malcolm *et al* (2002) present a paper for policy search using paired comparisons. The paired-comparisons method offers a more accurate and precise approach.

3.4 Advantages of Paired Comparison Method

In paired-comparison designs the consumer is asked to use two products and determine which product is better. The paired comparison is a wonderful design if presenting evidence to a jury, because of its "face value" or "face validity". It can be a very sensitive testing technique (i.e., it can measure very small differences) between two products.

The paired-comparison test is often less expensive than other methods, because sample sizes can be smaller in some instances. The repeated paired comparison test is another feature of paired comparisons. The purpose of the repeated paired-comparison taste test is to identify non-discriminators.

Paired-comparison testing, however, is limited in value for a serious, ongoing product-testing program. The paired-comparison test does not tell does and us when both products are bad not lend itself to the use of normative data. It is heavily influenced by the "interaction effect" (i.e., any variations in the control product will create corresponding variance in the test product's scores).

3.5 Review of the Paired Comparison Models

The basic model for paired comparisons has been discovered and rediscovered by various authors. It is also said that it arises as one of the special simple realization of a generalized model developed from psychophysical approaches.

Thurstone (1927) proposes a linear model for paired comparisons .He models paired comparisons through the concept of a subjective scale, an inbuilt sensation scale on which order but not physical measurement could be distinguish. He assumes that each element involved in paired comparisons generates a sensation. The element with largest sensation is the one that is chosen. It is an attempt to measure relative differences and to indicate one possible procedure in extending at least some of the ideas of psychophysical measurement to social values. The experiment data consists in the observed proportion of judgments and from these data the best fitting scale values of the stimuli are determined. Mosteller (1951) derives the same form when sensations are correlated with each other by a constant correlation coefficient.

Zermelo (1929) seems to have proposed it first when dealing with the estimation of the strengths of chess players in an uncompleted Round Robin tournament. His model is independently rediscovered by Bradley-Terry (1952) who demonstrates its usefulness in sensory testing. Ford (1957) also presents his research paper independently.

Kendall & Smith (1940) examine the method of paired comparisons. They say that suppose we have number of objects A, B, C etc which possesses some common quality. According to them, quality may be measurable and it may not be measurable. If the quality is measurable then we assigned variable values to the objects. If the quality is not measurable then we use the method known as Ranking. They say that Ranking is the arrangement of objects in order to possess certain quality.

They observe that the raking method is not good when the quality considered is not known with certainty. For example, if an observer ranks a number of individuals in order of intelligence, it is not impossible that the observer may judge A as more intelligent than B, B than C and C than A; if the individual are presented for his consideration one pair at a time. The chance of this happening is obviously increased when we are dealing with tastes in music, eatable or film stars; and in practice the event is not uncommon. They claim that such "inconsistent" preferences can never appear in ranking , if A is preferred to B and B to C, then A must automatically be shown as preferred to C .So the use of ranking thus destroys what may be valuable information about preferences.

They explain that the method of offering for judgment objects two at a time is known as the method of paired comparisons. This method has interesting application in animal experimentation.

Mosteller (1951) provides a detailed formulation and analysis of Thurstone model.

Glenn & David (1951) observe that in paired comparison experiments a judge is often enable to express any real preference in a number of the pair he judges. They examine that some of the methods in current use do not permit the judge to declare a tie. In other cases ties are permitted, but are ignored in performing the analysis. In some other cases, ties are divided, equally or randomly, between the tied members of a pair.

Using a model of the Thurstone-Mosteller (1951) type they develop a method which makes provisions for tied observations.

Scheffe (1952) develops a method of paired comparisons which differs from the others in that it uses a scoring method and the analysis of variance. The method has the feature that the effect of order of presentation of paired samples to the judges is taken into account. This method seems worthily suited to consumer preference studies wherein a considerable time lag may occur between the testing of the two samples of a pair.

Bradley (1952) by using the logistic density function provides a model for paired comparisons. He considers a case with no tie and order effect.

Bradley (1953) considers typical examples of statistical methods used in taste testing and procedures which are applicable in taste testing. The main emphasis is placed on the method of paired comparisons. The author is interested in the application of discriminate function techniques to the establishment of weights for the scores of various attributes in grading. Two methods of paired comparisons are of interest.

Bradley (1954) also develops a procedure for testing the appropriateness of the model for the method of paired comparisons. The proposed test is applied to a variety of experiments involving taste, preference and appearance. The writer presents two tests for treatment affects and one test for agreement among judges. A basic representation of paired comparisons experiment may be defined to be a set of incomplete blocks. The method of analysis for paired comparison depends on the form in which data are recorded. Writer's main consideration is to test of goodness of fit for the simple paired comparisons model. The writer shows enough data to provide some assurance that the model may be appropriate for experiments involving subjective judgments.

Bradley (1955) investigates large-sample properties of the test with equal numbers of selection decision. He obtains the asymptotic relative efficiency for balanced paired comparisons with the analysis of variance for a comparable size.

Bose (1956) has used paired comparison designs for testing concordance between judges. The object of his paper is to obtain some paired comparisons designs which have a high degree of symmetry. Two special classes of these designs have been investigated and explicit designs for small values of number of objects to be compared. The method of analysis depend on what use the experimenter wants to make of the design. He concludes that for any solution of the symmetrical balanced incomplete block design, we can drive a corresponding solution of the linked paired comparison designs.

Ford (1957) proposes the model for paired comparisons independently. He concentrates on solution of normal equation for parameter estimation and proved under stated conditions that the iterative procedure converged to a unique maximum for likelihood solution.

David (1959) submits a paper "Tournaments and Paired comparisons ". In this paper, he feels convenient to use the language of tournaments rather than that of paired comparisons. The analogy between a paired comparison experiment and a Round Robin tournament, first pointed out by Kendall (1940). Round Robin Tournaments have received considerable attention both from statisticians and Psychometricians. Bradley-Terry (1952) has proposed a model to represent the strength of the players. However this model does not provide an answer to such question as what constitutes a significant difference in the scores of two specified players. To deal with problems of this type, the joint probability distribution of the scores of any players in a simple round robin tournament is found.

Gridgeman (1959) brings light the importance of tied observations in the analysis by taking different experimental situations. He observes that when discrimination is the objective, admission of tied decisions theoretically increases the power the test of the null hypothesis. But he claims that in practice the admission of tied decisions may be decreased the subject's efficiency of decision and in these circumstances it is better to prohibit ties. He examines that in preference studies, ties should be admitted as they add information.

Thompson and Singh (1967) propose a Psychophysical model for paired comparisons.

Rao & Kupper (1967) states that the Bradley-Terry model for a paired-comparison experiment with t treatment do not consider "ties" in the model. Rao-Kupper introduces an additional parameter, called threshold parameter, into the model. This permit "ties" in the model. They say that in paired-comparison experiment in which a panel of judges ranks pair of treatments on the basis of some quality, a judge may not be able to express any real preference in a number of pair he judges. The inability of a judge to express a preference in a particular case may be due to one or both of the following reasons:

(1) His sense of perception is not sharp enough to detect the existing treatment difference

(2) The treatments do not differ in the quality judged.

They observe that for detecting the treatment differences on the basis of ranks it is necessary to analyze the differences between "true" treatment ratings rather than panelists' abilities to detect these differences. They examine that most of the existing models do not allow for the probability of a judge declaring a tie. In such cases, the usual practice is either to force the judge to express a definite preference, or, if this is not done, to treat these ties in one of the followings ways:

(a) They are completely ignored; (b) they are dividing equally between the tied numbers; or (c) they are dividing randomly between the tied numbers.

They claim that any model, which does not allow for the possibility of ties is not making full use of the information contained in the no-preference class.

Draper et al (1969) elaborate methods for analyzing data of three preference categories. These methods take proper account of the "no-preference" results. Some of these treatments are listed by Odesky (1967). An exact solution is given, and some straightforward approximate methods are derived and illustrated. Bayes' theorem is used to combine the sample information and available prior information. The concept of highest probability density (H.P.D) regions is used for posterior probabilities. H.P.D contours are plotted and illustrated through numerical example.

Grizzle et al (1969) present a method for analyzing categorical data that relies on finding a transformation to produce a linear model of full rank in transformed parameters with the resulting analysis following the method of weighted least squares. Robert. J Beaver (1977) present a paper in which the implementation of the Grizzle approach to the analysis of several univariate paired-and triple-comparison models derived from the Bradley-Terry model, is discussed. Paired and triple-comparison models are included. This method of analysis produces noniterative weighted least square estimates of preference ratings corresponding to the treatment under test and allows for testing hypothesis relative to their values within the framework of general linear hypothesis. Examples with and without ties or order effect for paired comparisons experiments and an example of a triple comparison experiment are presented. A generalized chi-square computer program can be used to implement the computations.

Davidson (1969) observes the relationship between two representations of the Bradley-Terry model for paired comparisons. He examines that if the responses are pair wise independent and are distributed according to the extreme values distribution, a member of a class of distributions proposed by Lehman, then the representation in terms of the logistic distribution is obtained.

Davidson (1970) develops a model for a paired-comparison experiment with allowing ties. He uses the Maximum Likelihood method to estimate the parameters of the model. He observes that his results of ranking agree with the method of scoring system. His model also satisfies the criteria of the goodness of fit test.

Davidson & Bradley (1971) present a model for multivariate paired comparison and a test of significance of the responses to specified attributes in estimating the responses to overall quality. They examine the problem of relating the response pattern on overall quality to that on a specified set of attributes. They derive a regression equation for a joint distribution of responses to dichotomous items and are applied to the multivariate-paired comparison model.

Davidson & Solomon (1973) use two models namely Multinomial and the Bradley-Terry for paired comparison experiments. They apply natural conjugate family of priors to represent prior beliefs. They observe that the ranking results using both estimators of the models are identical.

Beaver and Gokhale (1975) propose an additive order effect in Bradley-Terry – Luce Model. Davidson and Beaver (1976) present a multiplicative order effect.

While comparing a set of stimuli, the final rankings of the stimuli are affected by the order effect. This effect is usually eliminated by an appropriate design for balancing the order effect of the stimuli in a pair. Beaver and Gokhale (1975) modified the Bradley-Terry model to incorporate within pair order effect. Scheffe (1952) developed an

ANOVA for scored differences from pairs that allows for order of presentation of items within a pair.

Leonard (1977) explains that new Bayesian approaches have been developed for the estimation of parameters like conjugate prior distributions are replaced by logistic transformations and employing multivariate normal distributions for the transformed parameters. He observes that logistic transformation is more flexible technique for prior relationship between the parameters. He says that posterior estimates enable the statistician to combine information from various sources and to cope with zero cell frequencies in a formal manner. The Bradley-Terry model uses the same technique for treatment parameters. He provides alternative technique to the procedure proposed by Davidson and Solomon. He observes that when the parameters are priori exchangeable, the maximum likelihood estimates are constrained towards a common value. He uses biological example to investigate the distribution of genital display in a colony of squirrel monkeys by using the Bayes Estimates. He includes also incomplete contingency table with several zero entries.

Buchanan & Morrison (1985) explain that when comparing alternate product formulations, or evaluating competitors' formulations, researcher often conduct blind, forced choice product tests. In this study, the authors present an approach to designing these tests and evaluating their results. Two types of tests are considered: repeat paired comparisons formats, and formats consisting of several triangle tests and a single paired comparison. The approach consists of psychophysical assumptions, discrimination and preference constructs, and a set of analytical technique. Using this approach, a researcher can compare product test format to see which one most efficiently estimates a given construct of interest.

Joe (1990) says that linear paired comparisons models are studied when ties or draws are allowed, where the probability of a win plus half the probability of a draw is modeled as a symmetric function evaluated at the difference of strength parameters. He observes that these models are extended to make use of covariate information and used for ranking 64 top chess players since 1800 with information on career periods. He examines that the Davidson model, which allows for draws does not fit chess data well because of the large variability in draw percentages from player to player. He also presents an appropriate goodness-of-fit test for this extended use of linear model.

Stern (1990) also scrutinizes the claim of previous authors (Jackson and Fleckenstein, 1957; Mosteller, 1958; Noether, 1960) that different model of paired comparisons data lead to similar result by means of a set of paired comparisons models, based on gamma random variables, that includes the frequently applied Bradley-Terry and Thurstone-Mosteller models. He also analyzes several sports data sets and concludes that all the paired comparisons models in the family provide adequate and almost identical fits to the data.

Groeneveld (1990) uses the numbers of wins and losses in games between each pair of team in a League, discussed and compared several method of ranking. The writer considers three nonparametric ranking methods. He also considers ranking based on the classic Bradley-Terry parametric procedure. The properties of the four ranking methods in the competitive situation described, examined and contrasted. The nonparametric ranking methods appeared very simple in explanations. The empirical evidence from the

National league data suggests that the Bradley-Terry model fits reasonably well. This method may have value for ranking teams in this type of competitive situations.

Hennery (1992) describes a simple extension to the Thurstone-Mosteller model that allows for widely differing properties of ties. The corresponding modification to the Bradley-Terry model would also be possible but would give very similar results for Keene-Divinsky data. They Keene-Divinsky idea of including only the top players in deciding the best players of all times results in dilemma and it is suggested that large number of top players should be included. Other source of bias may also reduce by including more games.

Smith & Roberts (1993) present the use of the Gibbs sampler for Bayesian computation. They describe other Markov chain Monte Carlo simulation methods and comment on the advantages of sample-based approaches for Bayesian Inference.

Kuk (1995) purposes a linear model that allows a large number of draws and large variability of draw percentages among the players as is the case for chess and soccer matches. The model can also be extended to allow home ground advantages. When only summary results are available, ML Estimation is not feasible. So a method is recommended based on matching the numbers of home wins, home draws, away wins and away draws for each team with their expected values. Problem of obtaining estimated standard error is also discussed. Mostly underlying variable models can be reduced to the linear paired comparison model. If the underlying variables are independently exponentially distributed with different scale parameter then we have the Bradley-Terry model. It is suggested that to avoid confounding, it is desirable to introduce draw parameters that must be separated from the strength parameters. To achieve this, the writer generalizes the Thurstone- Mosteller model along the lines suggested by Glenn and David (1960) and Hennery (1992). The resulting model allows large numbers of draws and large variability of draw percentages among teams as well as the possibility of home ground advantage.

Glickman and Stern (1998) develop predictive models for National Football League game scores using data from the period 1988-1993. The parameters of primary interest "measure of team strength" are expected to vary over time. The purposed model accounts for this source of variability by modeling football outcomes using a state-space model that assumes team strength parameters follow a first order autoregressive process. Two source of variation are addressed in model.

The aim of the analysis is to obtain plausible inferences concerning team strengths and other model parameters, and to predict future game outcomes. Iterative simulation is used to obtain samples from the joint posterior distribution of all model parameters.

Glickmen (1999) says that a likelihood-based analysis is computationally cumbersome when the population of objects being compared is large. He examines that this problem is overcome through a computationally simple non-iterative algorithm for fitting a particular dynamic paired comparison model. The method is evaluated on simulated data and is applied to ranking the best chess players of all time, and to ranking the top current tennis-players.

Yao & Bockenholt (1999) present a Gibbs sampler for the estimation of Thurstonian ranking models. This approach is useful for the analysis of ranking data with a large number of options. They discuss the goodness-of-fit based on posterior predictive distributions.

Martin (1999) presents an analysis of the effect of various baseball play-off configurations on the probability of World Series. Play-off games are assumed to be independent. He considers several paired comparisons models for modeling the probability of a home team winning a single game as a function of the winning percentages of the contestants over the course of the season. He examines that the uniform and the logistic regression models are both adequate, whereas the Bradley-Terry model is not. He uses the single-game probabilities to compute the probability of winning the play-offs under various structure

Conner & Grant (2000) give an extension of Zermelo's model for ranking by paired comparison. Zermelo (1929) proposed a probabilistic model for ranking by paired comparison and showed that his model produces a unique ranking of the objects under consideration when the outcomes matrix is irreducible. When the matrix is reducible, the model may yield only a partial ordering of the objects. In this paper, they analyze a natural extension of Zermelo's model resulting from a singular pertubaration. They show that this extension produces a ranking for arbitrary (nonnegative) outcome matrices and retains several of the desirable properties of the original model. In addition, they discuss computational techniques and provide examples of their use.

Aslam (2002) presents method for testing of hypotheses in paired comparison experiments. He uses the Rao-Kuper model with allowing ties. He applies the non-informative prior for the parameters of the model .He finds posterior means of the parameters and include predictive probabilities in the analysis .He also examines the goodness of fit criteria to Rao-Kupper model and presents the graphs of the posterior distributions of the individual parameters .

Aslam (2003) says that the Posterior distribution is the workbench in Bayesian analysis. He derives the Posterior distribution for the parameters of the Rao-Kupper model for paired comparison data using the (informative) Dirichlet-gama prior distribution. In this study, he tries to find analytical expression for the posterior (marginal) distributions of the parameters of the model. He presents five approximate analytical expressions for the marginal posterior distributions of the threshold parameters and three analytical expressions for the posterior distribution of the treatment parameters.

Annis & Craig (2005) observe that the existing paired comparison models used for ranking football teams basically focus on either wins and losses or points scored. They say that each approach fails to produce satisfactory rankings in frequently arising situations due to its ignorance of additional data. They propose a new hybrid model incorporating both wins and constituent scores and show that it outperforms its competitors and is robust against model misspecifications based on a series of simulation studies.

Aslam (2005) says that sometime it may be difficult for a panelist to rank or compare more than two objects or treatment at the same time. For this reason paired comparison method is used. He compares the Rao-Kupper model (1967) with Davidson model (1970) for paired comparisons allowing ties. For this purpose, he compares the posterior means, the posterior probabilities of the hypotheses for the comparison of two treatment parameters and predictive probabilities. He also presents the graphs of the parameters of both models.

Kim (2005) proposes a Bayesian Method for finding an optimal ranking in scalar functions of k population parameters. The writer presents a completely oriented graph that is closely related to the paired comparison ranking which can be used to make an optimal ranking. The writer also constructs a graphical model designed particularly for the paired comparison ranking in parameter functions of several probability models.

Allsopp (2005) explain the role of home ground advantage in Test, ODI and domestic cricket. He proposes a method to estimate a projected score for team batting second in ODI cricket. He uses linear and logistic modeling technique to explain the factors, which effect team performance such as home advantage. He also finds team rating by using linear model that account for the size of a victory in ODI cricket and the magnitude of the first innings lead in Test and domestic cricket. He recommends that new methods be investigated to officially rate and rank teams in international cricket competitions because his rating provides a robust measure of team quality.

Ovens & Bukiet (2006) present a mathematical modeling approach to ONE-DAY cricket batting orders. By using this model, they give expected performance (runs distribution) of a cricket batting order in an innings. They prove that their model enables one to solve for the probability of one team beating another or to find the optimal batting order for a set of 11 players. They use the same Markov Chain approach to study the progress of runs for a batting order of non-identical players as Bukiet *et al.* (1997) use in baseball modeling. They show that batting order does effect the expected runs distribution in ONE-DAY cricket.

Aslam (2007) says that the computation problems of the complicated, complex and intractable multiple integrands become very much easier through implementation of the Gibbs sampler. He explains that the Bayesian Statistician needs computational tools to calculate a variety of summaries from posterior distributions that are mathematical complex and also often high dimensional. Geman and Geman (1984) introduced the Gibbs sampler via simulation from the high-dimensional distributions arising in image restoration. In this study, he presents a program in the SAS (Statistical Analysis System) package for the application of the Gibbs Sampler.

Chapter 4

RANKING OF TOP SEVEN CRICKET TEAMS

4.1 Introduction

In this chapter, we present the Bayesian analysis of the Rao-Kupper and Davidson models for top seven cricket teams (Australia, South Africa, Pakistan, India, New Zealand, Sri Lanka and England) using collected data sets from the websites (www.cricmania.com & www.cricinfo.com). In section 2, the origin of cricket is discussed. In sections 3 and 4, the Rao-Kupper and Davidson models for paired comparisons are defined with basic notations and likelihood function for the parameters of the models. In section 5, the Ranking of top seven Cricket Teams using Posterior Modes and Posterior Means of the parameter of the Rao-Kupper model are presented.. In sections 6 and 7, the Predictive Probabilities and the Preference Probabilities are shown in the tables. In section 8, the Bayesian hypothesis testing is presented for overall ODI and Test matches. In section 9, the graphs of posterior marginal distributions for the parameters of the Rao-Kupper Model are shown. Appropriateness of the model is tested in section 10. In section 11, Latest ICC Rankings are presented with Formula for calculating points. In the last section 12, the comparisons of ICC Ranking and Bayesian Ranking are presented.

4.2 Origin of Cricket

Most probably cricket has originated in the dark ages. It has probably originated after the Roman Empire and surely before the invasion of Norman's in England. This game has originated some where in Northern Europe. The conclusion of the research is that this game is derived from an old pastime by which one player throw a small piece of wood or a ball and another hit it with a heavy and thick stick named as club.

It is unknown that when and how this club-ball game developed into one where the hitter defended a target against the thrower. It is also not known that when points were awarded dependent upon how far the hitter was able to dispatch the ball. It is also a mystery that when helpers joined the two-player contest, and this contest changed into a team game.

All researchers are agreed on it that by Tudor times cricket was developed enough from club-ball to be recognizable as the game played today. It was well established in many parts of Kent, Sussex and Surry; that within a few years it has become a feature of leisure time at a significant number of schools.

We present here important events of Cricket in different era.

(a) Cricket in 16th Century

1550 (approx) Evidence of cricket being played in Guildford, Surrey 1598 Cricket mentioned in Florio's Italian–English dictionary.

(b) Cricket in 17th Century

1610 Reference to "cricketing" between Weald and Upland near Chevening, Kent. 1611 Randle Cotgrave's French–English dictionary translates the French word "crosse" as a cricket staff. Two youths fined for playing cricket at Sidlesham, Sussex. 1624 Jasper Vinall becomes first man known to be killed playing cricket: hit by a bat while trying to catch the ball – at Horsted Green, Sussex. 1676 First reference to cricket being played abroad, by British residents in Aleppo, Syria. 1694 Two shillings and sixpence paid for a "wagger" (wager) about a cricket match at

Lewes. 1697 First reference to "a great match" with 11 players a side for fifty guineas, in Sussex.

(c) Cricket in 18th Century

1700 Cricket match announced on Clapham Common. 1709 First recorded inter-county match: Kent v Surrey. 1710 First reference to cricket at Cambridge University. 1727 Articles of Agreement written governing the conduct of matches between the teams of the Duke of Richmond and Mr Brodrick of Peperharow, Surrey. 1729 Date of earliest surviving bat, belonging to John Chitt, now in the pavilion at The Oval. 1730 First recorded match at the Artillery Ground, off City Road, central London, still the cricketing home of the Honourable Artillery Company. 1744 Kent beat All England by one wicket at the Artillery Ground. First known version of the Laws of Cricket, issued by the London Club, formalising the pitch as 22 yards long. 1767 (approx) Foundation of the Hambledon Club in Hampshire, the leading club in England for the next 30 years. 1769 First recorded century, by John Minshull for Duke of Dorset's XI v Wrotham. 1771 Width of bat limited to 4 1/4 inches, where it has remained ever since. 1774 LBW law devised. 1776 Earliest known scorecards, at the Vine Club, Sevenoaks, Kent. 1780 The first six-seamed cricket ball, manufactured by Dukes of Penshurst, Kent. 1787 First match at Thomas Lord's first ground, Dorset Square, Marylebone – White Conduit Club v Middlesex. Formation of Marylebone Cricket Club by members of the White Conduit Club. 1788 First revision of the Laws of Cricket by MCC. 1794 First recorded inter-schools match: Charterhouse v Westminster. 1795 First recorded case of a dismissal "leg before wicket".

(d) Cricket in 19th Century

1806 First Gentlemen v Players match at Lord's. 1807 First mention of "straight-armed" (i.e. round-arm) bowling: by John Willes of Kent. 1809 Thomas Lord's second ground opened at North Bank, St John's Wood. 1811 First recorded women's county match: Surrey v Hampshire at Ball's Pond, London. 1814 Lord's third ground opened on its present site, also in St John's Wood. 1827 First Oxford v Cambridge match, at Lord's. A draw. 1828 MCC authorised the bowler to raise his hand level with the elbow. 1833 John Nyren published his classic *Young Cricketer's Tutor and the Cricketers of My Time*. 1836 First North v South match, for many years regarded as the principal fixture of the season. 1836 (approx) Batting pads invented. 1841 General Lord Hill, commander-in-chief of the British Army, orders that a cricket ground be made an adjunct of every military barracks. 1844 First official international match: Canada v United States. 1845 First match played at The Oval. 1846 The All-England XI, organized by William Clarke, begins playing matches, often against odds, throughout the country. 1849 First Yorkshire v Lancashire match. 1850 Wicket-keeping gloves first used. 1850 John Wisden bowls all ten batsmen in an innings for North v South. 1853 First mention of a champion county: Nottinghamshire. 1858 First recorded instance of a hat being awarded to a bowler taking three wickets with consecutive balls.1859 First touring team to leave England, captained by George Parr, draws enthusiastic crowds in the US and Canada. 1864 "Overhand bowling" authorised by MCC. John Wisden's The Cricketer's Almanack first published. 1868 Team of Australian aborigines tour England. 1873 WG Grace becomes the first player to record 1,000 runs and 100 wickets in a season. First regulations restricting county qualifications, often regarded as the official start of the County Championship. 1877 First Test match: Australia beat England by 45 runs in Melbourne. 1880 First Test in England: a five-wicket win against Australia at The Oval. 1882 Following England's first defeat by Australia in England, an "obituary notice" to English cricket in the Sporting Times leads to the tradition of The Ashes. 1889 South Africa's first Test match. Declarations first authorised, but only on the third day, or in a one-day match. 1890 County Championship officially constituted. Present Lord's pavilion opened. 1895 WG Grace scores 1,000 runs in May, and reaches his 100th hundred. 1899 AEJ Collins scores 628 not out in a junior house match at Clifton College, the highest individual score in any match. Selectors choose England team for home Tests, instead of host club issuing invitations.

(e) Cricket in 20th Century

1900 Six-ball over becomes the norm, instead of five. 1909 Imperial Cricket Conference (ICC – now the International Cricket Council) set up, with England, Australia and South Africa the original members. 1910 Six runs given for any hit over the boundary, instead of only for a hit out of the ground. 1912 First and only triangular Test series played in England, involving England, Australia and South Africa. 1915 WG Grace dies, aged 67. 1926 Victoria score 1,107 v New South Wales at Melbourne, the record total for a first-class innings. 1928 West Indies' first Test match. AP "Tich" Freeman of Kent and England becomes the only player to take more than 300 first-class wickets in a season: 304. 1930 New Zealand's first Test match. Donald Bradman's first tour of England: he scores 974 runs in the five Ashes Tests, still a record for any Test series. 1931 Stumps made higher (28 inches not 27) and wider (nine inches not eight – this was optional until 1947). 1932 India's first Test match. Hedley Verity of Yorkshire takes ten wickets for ten runs v Nottinghamshire, the best innings analysis in first-class cricket. 1932-33 The Bodyline tour of Australia in which England bowl at batsmen's bodies with packed leg-side field to neutralise Bradman's scoring. 1934 Jack Hobbs retires, with 197 centuries and 61,237 runs, both records. First women's Test: Australia v England at Brisbane. 1935 MCC condemn and outlaw Bodyline. 1947 Denis Compton of Middlesex and England scores a record 3,816 runs in an English season. 1948 First five-day Tests in England. Bradman concludes Test career with a second-ball duck at The Oval and a batting average of 99.94 – four runs short of 100. 1952 Pakistan's first Test match. 1953 England regain the Ashes after a 19-year gap, the longest ever. 1956 Jim Laker of England takes 19 wickets for 90 v Australia at Manchester, the best match analysis in first-class cricket. 1957 Declarations authorised at any time. 1960 First tied Test, Australia v West Indies at Brisbane. 1963 Distinction between amateur and professional cricketers abolished in English cricket. The first major one-day tournament begins in England: the Gillette Cup. 1969 Limited-over Sunday league inaugurated for first-class counties. 1970 Proposed South African tour of England cancelled: South Africa excluded from international cricket because of their government's apartheid policies. 1971 First one-day international: Australia v England at Melbourne. 1975 First World Cup: West Indies beat Australia in final at Lord's. 1976 First women's match at Lord's, England v Australia. 1977 Centenary Test at Melbourne, with identical result to the first match: Australia beat England by 45 runs. Australian media tycoon Kerry Packer, signs 51 of the world's leading players in defiance of the cricketing authorities. 1978 Graham Yallop of Australia wears a protective helmet to bat in a Test match, the first player to do so. 1979 Packer and official cricket agree peace deal. 1980 Eight-ball over abolished in

Australia, making the six-ball over universal. 1981 England beat Australia in Leeds Test, after following on with bookmakers offering odds of 500 to 1 against them winning. 1982 Sri Lanka's first Test match. 1991 South Africa return, with a one-day international in India. 1992 Zimbabwe's first Test match. Durham becomes the first county since Glamorgan in 1921 to attain first-class status. 1993 The ICC ceases to be administered by MCC, becoming an independent organization with its own chief executive. 1994 Brian Lara of Warwickshire becomes the only player to pass 500 in a first-class innings: 501 not out v Durham.

(f) Cricket in 21th Century

2000 South Africa's captain Hansie Cronje banned from cricket for life after admitting receiving bribes from bookmakers in match-fixing scandal. Bangladesh's first Test match. County Championship split into two divisions, with promotion and relegation. The Laws of Cricket revised and rewritten. 2001 Sir Donald Bradman dies, aged 92. 2003 Twenty20 Cup, a 20-over-per-side evening tournament, inaugurated in England. 2004 Lara becomes the first man to score 400 in a Test innings, against England. Highest Score at batting position no 2 in Test match: 380 by Mathew Hayden (Australia v Zimbabwe). 2005 Fastest Test fifty in Balls: Jaquas Kallis in 24 balls (South Africa v Zimbabwe). Shane Warne becomes the first bowler to take 600 Test wickets.

2006 Highest score in ODI: Sri Lanka 443-9 vs Netherlands. Pakistan's Wasim Raja, 54, dies on the cricket field after a heart attack.

2007 Australia won the Cricket World Cup consecutive third time.

4.3 THE RAO-KUPPER MODEL

Rao and Kupper (1967) propose a modification of the Bradley-Terry (1952) model to allow the tied observations. They introduce a threshold parameter $\delta = \ln\lambda$ and suppose that if the observed difference (X_i-X_j) is less than δ then the panelist is unable to distinguish between the treatment T_i and T_j and will declare a tie. Now the probability $P\{(X_i - X_j) > \delta \mid \theta_i, \theta_j\}$ that the treatment T_i is preferred to the treatment T_j ($i \neq j$) when the treatment T_i and T_j are compared is denoted by $\psi_{i.ij}$ i.e.,

$$\psi_{i.ij} = \frac{1}{4} \int_{-(\ln\theta i - \ln\theta j)+\delta}^{\infty} \sec h^2(y/2)dy$$

$$= \frac{\theta_i}{\theta_i + \lambda\theta_j} \qquad (4.1)$$

The probability that treatment T_j is preferred to treatment T_i is denoted by $\psi_{j.ij}$ and may be obtained by swapping i with j in above equation. The probability that treatment T_j and T_j have no preference is denoted by $\psi_{o.ij}$. It is given by

$$\psi_{o.ij} = \frac{1}{4} \int_{-(\ln\theta i - \ln\theta j)-\delta}^{-(\ln\theta i - \ln\theta j)+\delta} \sec h^2(y/2)dy$$

$$= \frac{(\lambda^2 - 1)\theta_i\theta_j}{(\theta + \lambda\theta_j)(\lambda\theta_i + \theta_j)} \qquad (4.2)$$

The Rao-Kupper model is given by (4.1) and (4.2). If λ= 1 then the Rao- Kupper model yields the Bradley-Terry model.

4.3.1 Notations for the Model

The following notations are defined for the Rao- Kupper model:

$n_{i.ijk} = 1$ or 0 according as treatment T_i is preferred to treatment T_j or not in the k'th repetition of comparison.

$n_{o.ijk} = 1$ or 0 according as treatment T_i is tied with treatment T_j or not.

We note that $n_{o.ijk} + n_{i.ijk} + n_{j.ijk} = 1$ and $n_{i.ijk} = n_{i.jik}$

$n_{i.ij} = \sum_k n_{i.ijk}$ = the number of times treatment T_i is preferred to treatment T_j.

$n_{o.ij} = \sum_k n_{o.ijk}$ = the number of times treatments T_i and T_j are tied.

r_{ij} = the number of times treatments T_i is computed with treatment T_j and

$r_{ij} = n_{o.ij} + n_{i.ij} + n_{j.ij} = r_{ji}$.

The following notation is useful for further simplification of the likelihood function.

$n_{ijk} = n_{o.ijk} + n_{i.ijk},\qquad n_{jik} = n_{o.ijk} + n_{j.ijk} = r_{ij} - n_{i.ijk},$

$n_{ij} = \sum_k n_{ijk}$ = the number of times treatment T_i is preferred to treatment T_j and

the number of times treatments T_i and T_j are tied.

$n_i = \sum_{j \neq i}^{m} n_{ij}$ = the total number of times treatment T_i is preferred to any other treatment, and the number of times treatments T_i and T_j are tied.

$n_o = \sum_{i<j}^{m} n_{o.ij}$ = the total number of times treatments T_i and T_j are tied.

Through the use of (1), (2) and the constraint $\psi_{i.ij} + \psi_{o.ij} + \psi_{j.ij} = 1$, he proposes the following model:

$$\psi_{h.ij} = \frac{\theta_h}{\theta_i + \theta_j + v\sqrt{\theta_i \theta_j}}, \qquad h=i,j \text{ and } i \neq j \tag{4.8}$$

$$\psi_{o.ij} = \frac{v\sqrt{\theta_i \theta_j}}{\theta_i + \theta_j + v\sqrt{\theta_i \theta_j}}, \tag{4.9}$$

which is the Davidson model for paired comparison with ties. If V=0 then the model yields the Bradley-Terry model. Davidson uses the method of maximum likelihood to estimate the parameters and likelihood ratio statistic for testing the equality of parameters.

4.4.1 Notations for the Model

The notation to describe the data and the likelihood function for the model are presented in this Section. We use the same notation as for the Rao-Kupper model except for the following.

$s_{ij} = n_{i.ij} + n_{o.ij} / 2$, where $n_{i.ij}$ and $n_{o.ij}$ have been defined above,

$s_i = \sum_{j=1}^{m} s_{ij} =$ the total number of times treatment T_i is preferred to any other treatment plus half of the total number of times treatment T_i and treatment T_j are tied.

4.4.2 Likelihood Function for the Parameters of the Model

Now the probability of the observed result in the K'th repetition of the pair (T_i, T_j) according to the Davidson model can be presented as:

$$P_{ijk}=\left(\frac{\nu\sqrt{\theta_i\theta_j}}{\theta_i+\theta_j+\nu\sqrt{\theta_i\theta_j}}\right)^{n_{o.ijk}}\left(\frac{\theta_i}{\theta_i+\theta_j+\nu\sqrt{\theta_i\theta_j}}\right)^{n_{ijk}}\left(\frac{\theta_j}{\theta_i+\theta_j+\nu\sqrt{\theta_i\theta_j}}\right)^{n_{jik}}$$

where $0\le\theta_i\le 1$ (i= 1, 2, …, m), $\sum_{i=1}^{m}\theta_i=1$, $\nu>0$.

Hence the likelihood function of the observed outcome **x** [where **x** represents the data $\{n_{i.ij}, n_{j.ij}, n_{o.ij}\}$] of the experiment is

$$l(\mathrm{x};\theta_1,\ldots,\theta_m,\nu)=\prod_{i(<j)=1}^{m}\prod_{k=1}^{r_{ij}}P_{ijk}$$

$$l(\mathrm{x};\theta_1,\ldots,\theta_m,\nu)=\nu^{n_{o..}}\prod_{i(<j)=1}^{m}\frac{K_{ij}}{(\theta_i+\theta_j+\nu\sqrt{\theta_i\theta_j})^{r_{ij}}}\prod_{i=1}^{m}\theta_i^{s_i}, \tag{4.10}$$

where $0\le\theta_i\le 1$ (i= 1, 2, …, m), $\sum_{i=1}^{m}\theta_i=1$, $\nu>0$ is the tied parameter, $\theta_1,\ldots,\theta_m$ are the treatment parameters, s_i is defined above, r_{ij} is mentioned in Section 2 and $K_{ij}=r_{ij}/(n_{o.ij}!\,n_{i.ij}!\,n_{j.ij}!)$.

4.4.3 The Posterior Distribution for the Parameters of the Model

The (joint) posterior distribution of the parameters $\theta_1,\theta_2,\ldots\ldots..\theta_m$ and λ using likelihood function (4.10) and prior distribution (4.4) is:

$$p(\theta_1,\theta_2,\ldots..\theta_m;\lambda|x)=\prod_{k=1}^{\mathrm{r_{ij}}}\prod_{i<j=1}^{\mathrm{m}}\mathrm{P_{ijk}}$$

$$p(\theta_1,\theta_2,\ldots..\theta_m;\lambda|x)\propto\nu^{n_{o..}}\prod_{i(<j)=1}^{m}\frac{K_{ij}}{(\theta_i+\theta_j+\nu\sqrt{\theta_i\theta_j})^{r_{ij}}}\prod_{i=1}^{m}\theta_i^{s_i}, \tag{4.11}$$

4.5 RANKING OF TOP CRICKET TEAMS

We have considered the following top seven cricket teams for ranking through the paired comparison models using Bayesian approach:

1. Australia 2. South Africa 3. Pakistan 4. India

5. New Zealand 6. Sri Lanka 7. England

Here for convenience we suppose the following parameters for teams:

Australia (AU) $=\theta_1$, South Africa (SA) $=\theta_2$, Pakistan (PA) $=\theta_3$, India (IN) $=\theta_4$

New Zealand (NZ) $=\theta_5$, Sri Lanka (SL) $=\theta_6$, England (EN) $=\theta_7$

The posterior modes and means of the parameters are considered for the ranking of the teams. The ranking for ODI and test matches are presented separately for different time period. These ranking are also compared with the ICC ranking.

4.5.1 RANKING USING POSTERIOR MODES

The following equations are derived to find the posterior modes of the parameters for the Rao-Kupper model:

Now the Posterior distribution for seven teams using uniform prior is:

$$p(\theta_1,\theta_2,\ldots\theta_7|x)=\prod_{k=1}^{21}\prod_{i<j=1}^{7}P_{ijk}$$

$$=\prod_{k=1}^{21}\prod_{i<j=1}^{7}(\lambda^2-1)^{n_{0ijk}}\left(\frac{\theta_i}{\theta_i+\lambda\theta_j}\right)^{n_{ijk}}\left(\frac{\theta_j}{\theta_j+\lambda\theta_i}\right)^{n_{jik}}$$

$$=\frac{(\lambda^2-1)^{n_0}\prod_{i<j=1}^{7}K_{ij}\prod_{i=1}^{7}\theta_i^{n_i}}{\prod_{i\neq j}^{7}(\theta_i+\lambda\theta_j)^{n_{ij}}}$$

Now taking natural log on both sides

$$\text{Log}\,p(.|x)=\log(\lambda^2-1)^{n_0}+\sum_{j=1}\log K_{ij}+\sum_{i=1}^{7}\log\theta_i^{n_i}-\sum_{i\neq j}\log(\theta_i+\lambda\theta_j)^{n_{ij}}$$

Now partially differentiating w.r.t $\theta_1,\ldots,\theta_7,\lambda$ and putting equal to zero, we get

$$\frac{n_i}{\theta_i}-\sum_{i\neq j}^{7}\frac{n_{ij}}{(\theta_i+\lambda\theta_j)}-\sum_{i\neq j}^{7}\frac{\lambda n_{ij}}{(\theta_i+\lambda\theta_j)}=0, \quad \text{for } i=1,2,\ldots 7 \qquad (4.12)$$

$$\frac{2n_o\lambda}{\lambda^2-1}-\sum_{i\neq j}\frac{\theta_i n_{ij}}{\theta_i+\lambda\theta_j}=0 \qquad (4.13)$$

The following equations are derived to find the posterior modes of the parameters for the Davidson model:

The Posterior distribution for seven teams using uniform prior is:

$$p(\theta_1,\theta_2,\ldots\theta_7|x)=\prod_{k=1}^{21}\prod_{i<j=1}^{7}P_{ijk}$$

$$=v^{n_o}\prod_{i(<j)=1}^{m}\frac{K_{ij}}{(\theta_i+\theta_j+v\sqrt{\theta_i\theta_j})^{r_{ij}}}\prod_{i=1}^{m}\theta_i^{s_i},$$

taking natural log on both sides

$$Logp(.\,|\,x)=n_o\log v+\sum_{j=1}\log K_{ij}+\sum_{i=1}^{7}s_i\log\theta_i-\sum_{i\neq j}r_{ij}(\theta_i+\theta_j+v\sqrt{\theta_i\theta_j})$$

Now partially differentiating w.r.t $\theta_1,\ldots,\theta_7,v$ and putting equal to zero, we get

$$\frac{s_i}{\theta_i}-\sum_{i\neq j}^{7}\frac{r_{ij}}{\theta_i+\theta_j+v\sqrt{\theta_i\theta_j}}(1+\frac{v\theta_j}{2\sqrt{\theta_i\theta_j}})-\sum_{i\neq j}^{7}\frac{r_{ij}}{\theta_i+\theta_j+v\sqrt{\theta_i\theta_j}}(1+\frac{v\theta_i}{2\sqrt{\theta_i\theta_j}})=0, \qquad (4.14)$$

for $i=1,2,\ldots 7$

$$\frac{n_o}{v}-\sum_{i\neq j}^{7}\frac{r_{ij}(\sqrt{\theta_i\theta_j})}{\theta_i+\theta_j+v\sqrt{\theta_i\theta_j}}=0, \qquad (4.15)$$

The results obtained from the Rao-Kupper model and Davidson model are presented in the following Tables. The posterior modes obtained from both the models using data at given in Tables 4(a), 4(b) , 4(c) and 4(d) are presented in Tables 4.1, 4.2, 4.3, and 4.4 respectively. A program is designed in SAS package given in (Appendix A) to obtain the posterior mode from both the models.

It is to be noted that:

$n_{i.ij} = \sum_k n_{i.ijk}$ = the number of times team 'i' beats to team 'j'.

$n_{o.ij} = \sum_k n_{o.ijk}$ = the number of times the matches between team 'i' and team 'j' are tied.

r_{ij} = total number of matches played between team 'i' and team 'j'.

$n_{ij} = \sum_k n_{ijk}$ = the number of times team 'i' beats to team 'j' and the number of times the matches between team 'i' and team 'j' are tied.

Table 4(a) Data for One Day Internationals (1971 to December2006) of Top Seven Cricket Teams

Pairs	$n_{i.ij}$	$n_{j.ij}$	$n_{o.ij}$	n_{ij}	n_{ji}	r_{ij}
(θ_1, θ_2)	37	28	3	40	31	68
(θ_1, θ_3)	43	27	4	47	31	74
(θ_1, θ_4)	50	27	4	54	31	81
(θ_1, θ_5)	70	27	3	73	30	100
(θ_1, θ_6)	41	19	2	43	21	62
(θ_1, θ_7)	47	34	4	51	38	85
(θ_2, θ_3)	30	13	1	31	14	44
(θ_2, θ_4)	30	18	2	32	20	50
(θ_2, θ_5)	27	14	4	31	18	45
(θ_2, θ_6)	20	21	2	22	23	43
(θ_2, θ_7)	21	11	2	23	13	34

(θ_3, θ_4)	64	40	4	68	44	108
(θ_3, θ_5)	47	28	2	49	30	77
(θ_3, θ_6)	64	38	4	68	42	106
(θ_3, θ_7)	24	33	1	25	34	58
(θ_4, θ_5)	36	35	4	40	39	75
(θ_4, θ_6)	47	35	7	54	42	89
(θ_4, θ_7)	29	26	2	31	28	57
(θ_5, θ_6)	32	25	3	35	28	60
(θ_5, θ_7)	25	25	4	29	29	54
(θ_6, θ_7)	18	19	0	18	19	37
Total	802	543	62	864	605	1407

Table 4.1 Posterior Mode for Overall One Day Internationals Matches

	Using Models		
Parameters	**Rao-Kupper**	**Davidson**	**Team-Ranking**
$\theta_1 = AU$	0.220	0.223	1
$\theta_2 = SA$	0.184	0.187	2
$\theta_3 = PA$	0.146	0.145	3
$\theta_4 = IN$	0.115	0.114	5
$\theta_5 = NZ$	0.105	0.103	6
$\theta_6 = SL$	0.100	0.099	7
$\theta_7 = EN$	0.131	0.130	4

V,λ	1.096	0.094	-

Table 4(b) Data for Overall Test Matches (from beginning to December 2006) of Top Seven Cricket Teams

Pairs	$n_{i.ij}$	$n_{j.ij}$	$n_{o.ij}$	n_{ij}	n_{ji}	r_{ij}
(θ_1, θ_2)	45	15	18	63	33	78
(θ_1, θ_3)	24	11	17	41	28	52
(θ_1, θ_4)	32	15	21	53	36	68
(θ_1, θ_5)	22	7	17	39	24	46
(θ_1, θ_6)	11	1	6	17	7	18
(θ_1, θ_7)	126	97	88	214	185	311
(θ_2, θ_3)	5	2	4	9	6	11
(θ_2, θ_4)	7	3	6	13	9	16
(θ_2, θ_5)	18	4	11	29	15	33
(θ_2, θ_6)	8	4	5	13	9	17
(θ_2, θ_7)	26	54	50	76	104	130
(θ_3, θ_4)	12	8	36	48	44	63
(θ_3, θ_5)	21	6	18	39	24	45
(θ_3, θ_6)	15	7	10	25	17	32
(θ_3, θ_7)	12	18	36	48	54	66
(θ_4, θ_5)	14	9	21	35	30	44
(θ_4, θ_6)	10	3	13	23	16	26

(θ_4, θ_7)	17	34	43	60	77	94
(θ_5, θ_6)	8	4	10	18	14	22
(θ_5, θ_7)	7	41	40	47	81	88
(θ_6, θ_7)	5	8	5	10	13	18
Total	445	351	475	920	826	1271

Table 4.2 Posterior Mode for Overall Test Matches

	Using Models		
Parameters	**Rao-Kupper**	**Davidson**	**Team-Ranking**
$\theta_1 = AU$	0.254	0.301	1
$\theta_2 = SA$	0.129	0.120	4
$\theta_3 = PA$	0.146	0.141	3
$\theta_4 = IN$	0.122	0.110	5
$\theta_5 = NZ$	0.082	0.060	6
$\theta_6 = SL$	0.072	0.053	7
$\theta_7 = EN$	0.195	0.216	2
V,λ	2.320	1.283	-

Table 4(c) Data for One Day Internationals of Top Seven Cricket Teams from 1999 to 2003

Pairs	$n_{i.ij}$	$n_{j.ij}$	$n_{o.ij}$	n_{ij}	n_{ji}	r_{ij}
(θ_1, θ_2)	11	5	3	14	8	19
(θ_1, θ_3)	12	5	1	13	6	18
(θ_1, θ_4)	15	4	0	15	4	19
(θ_1, θ_5)	10	5	1	11	6	16
(θ_1, θ_6)	11	4	0	11	4	15
(θ_1, θ_7)	15	2	0	15	2	17
(θ_2, θ_3)	14	6	0	14	6	20
(θ_2, θ_4)	10	8	1	11	9	19
(θ_2, θ_5)	15	4	3	18	7	22
(θ_2, θ_6)	10	8	1	11	9	19
(θ_2, θ_7)	6	4	0	6	4	10
(θ_3, θ_4)	10	5	0	10	5	15
(θ_3, θ_5)	17	6	0	17	6	23
(θ_3, θ_6)	14	13	1	15	14	28
(θ_3, θ_7)	7	5	0	7	5	12
(θ_4, θ_5)	10	14	2	12	16	26
(θ_4, θ_6)	10	7	2	12	9	19
(θ_4, θ_7)	10	4	1	11	5	15
(θ_5, θ_6)	4	11	0	4	11	15

(θ_5, θ_7)	3	2	0	3	2	5
(θ_6, θ_7)	8	9	0	8	9	17
Total	222	131	16	238	147	369

Table 4.3 Posterior Mode for One Day Internationals Matches from 1999 to 2003

	Using Models		
Parameters	**Rao-Kupper**	**Davidson**	**Team-Ranking**
$\theta_1 = AU$	0.314	0.325	1
$\theta_2 = SA$	0.178	0.178	2
$\theta_3 = PA$	0.137	0.134	3
$\theta_4 = IN$	0.101	0.099	5
$\theta_5 = NZ$	0.082	0.078	6
$\theta_6 = SL$	0.113	0.111	4
$\theta_7 = EN$	0.075	0.074	7
V,λ	1.101	0.096	-

Table 4(d) Data for Test Matches of top Seven Cricket Teams from 1999 to 2003

Pairs	$n_{i.ij}$	$n_{j.ij}$	$n_{o.ij}$	n_{ij}	n_{ji}	r_{ij}
(θ_1, θ_2)	5	1	0	5	1	6
(θ_1, θ_3)	6	0	0	6	0	6

(θ_1, θ_4)	5	3	1	6	4	9
(θ_1, θ_5)	3	0	3	6	3	6
(θ_1, θ_6)	0	1	2	2	3	3
(θ_1, θ_7)	9	2	0	9	2	11
(θ_2, θ_3)	2	1	1	3	2	4
(θ_2, θ_4)	3	0	1	4	1	4
(θ_2, θ_5)	3	0	3	6	3	6
(θ_2, θ_6)	5	1	2	7	3	8
(θ_2, θ_7)	4	3	3	7	6	10
(θ_3, θ_4)	2	1	0	2	1	3
(θ_3, θ_5)	3	1	2	5	3	6
(θ_3, θ_6)	6	3	2	8	5	11
(θ_3, θ_7)	1	2	2	3	4	5
(θ_4, θ_5)	1	2	5	6	7	8
(θ_4, θ_6)	1	2	1	2	3	4
(θ_4, θ_7)	2	1	4	6	5	7
(θ_5, θ_6)	0	0	2	2	2	2
(θ_5, θ_7)	3	2	2	5	4	7
(θ_6, θ_7)	2	4	3	5	7	9
Total	66	30	39	105	69	135

Table 4.4 Posterior Mode for Test Matches

From 1999 to 2003

	Using Models		
Parameters	**Rao-Kupper**	**Davidson**	**Team-Ranking**
$\theta_1 = AU$	0.457	0.383	1
$\theta_2 = SA$	0.182	0.174	2
$\theta_3 = PA$	0.085	0.101	3
$\theta_4 = IN$	0.067	0.082	6
$\theta_5 = NZ$	0.068	0.090	5
$\theta_6 = SL$	0.061	0.078	7
$\theta_7 = EN$	0.801	0.093	4
V,λ	0.919	1.980	-

We use posterior mode of the Rao-Kupper and Davidson Models for ranking of top Seven Cricket Teams. Both the models give us the same ranking of the Teams. So we conclude here that either of the models can be used for ranking.

According to the results of overall one day international matches Australia, South Africa and Pakistan are in 1^{st} ,2^{nd} and 3^{rd} positions respectively .Also England , India, New Zealand and Sri Lanka are in 4^{rth} , 5^{th} , 6^{th} and 7^{th} positions respectively. Furthermore, we examine the data from 1999 to 2003 for these Teams. We observe from the results that ranking remains the same for first three, fifth and six positions. Here changes occur between seventh and fourth positions, England loses its fourth position and Sri Lanka gets fourth position.

The results of overall Test Matches results for these Top Seven Teams show that Australia, England and Pakistan are in first three positions respectively. The results from 1999 to 2003 show that Australia and Pakistan maintains its 1st and 3rd positions whereas England loses its second position. Now England is in fourth position. The Ranking of India and New Zealand also Changes. But Sri Lanka can't change its seventh position.

According to these findings, we can conclude that Australia, South Africa and Pakistan are best three Cricket Teams of the world.

4.5.2 RANKING USING POSTERIOR MEANS

The results obtained from the Rao-Kupper model are presented in the following Tables. The posterior means obtained from the model using a uniform prior and the data given in Tables 4(a), 4(b), 4(c), and 4(d) are presented in Tables 4.5, 4.6, 4.7 and 4.8 respectively. Programs are designed in C++ Language given in (Appendix B) and in SAS package given in (Appendix E) to obtain the Posterior means.

Table 4.5 Posterior Mean For Overall ODI Matches

Parameters	Posterior Means by Quadrature Method	Posterior Means using Gibbs Sampling	Team Ranking
$\theta_1 = AU$	0.226856	0.190196	1
$\theta_2 = SA$	0.189690	0.161496	2
$\theta_3 = PA$	0.144115	0.128744	3
$\theta_4 = IN$	0.115069	0.100846	5
$\theta_5 = NZ$	0.100144	0.093684	6
$\theta_6 = SL$	0.098432	0.091340	7
$\theta_7 = EN$	0.125694	0.116375	4

V	1.101190	1.105587	

Table 4.6 Posterior Mean For Overall Test Matches

Parameters	Posterior Means by Quadrature Method	Posterior Means using Gibbs Sampling	Team Ranking
$\theta_1 = AU$	0.248978	0.207223	1
$\theta_2 = SA$	0.125480	0.103009	5
$\theta_3 = PA$	0.146775	0.125368	3
$\theta_4 = IN$	0.126399	0.106285	4
$\theta_5 = NZ$	0.087150	0.075044	6
$\theta_6 = SL$	0.077496	0.072286	7
$\theta_7 = EN$	0.187722	0.155463	2
V	2.405050	2.386849	

Here we initially use Gibbs Sampling technique to find the Posterior means for the parameters of the Rao-Kupper Model. We examine that both the techniques give us the same ranking result, therefore for further estimation work we use Quadrature method.

Table 4.7 Posterior Mean For ODI Matches From 1999 To 2003

Parameters	Posterior Means by Quadrature Method	Team Ranking
$\theta_1 = AU$	0.306130	1
$\theta_2 = SA$	0.174978	2
$\theta_3 = PA$	0.139658	3
$\theta_4 = IN$	0.098267	5
$\theta_5 = NZ$	0.073943	6
$\theta_6 = SL$	0.112875	4
$\theta_7 = EN$	0.073573	7
V	1.107490	

Table 4.8 Posterior Mean for Test Matches from 1999 To 2003

Parameters	Posterior Means by Quadrature Method	Team Ranking
$\theta_1 = AU$	0.344366	1
$\theta_2 = SA$	0.177096	2
$\theta_3 = PA$	0.107871	3
$\theta_4 = IN$	0.089531	6
$\theta_5 = NZ$	0.096660	5
$\theta_6 = SL$	0.085972	7
$\theta_7 = EN$	0.098503	4
V	2.018240	

We use posterior mean of Rao-Kupper Model for ranking of top Seven Cricket Teams. We have observed that posterior mean using both Quadrature Method and Gibbs Sampling give the same ranking results. According to the results of overall one day international matches Australia, South Africa and Pakistan are in 1^{st}, 2^{nd} and 3^{rd} positions respectively Also England, India, New Zealand and Sri Lanka are in 4^{th}, 5^{th}, 6^{th} and 7^{th} positions respectively. Furthermore, we examine the data from 1999 to 2003 for these Teams. We observe from the results that ranking remains the same for first three, fifth and six positions. Here changes occur between seventh and fourth positions, England loses its fourth position and Sri Lanka gets fourth position.

The results of overall Test Matches results for these Top Seven Teams show that Australia, England and Pakistan are in first three positions respectively. The results from 1999 to 2003 show that Australia and Pakistan maintains its 1^{st} and 3^{rd} positions whereas England loses its second position. Now England is in fourth position. The Ranking of India and New Zealand also Changes. But Sri Lanka does not change its seventh position.

According to these findings, we can conclude that Australia, South Africa and Pakistan are best three Cricket Teams of the world.

4.6 PREDICTIVE PROBABILITIES

Suppose we are trying to predict a random variable Z□ $p(z \mid \theta)$. The idea of Bayesian predictive inference is that, since $p(\theta \mid x)$ is the believed posterior distribution of θ, then $p(z \mid \theta)\ p(\theta \mid x)$ is the joint distribution of z and θ given x, and integrating out over θ will give the believed distribution of z given x. So the predictive density of Z given x is

$$p(z \mid x) = \int_{\theta} p(z \mid \theta) p(\theta \mid x) d\theta, \tag{4.16}$$

and the predictive probability that Z_i exceeds Z_j is given as

$$p(Z_i > Z_j | x) = \int_{\theta} p(Z_i > Z_j | \theta) p(\theta | x) d\theta \tag{4.17}$$

[See Geisser (1971, 1980) for more detail]

4.6.1 PREDICTIVE PROBABILITIES OF OVERALL ODI MATCHES

The Predictive Probabilities $P_{(ij)}$ (i<j=1,…,m) that team i would be preferred to team j in One-Day Internationals Matches in a future single comparison of these teams is shown in Table 4.9 below. The Predictive probability $P_{(12)}$ that team T_1 would be preferred to team T_2 in a future single comparison is obtained

from the following expression:

$$P_{(12)} = \frac{1}{K} \int_{\theta_1=0}^{1} \int_{\theta_2=1}^{1-\theta_1} \int_{\theta_3=1}^{1-\theta_1-\theta_2} \int_{\theta_4=1}^{1-\theta_1-\theta_2-\theta_3} \int_{\theta_5=1}^{1-\theta_1-\theta_2-\theta_3-\theta_4} \int_{\theta_6=1}^{1-\theta_1-\theta_2-\theta_3-\theta_4-\theta_5} \int_{\lambda=1}^{\infty} \mathrm{P}(\theta|\mathrm{x}) d\lambda d\theta_6 d\theta_5 d\theta_4 d\theta_3 d\theta_2 d\theta_1$$

where

$$\mathrm{P}(\theta|\mathrm{x}) = \frac{(\lambda^2-1)^{n_0} \theta_1^{n_1+1} \theta_2^{n_2} \theta_3^{n_3} \theta_4^{n_4} \theta_5^{n_5} \theta_6^{n_6} (1-\theta_1-\theta_2-\theta_3-\theta_4-\theta_5-\theta_6)^{n_7}}{(\theta_1+\lambda\theta_2) \prod_{i(<j)=1}^{7} (\theta_i+\lambda\theta_j)^{n_{ij}} (\lambda\theta_i+\theta_j)^{n_{ji}}}$$

$$\text{Here } \theta_7 = 1 - \sum_{i=1}^{6} \theta_i$$

A program is designed in C++ Language given in (Appendix C) to solve above integral. Similarly the expressions for the predictive probabilities $P_{(13)}, P_{(14)}$ and so on

$P_{(67)}$ can be derived. We also calculate the predictive probabilities $P_{(0.12)}$ that in a future match played between team T_1 and team T_2 will be tie and so on up to $P_{(0.67)}$. The Predictive Probabilities are shown in Table 4.9 below:

Table 4.9 Predictive Probabilities For Overall ODI Matches

Predictive probabilities	$P_{(ij)}$	$P_{(o.ij)}$
$p_{(12)}$	0.522437	0.042832
$p_{(13)}$	0.587283	0.029801
$p_{(14)}$	0.640968	0..078909
$p_{(15)}$	0.671236	0.078061
$p_{(16)}$	0.674700	0.074988
$p_{(17)}$	0.621097	0.053795
$p_{(23)}$	0.541635	0.031835
$p_{(24)}$	0.597008	0.085485
$p_{(25)}$	0.628488	0.085040
$p_{(26)}$	0.632207	0.081817
$p_{(27)}$	0.657641	0.021501
$p_{(34)}$	0.533597	0.039986
$p_{(35)}$	0.566079	0.039132
$p_{(36)}$	0.569976	0.039140
$p_{(37)}$	0.512173	0.041118
$p_{(45)}$	0.509022	0.042325
$p_{(46)}$	0.512883	0.042365
$p_{(47)}$	0.455386	0.040601
$p_{(56)}$	0.480228	0.043228
$p_{(57)}$	0.423085	0.040956
$p_{(67)}$	0.419281	0.040950

This Table shows e.g., $p_{(12)}$ =0.522437 that in future matches 52.2% chance that Australia will win against South Africa in any ODI match. Also $p_{(0.12)}$=0.042832 shows that in future matches played between Australia and South Africa, there will be 4.28% chance of being tie.

4.6.2 REDICTIVE PROBABILITIES OF OVERALL TEST MATCHES

The Predictive Probabilities $P_{(ij)}$ $(i<j=1,\ldots,m)$ that team i would be preferred to team j in Overall Test Matches in a future single comparison of these teams are shown in Table 4.10 below.

Table 4.10 Predictive Probabilities For Overall Test Matches

Predictive probabilities	$P_{(ij)}$	$P_{(o.ij)}$
$p_{(12)}$	0.482727	0.347462
$p_{(13)}$	0.418248	0.372104
$p_{(14)}$	0..458210	0.364550
$p_{(15)}$	0.516183	0.351734
$p_{(16)}$	0.537617	0.323290
$p_{(17)}$	0.366115	0.394360
$p_{(23)}$	0.257304	0.376921
$p_{(24)}$	0.299705	0.380791
$p_{(25)}$	0.356917	0.390477
$p_{(26)}$	0.348490	0.388927
$p_{(27)}$	0.213697	0.377357
$p_{(34)}$	0.332878	0.382858
$p_{(35)}$	0.421136	0.374863

$p_{(36)}$	0.460508	0.375593
$p_{(37)}$	0.249179	0.391538
$p_{(45)}$	0.387585	0.388156
$p_{(46)}$	0.427008	0.386908
$p_{(47)}$	0.222921	0.379983
$p_{(56)}$	0.342540	0.401854
$p_{(57)}$	0.165719	0.361894
$p_{(67)}$	0.149600	0.364344

This Table shows e.g., $p_{(12)}$=0.482727 that in future matches 48.3% chance that Australia will win against South Africa in any Test match. Also $p_{(0.12)}$=0.347462 shows that in future matches played between Australia and South Africa, there will be 34.75% chance of being tie.

4.7 Preference Probabilities

As we know that the preference probabilities are calculated by just calculating the probability of preferring team T_i to team T_j using the values of parameters, the preference probabilities for the Overall ODI Matches by using the data given in Table 4(a) and preference probabilities for the Overall Test Matches using the data in Table 4(b) are given below in Tables 4.11 & 4.12 respectively.

Table 4.11 Preference Probabilities of Overall ODI Matches

Preference probabilities	$\psi_{i.ij}$	$\psi_{j.ij}$	$\psi_{o.ij}$
$\psi_{1.12}$	0.5207	0.4328	0.0454
$\psi_{1.13}$	0.5885	0.3771	0.0439
$\psi_{1.14}$	0.6416	0.3229	0.0413
$\psi_{1.15}$	0.6730	0.3034	0.0401
$\psi_{1.16}$	0.6768	0.2932	0.0394
$\psi_{1.17}$	0.6211	0.3520	0.0429
$\psi_{2.23}$	0.5445	0.4199	0.0452
$\psi_{2.24}$	0.5995	0.3632	0.0434
$\psi_{2.25}$	0.6325	0.3424	0.0424
$\psi_{2.26}$	0.6365	0.3315	0.0418
$\psi_{2.27}$	0.5781	0.3938	0.0445
$\psi_{3.34}$	0.5320	0.4182	0.0452
$\psi_{3.35}$	0.5666	0.3962	0.0446
$\psi_{3.36}$	0.5708	0.3846	0.0442
$\psi_{3.37}$	0.5101	0.4501	0.0457
$\psi_{4.45}$	0.5108	0.4545	0.0457
$\psi_{4.46}$	0.5151	0.4424	0.0456
$\psi_{4.47}$	0.4540	0.5096	0.0456
$\psi_{5.56}$	0.4802	0.4649	0.0458
$\psi_{5.57}$	0.4197	0.5323	0.0452

$\psi_{6.67}$	0.4156	0.5445	0.0450

From the above table it is observed that the preference scheme is like as AU→ SA→ PA→ EN→ IN→ NZ→ SL. That is Australia possesses maximum preference and Sri Lanka holds minimum preference.

Table 4.12 Preference Probabilities of Overall Test Matches

Preference probabilities	$\psi_{i.ij}$	$\psi_{j.ij}$	$\psi_{o.ij}$
$\psi_{1.12}$	0.4591	0.1796	0.3613
$\psi_{1.13}$	0.4285	0.1986	0.3729
$\psi_{1.14}$	0.4730	0.1715	0.3555
$\psi_{1.15}$	0.5718	0.1222	0.3061
$\psi_{1.16}$	0.6033	0.1089	0.2879
$\psi_{1.17}$	0.3596	0.2486	0.3918
$\psi_{2.23}$	0.2758	0.3279	0.3963
$\psi_{2.24}$	0.3131	0.2896	0.3973
$\psi_{2.25}$	0.4041	0.2151	0.3809
$\psi_{2.26}$	0.4358	0.1939	0.3703
$\psi_{2.27}$	0.2219	0.3945	0.3836
$\psi_{3.34}$	0.3403	0.2648	0.3949
$\psi_{3.35}$	0.4342	0.1949	0.3709
$\psi_{3.36}$	0.4664	0.1753	0.3583
$\psi_{3.37}$	0.2440	0.3654	0.3907
$\psi_{4.45}$	0.3907	0.2246	0.3846

$\psi_{4.46}$	0.4221	0.2029	0.3751
$\psi_{4.47}$	0.2124	0.4079	0.3797
$\psi_{5.56}$	0.3293	0.2746	0.3962
$\psi_{5.57}$	0.1534	0.5062	0.3404
$\psi_{6.67}$	. 0.1373	0.5386	0.3241

From the above table it is observed that the scheme is like as AU→ EN→ PA→ SA→ IN→ NZ→ SL. That is Australia possesses maximum preference and Sri Lanka holds minimum preference.

4.8 Bayesian Hypothesis Testing

The following two hypotheses H_{ij} and $\bar{H}_{ij}$ (i<j=1, 2, 3, 4, 5, 6, 7) are compared:

$$H_{ij} : \theta_i > \theta_j ,$$

and $$\bar{H}_{ij} : \theta_j \geq \theta_i$$

The posterior probability p_{ij} for H_{ij} is $p_{ij} = p(\theta_i > \theta_j)$, and $q_{ij} = 1 - p_{ij}$ is the posterior probability for $\bar{H}_{ij}$

The decision rule used here, for accepting or rejecting the above hypothesis is; let s =min(p_{ij}, q_{ij}), if p_{ij} is small(s<0.1) then $\bar{H}_{ij}$ is accepted and if q_{ij} is small(s<0.1), H_{ij} is accepted. And if s>0.1, the decision is inconclusive (Aslam 2002)

The posterior probability p_{12} for H_{12} is obtained as:

$$p_{12} = p(\theta_1 > \theta_2) = p(\theta_1 - \theta_2 > 0) = p(\phi > 0)$$

A program is designed in C++ Language given in (Appendix D) to find the posterior probabilities for the several pair of parameters using the data given in table 4(a), table 4(b) for the overall ODI and Test matches respectively.

Table 4.13 Posterior Probabilities of Overall ODI Matches

Posterior probabilities	p_{ij}	q_{ij}
p_{12}	0.6417	0.3583
p_{13}	0.8698	0.1302
p_{14}	0.9697	0.0303
p_{15}	0.9914	0.0005
p_{16}	0.9906	0.0094
p_{17}	0.8745	0.1255
p_{23}	0.5356	0.4644
p_{24}	0.8730	0.1270
p_{25}	0.9251	0.0749
p_{26}	0.9281	0.0719
p_{27}	0.9893	0.0170
p_{34}	0.5890	0.4110
p_{35}	0.6430	0.3570
p_{36}	0.6527	0.3473
p_{37}	0.6322	0.3678
p_{45}	0.1253	0.8747

p_{46}	0.1409	0.8591
p_{47}	0.1264	0.8736
p_{56}	0.0671	0.9329
p_{57}	0.0443	0.9557
p_{67}	0.0376	0.9624

For the data set given in Table 4(a), it is found that p_{12}=0.6417 and q_{12}=0.3583, the decision is inconclusive. Similarly we obtain, p_{14}= 0.9697 and q_{14}=0.0303, so H_{14} is accepted, it means that Australia is better than India. Now p_{67}=0.0376 and q_{67}=0.9624, so we accept the hypotheses $\overline{H}_{67}$ with high probability that England is better than Sri Lanka in Overall ODI matches.

Table 4.14 Posterior Probabilities of Overall Test Matches

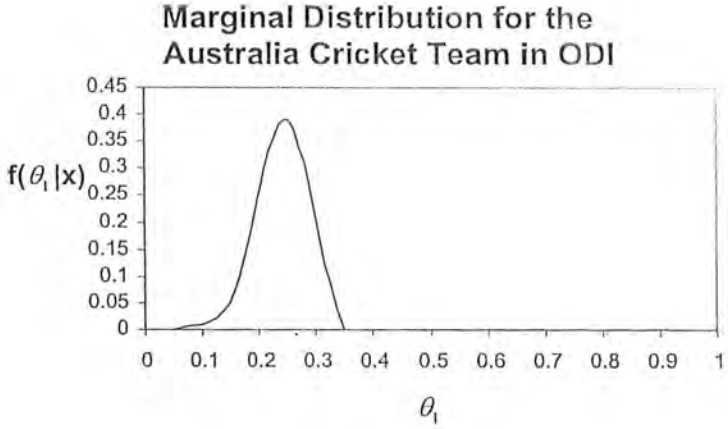
Posterior probabilities	p_{ij}	q_{ij}
p_{12}	0.9283	0.0717
p_{13}	0.8512	0.1489
p_{14}	0.9282	0.0718
p_{15}	0.9897	0.0103
p_{16}	0.9776	0.0224
p_{17}	0.4034	0.5967
p_{23}	0.2146	0.7854
p_{24}	0.2853	0.7147
p_{25}	0.6126	0.3874

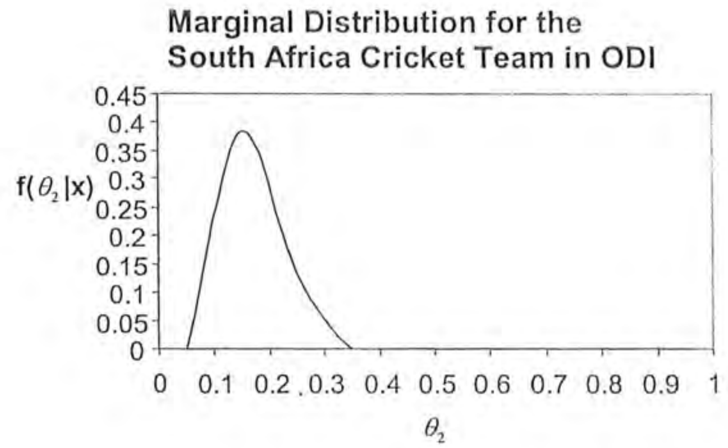
p_{26}	0.6899	0.3101
p_{27}	0.9896	0.0103
p_{34}	0.4366	0.5634
p_{35}	0.7420	0.2580
p_{36}	0.7965	0.2035
p_{37}	0.0755	0.9245
p_{45}	0.6321	0.3679
p_{46}	0.7105	0.2895
p_{47}	0.0267	0.9733
p_{56}	0.3482	0.6518
p_{57}	0.0018	0.9982
p_{67}	0.0095	0.9905

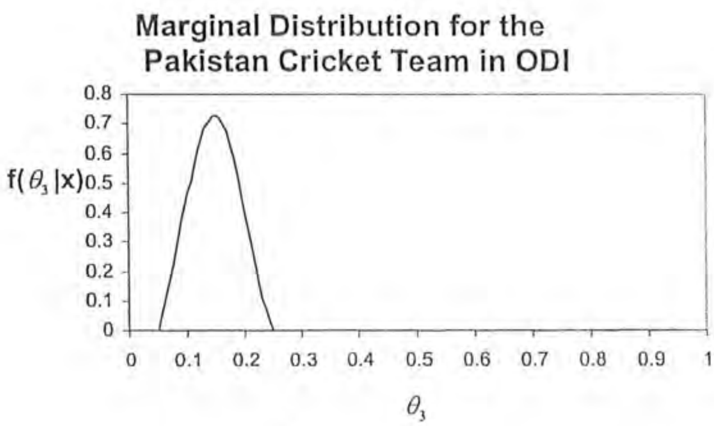
For the data set given in Table 4(b), it is found that p_{12}=0.9283 and q_{12}=0.0717, so H_{12} is accepted. Similarly we obtain, p_{14}= 0.9282 and q_{14}=0.0718, so H_{14} is accepted. Now p_{67}=0.0095 and q_{67}=0.9905, so we accept the hypotheses $\bar{H}_{67}$ with high probability that England is better than Sri Lanka in Overall Test matches.

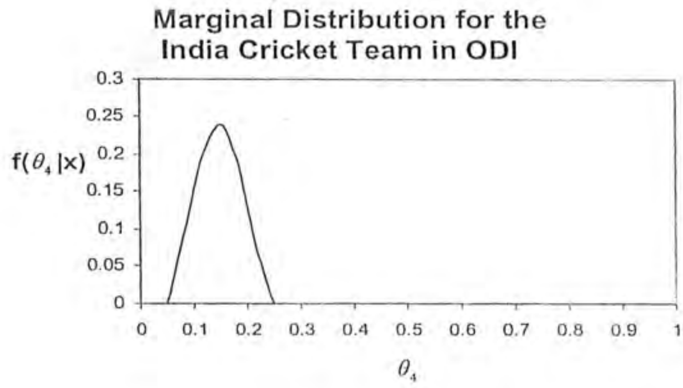
4.9 Graphs of Marginal Distributions

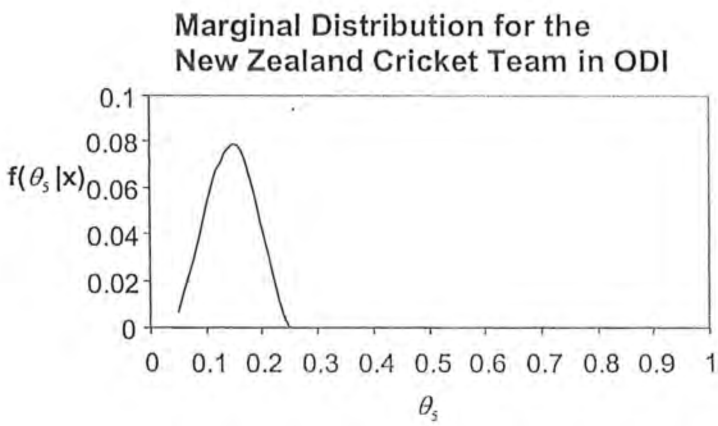
The graphs of the marginal posterior densities of the Top Seven Teams for Overall ODI and Test matches are shown in Figures 4.1, 4.2 using the data of Table 4(a), 4(b) respectively. A program is designed in C++ Language to obtain the ordinates of respective marginal densities.

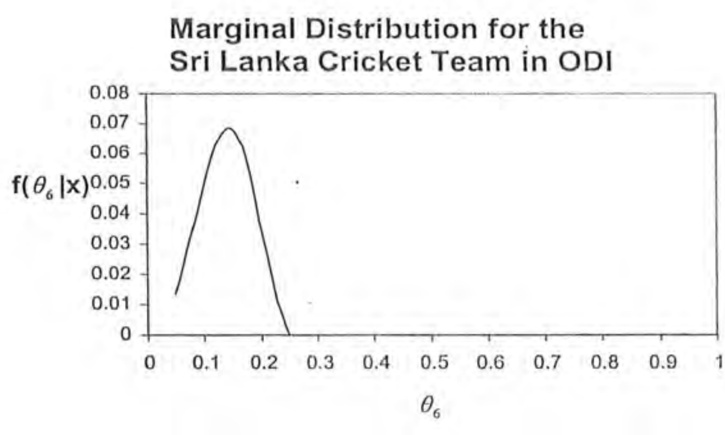

Marginal Distribution for the
Australia Cricket Team in ODI
f(θ1|x)
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
θ1


Marginal Distribution for the
South Africa Cricket Team in ODI
f(θ2|x)
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
θ2


Marginal Distribution for the
Pakistan Cricket Team in ODI
f(θ3|x)
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
θ3


Marginal Distribution for the
India Cricket Team in ODI
0.3
0.25
0.2
0.15
0.1
0.05
0
$f(\theta_4|x)$
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
θ_4


Marginal Distribution for the
New Zealand Cricket Team in ODI
0.1
0.08
0.06
0.04
0.02
0
$f(\theta_5|x)$
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
θ_5


Marginal Distribution for the
Sri Lanka Cricket Team in ODI
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
$f(\theta_6|x)$
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
θ_6

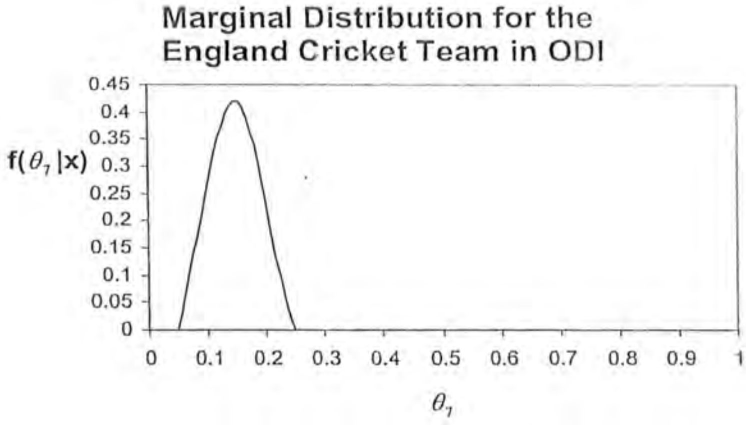
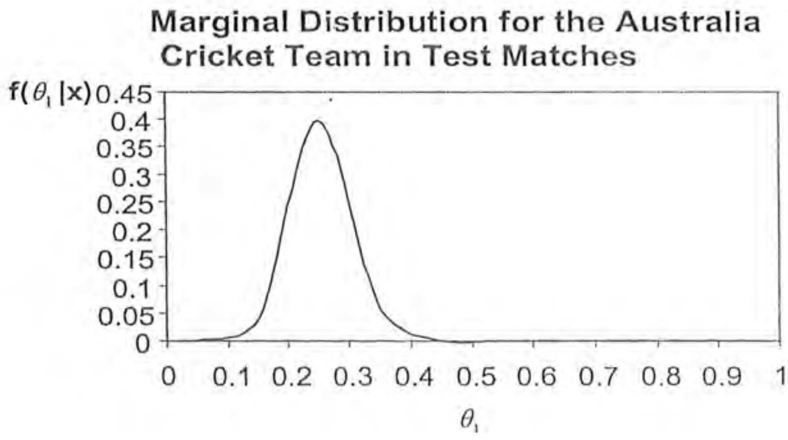
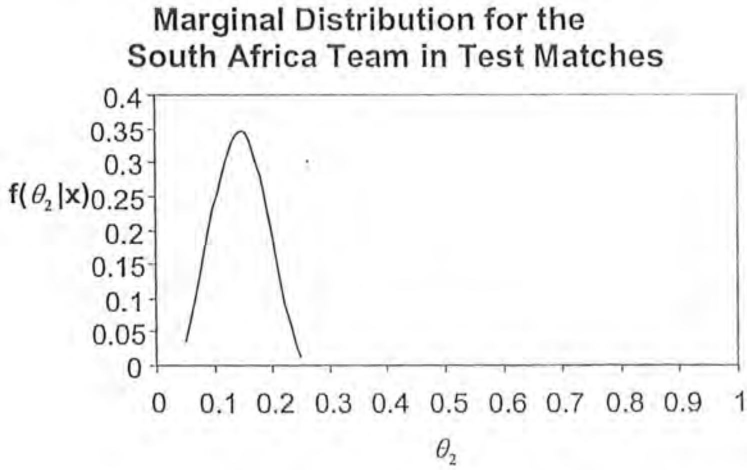
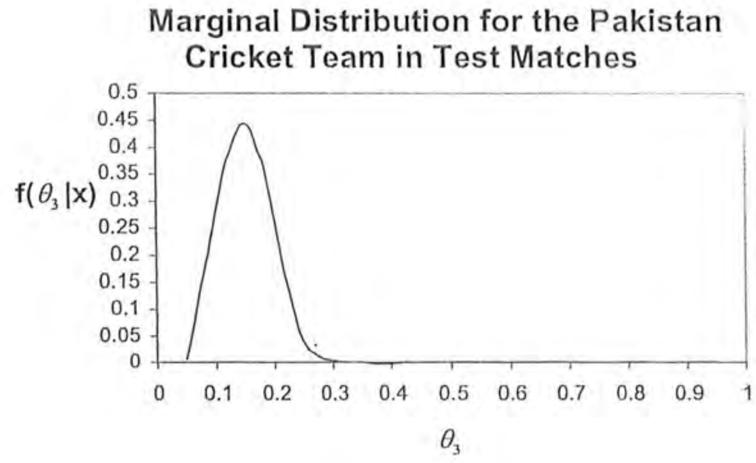
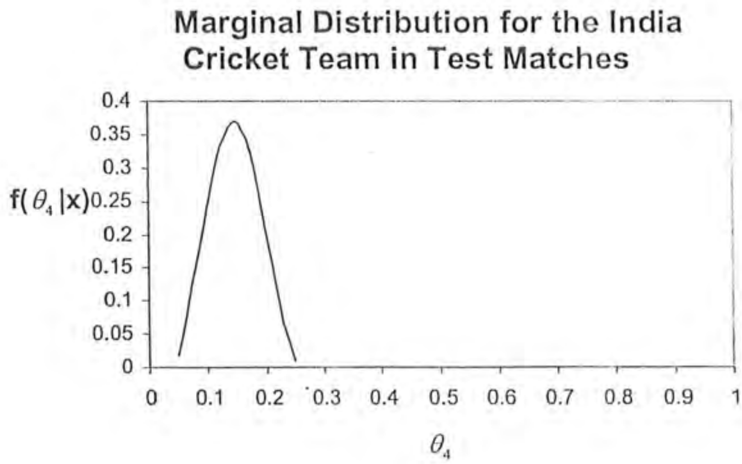

Marginal Distribution for the
England Cricket Team in ODI
0.45
0.4
0.35
f(θ7 |x) 0.3
0.25
0.2
0.15
0.1
0.05
0
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
θ7

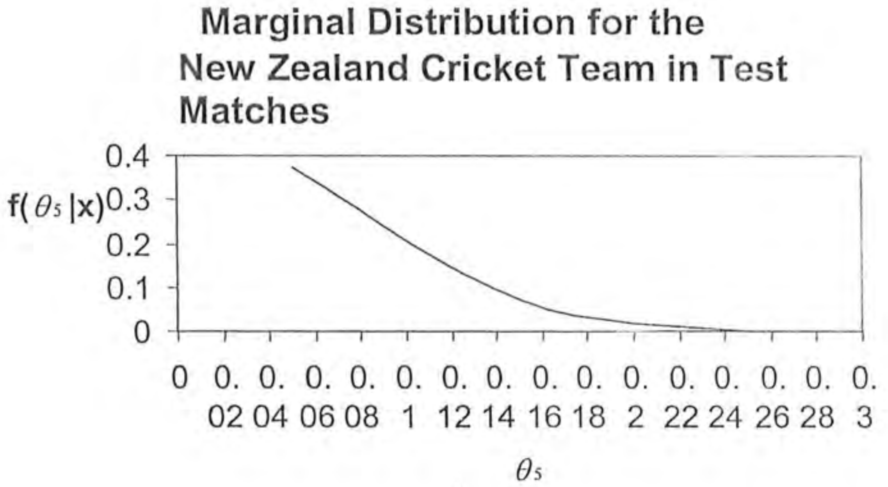
Figure 4.1

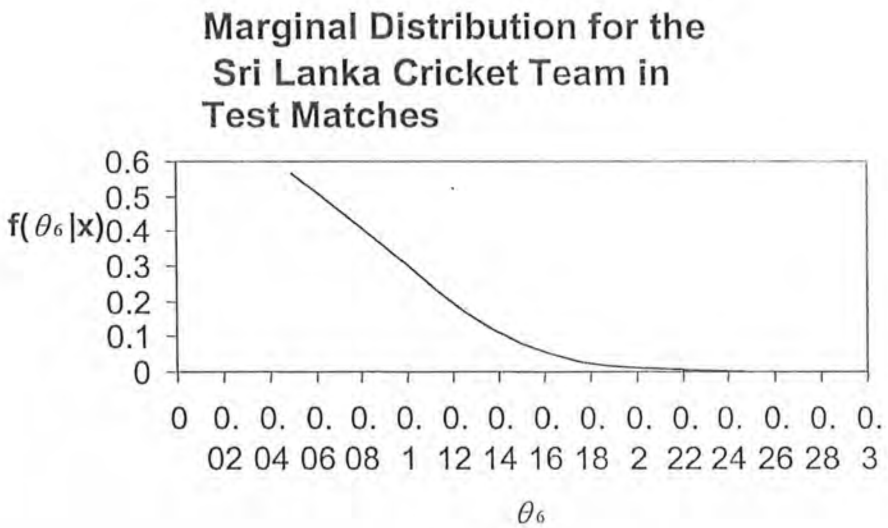

Marginal Distribution for the Australia
Cricket Team in Test Matches
f(θ1 |x) 0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
θ1


Marginal Distribution for the
South Africa Team in Test Matches
0.4
0.35
0.3
f(θ2 |x) 0.25
0.2
0.15
0.1
0.05
0
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
θ2


Marginal Distribution for the Pakistan Cricket Team in Test Matches
$f(\theta_3|x)$
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
θ_3


Marginal Distribution for the India Cricket Team in Test Matches
$f(\theta_4|x)$
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
θ_4


Marginal Distribution for the New Zealand Cricket Team in Test Matches
0.4
f(θ_5|x) 0.3
0.2
0.1
0
0 0.02 0.04 0.06 0.08 0.1 0.12 0.14 0.16 0.18 0.2 0.22 0.24 0.26 0.28 0.3
θ_5


Marginal Distribution for the Sri Lanka Cricket Team in Test Matches
0.6
0.5
f(θ_6|x) 0.4
0.3
0.2
0.1
0
0 0.02 0.04 0.06 0.08 0.1 0.12 0.14 0.16 0.18 0.2 0.22 0.24 0.26 0.28 0.3
θ_6

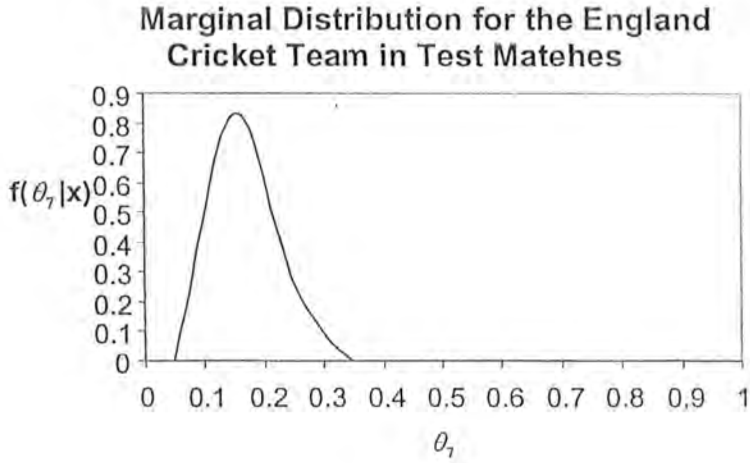


Figure 4.2

4.10 Appropriateness of the Model

For testing the appropriateness of the model or the good of fit of the model for paired comparisons, the observed numbers of preferences are compared with the expected number of preferences. If the discrepancies are small we consider the solution to be internally consistent. We employ the χ^2 statistic for testing the hypothesis that the model is true for some value of θ_o (the vector of parameter values). This is equivalent to the assertion that the observed and expected number of preferences are 'in agreement'. Glenn and David (1960), Rao and Kupper (1976) and Davidson (1970) employ the χ^2 statistic for testing the hypothesis that the observed and expected number of preferences are in agreement. Let the expected number of preferences be denoted by

$\hat{n}_{i.ij}$ = the expected number of times treatment T_i is preferred to treatment T_j.

$\hat{n}_{o.ij}$ = the expected number of times treatments T_i and T_j are tied.

The χ^2 statistic has the following form:

$$\chi^2 = \sum_{i<j}^{m} \left\{ \frac{(n_{i.ij} - \hat{n}_{i.ij})^2}{\hat{n}_{i.ij}} + \frac{(n_{j.ij} - \hat{n}_{j.ij})^2}{\hat{n}_{j.ij}} + \frac{(n_{o.ij} - \hat{n}_{o.ij})^2}{\hat{n}_{o.ij}} \right\} \quad (4.18)$$

with {m(m-2)} degrees of freedom.

Consider the following two hypotheses for the testing of Goodness of fit of Overall ODI Internationals Matches for the data given in Table 4.1(a).

H_0 : The model is true for some values $\theta = \theta_0$

H_1 : The model is not true for any value of θ

For testing the above hypotheses for possible rejection, we compare the observed number of preference with the expected number of preference and if differences are small then, model is considered consistent.

The observed and expected numbers of preference presented in Table 4.13

Table 4.13 Observed and Expected Number of Preference of overall ODI matches

Pairs	$n_{i.ij}$	$\hat{n}_{i.ij}$	$n_{j.ij}$	$\hat{n}_{j.ij}$	$n_{o.ij}$	$\hat{n}_{o.ij}$
(θ_1, θ_2)	37	35.48	28	29.43	3	3.09
(θ_1, θ_3)	43	42.84	27	27.91	4	3.25
(θ_1, θ_4)	50	51.50	27	26.16	4	3.45
(θ_1, θ_5)	70	65.66	27	30.34	3	4.01
(θ_1, θ_6)	41	41.38	19	18.18	2	2.44
(θ_1, θ_7)	47	51.43	34	29.92	4	3.64
(θ_2, θ_3)	30	23.53	13	18.48	1	1.99
(θ_2, θ_4)	30	29.67	18	18.16	2	2.17

(θ_2, θ_5)	27	27.68	14	15.41	4	1.91
(θ_2, θ_6)	20	26.94	21	14.25	2	1.80
(θ_2, θ_7)	21	19.10	11	13.39	2	1.51
(θ_3, θ_4)	64	57.96	40	45.16	4	4.88
(θ_3, θ_5)	47	43.06	28	30.51	2	3.43
(θ_3, θ_6)	64	60.55	38	40.77	4	4.69
(θ_3, θ_7)	24	29.24	33	26.11	1	2.65
(θ_4, θ_5)	36	37.49	35	34.09	4	3.43
(θ_4, θ_6)	47	45.57	35	39.37	7	4.06
(θ_4, θ_7)	29	25.35	26	29.05	2	2.60
(θ_5, θ_6)	32	29.36	25	27.90	3	2.75
(θ_5, θ_7)	25	22.81	25	28.75	4	2.44
(θ_6, θ_7)	18	15.19	19	20.15	0	1.66

It is to be noted that:

$\hat{n}_{i.ij}$ = the expected number of times team T_i is better to team T_j.

$\hat{n}_{o.ij}$ = the expected number of times matches played between team T_i and T_j are tied.

We calculate the value of Chi-square with 35 degree of freedom by using the formula given in (4.9). The value of χ^2 =29.81235 with p-value=0.716616, so there is no evidence that the model does not fit the data.

Now we apply the same test for Overall Test Matches.

Table 4.14 Observed and Expected Number of Preference of overall Test matches

Pairs	$n_{i.ij}$	$\hat{n}_{i.ij}$	$n_{j.ij}$	$\hat{n}_{j.ij}$	$n_{o.ij}$	$\hat{n}_{o.ij}$
(θ_1, θ_2)	45	35.80	15	14.01	18	28.18
(θ_1, θ_3)	24	22.28	11	10.33	17	19.39
(θ_1, θ_4)	32	32.16	15	11.66	21	24.18
(θ_1, θ_5)	22	26.30	7	5.62	17	14.18
(θ_1, θ_6)	11	10.86	1	1.96	6	5.18
(θ_1, θ_7)	126	120.63	97	87.33	88	95.03
(θ_2, θ_3)	5	3.03	2	3.61	4	4.36
(θ_2, θ_4)	7	5.01	3	4.63	6	6.36
(θ_2, θ_5)	18	13.33	4	7.10	11	12.57
(θ_2, θ_6)	8	7.41	4	3.30	5	6.30
(θ_2, θ_7)	26	28.84	54	51.29	50	49.87
(θ_3, θ_4)	12	14.14	8	10.85	36	28.53
(θ_3, θ_5)	21	19.54	6	8.77	18	16.69
(θ_3, θ_6)	15	14.92	7	5.61	10	11.47
(θ_3, θ_7)	12	16.10	18	24.11	36	29.36
(θ_4, θ_5)	14	17.19	9	9.88	21	16.92
(θ_4, θ_6)	10	10.97	3	5.27	13	9.75
(θ_4, θ_7)	17	19.97	34	38.34	43	36.69
(θ_5, θ_6)	8	7.24	4	6.04	10	8.72
(θ_5, θ_7)	7	10.52	41	44.54	40	35.53

(θ_6, θ_7)	5	3.42	8	9.70	5	5.83

We calculate the value of Chi-square with 35 degree of freedom by using the formula given in (4.9). The value of χ^2 =38.20681 p-value=0.325861, so there is no evidence that the model does not fit the data.

4.11 Latest ICC Team Rankings

The International Cricket Council (ICC) has developed an innovative method for Test and One-day International (ODI) rankings of 10 Test Playing countries and Kenya, which has only the ODI status. The ranking of other teams with ODI status aren't included, as the criteria of a minimum number of matches to be played within a certain time limit, is not met by them. The ODI rankings of teams were started by the ICC on August 01, 2003, and are based on results of all ODI's played since then.

The current cricket ratings of teams will be based on three years of results, and so all ODI's played until August 2006 will be taken into account for team rankings. In August 2006, the first year of results will be dropped from the table, so it will then cover the most recent two years results. With this, the ODI rankings of teams would change overnight, even without any new ODI being played.

The rating of each team is obtained by dividing total points scored and total matches played up to a particular time period, and the nearest whole number is considered the rating of the team that decides its ranking. The points earned by teams are calculated by a mathematical formula, which depends on two factors, the result and the rating of the opponent against whom the result was achieved. Higher rating teams beating lower rating teams get less points and vice-versa. Currently Australia is leading in both ODI and Test rankings.

4.11.1 Principles Underpinning the Rankings for ODI Matches

- Based on individual matches
- All ODIs treated equally
- No account taken of venue
- No account taken of margin of victory
- Recent results carry more weight
- Account taken of strength of opposition
- Transparency

4.11.2 Formula For Calculating Points

Case 1- Gap in rating less than 40 points

- If team wins, it scores 50 points more than its opponent's rating.
- If team loses, it scores 50 points less than its opponent's rating.
- If team tie, it scores its opponent's rating.

Case 2- Gap in ratings 40 points or more

If the stronger team wins, it scores 10 points more than its own rating while the weaker team scores 10 points less than its own rating. If the weaker team wins, it scores 90 points more than its own rating while the stronger team scores 90 points less than its own rating. If the match is tied, the stronger team scores 40 points less than its own rating and the weaker team scores 40 points more than its own rating. To work out the 'rating' divide the points scored by matches played.

Example

The first ODI in November 2002 was between India and West Indies. Their launch positions were as follows:

Team	Matches Played	Points	Rating
India	31	3301	106
West Indies	22	2074	94

The difference between their ratings was only 12, so Case 1 applies. Also the sum of their ratings was 106+94=200, so there 200 rating points available from the match whatever the result. If India had won, they would have scored 94+50=144 points and the West Indies would have scored 106-50=56.

The updated table would have then shown India with 3301+144=3445 points from 32 matches, giving them a new rating of 3445/32=108 to the nearest whole number. The West Indies would have 2074+56=2130 points from 23 matches and a new rating of 93.

As the West Indies won, they scored 106+50=156 points and India scored 94-50=44 points. The new table then showed India with 3345 points from 32 matches and a rating of 105 with West Indies moving to 2230 from 23 matches and a rating of 97.

{See also the Principles Underpinning the Rankings for Test Matches in Appendix F}

4.12 ICC Ranking VS Bayesian Ranking

Here we give below in Table 4.14 & 4.15, the Latest ICC Team Ranking in December (2006) of ODI and Test Matches.

Table 4.15 ICC ODI Ranking			
Team	Matches	Points	Rating
AU	33	4335	131
SA	32	4047	126
NZ	22	2477	113
PA	30	3373	112
SL	37	4011	108
IN	41	4355	106
EN	26	2573	99

Table 4.16 ICC Test Ranking			
Team	Matches	Points	Rating
AU	37	4793	130
SA	41	4864	119
NZ	34	3800	112
PA	34	3780	111
SL	36	3686	102
IN	34	3182	94
EN	28	2602	93

Now for comparison purposes, we collect recent two years data for ODI and recent three years data for Test Matches.

Table 4(e) Data for One Day Internationals of top Seven Cricket Teams of recent two years

Pairs	$n_{i.ij}$	$n_{j.ij}$	$n_{o.ij}$	n_{ij}	n_{ji}	r_{ij}
(θ_1, θ_2)	5	4	0	5	4	9
(θ_1, θ_3)	4	1	0	4	1	5
(θ_1, θ_4)	2	0	1	3	1	3
(θ_1, θ_5)	8	1	0	8	1	9
(θ_1, θ_6)	5	2	0	5	2	7
(θ_1, θ_7)	4	2	2	6	4	10
(θ_2, θ_3)	1	0	0	1	0	1

(θ_2, θ_4)	6	2	0	6	2	8
(θ_2, θ_5)	4	1	1	5	2	6
(θ_2, θ_6)	3	2	0	3	2	5
(θ_2, θ_7)	4	1	2	6	3	7
(θ_3, θ_4)	6	6	0	6	6	12
(θ_3, θ_5)	0	1	0	0	1	1
(θ_3, θ_6)	3	0	1	4	1	4
(θ_3, θ_7)	5	4	1	6	5	10
(θ_4, θ_5)	1	2	0	1	2	3
(θ_4, θ_6)	6	4	1	7	5	11
(θ_4, θ_7)	6	1	0	6	1	7
(θ_5, θ_6)	4	3	0	4	3	7
(θ_5, θ_7)	0	0	0	0	0	0
(θ_6, θ_7)	5	0	0	5	0	5
Total	82	36	9	91	46	127

Table 4(f) Data for Test Matches of top Seven Cricket Teams of recent three years

Pairs	$n_{i.ij}$	$n_{j.ij}$	$n_{o.ij}$	n_{ij}	n_{ji}	r_{ij}
(θ_1, θ_2)	5	0	1	5	1	6
(θ_1, θ_3)	3	0	0	3	0	3

(θ_1, θ_4)	2	1	2	3	3	5
(θ_1, θ_5)	4	0	1	4	1	5
(θ_1, θ_6)	4	0	1	4	1	5
(θ_1, θ_7)	5	2	2	7	4	9
(θ_2, θ_3)	0	0	0	0	0	0
(θ_2, θ_4)	1	2	1	3	3	4
(θ_2, θ_5)	3	1	2	4	3	6
(θ_2, θ_6)	0	3	1	3	4	4
(θ_2, θ_7)	1	2	2	3	4	5
(θ_3, θ_4)	3	3	3	6	6	9
(θ_3, θ_5)	0	0	0	0	0	0
(θ_3, θ_6)	2	1	1	3	2	4
(θ_3, θ_7)	2	3	2	5	5	7
(θ_4, θ_5)	0	0	0	0	0	0
(θ_4, θ_6)	2	0	1	2	1	3
(θ_4, θ_7)	1	1	1	2	2	3
(θ_5, θ_6)	2	1	1	3	2	4
(θ_5, θ_7)	0	3	0	3	3	3
(θ_6, θ_7)	1	1	1	2	2	3
Total	40	24	23	65	47	87

Table 4.17 Bayesian ODI Ranking			
Team	Posterior Mean	Posterior Mode	Ranking
$\theta_1 = AU$	0.2717	0.2875	1
$\theta_2 = SA$	0.2220	0.2296	2
$\theta_3 = PA$	0.1098	0.1021	4
$\theta_4 = IN$	0.1226	0.1182	3
$\theta_5 = NZ$	0.1040	0.0993	6
$\theta_6 = SL$	0.1047	0.0995	5
$\theta_7 = EN$	0.0652	0.0638	7

Table 4.18 Bayesian Test Ranking			
Team	Posterior Mean	Posterior Mode	Ranking
$\theta_1 = AU$	0.3404	0.3775	1
$\theta_2 = SA$	0.0902	0.0841	6
$\theta_3 = PA$	0.1190	0.1118	4
$\theta_4 = IN$	0.1472	0.1417	2
$\theta_5 = NZ$	0.0951	0.0864	5
$\theta_6 = SL$	0.0838	0.0776	7
$\theta_7 = EN$	0.1243	0.1208	3

The above comparison shows that the Bayesian Ranking almost equal to ICC Ranking for ODI Matches. But on the other hand the Bayesian Ranking is differing from ICC Ranking for Test Matches. The reason is that we take the data of top seven teams only. If we take all the teams, the Bayesian Ranking may be equal to ICC Ranking.

Chapter 5

CONCLUSION AND FURTHER RESEARCH

The present study comprises the ranking of top seven cricket teams (Australia, South Africa, Pakistan, India, New Zealand, Sri Lanka and England) using the Rao-Kupper (1967) and Davidson (1970) models for paired comparisons. Aslam (2005) presents Bayesian analysis of paired comparison data through these two models. He observes that both models give very similar results in application. We use these two models initially, as the results are similar of both models and also observe that the ranking of seven teams under these models are same, so further study we use only the Rao-Kupper model. The complete discussion on these models is detailed in chapter#4.

Aslam (2005) presents Bayesian analysis on these two models by using three parameters. Here, we ensue to perform Bayesian analysis using seven parameters (top seven cricket teams). We use only Uniform prior for the parameters of the models to obtain our desired results.

We consider Bayesian estimation and Bayesian testing of hypothesis about the comparison of the parameters. The posterior probabilities of the hypotheses for the comparison of two parameters have been calculated and decisions have been made about the hypotheses according to these probabilities. The posterior estimates in terms of posterior means and joint modes are determined. The predictive probabilities that one treatment would be preferred to another treatment when two treatments are being compared are also computed. The results are computed using collected data from the popular websites (www.cricmania, www.cricinfo) for these seven teams. We compute the posterior estimates (posterior means and joint mode), posterior probabilities, preference probabilities and predictive probabilities and then the results are compared. Posterior (marginal) densities for the parameters of the Rao-Kupper model are sketched for these seven teams using both ODI and Test matches data in Figures 4.1, 4.2 respectively.

The appropriateness of the model is also assessed using the chi-square goodness of fit test and it is observed that model fits with data of ODI matches but it does not fit with data of Test matches.

The results are calculated through Quadrature method as well as Gibbs sampling up to seven dimensional integration problems. It is found that the results are similar in both the models. Results of posterior means are compared in Tables 4.5. Hence, it is suggested that Gibbs sampler can be used for numerical solution especially for solving high dimensional integration. [For more detail see Aslam (2007)]. Since both Quadrature and Gibbs sampling technique give similar results therefore we use Quadrature method in major part of our study. In the last two sections, we explain the ICC teams ranking criteria and present the comparison of Bayesian ranking with ICC ranking. Here, we observe that the Bayesian Ranking almost same with the ICC Ranking for ODI Matches but on the other hand the Bayesian Ranking is differing from ICC Ranking for Test Matches. The reason is that we take the data of only top seven teams and find out the ranking, if we take all the teams then ranking result may be improved.

For further research, this work can be extended towards many course of action to the Bayesian analysis of paired comparisons data. One may increase the number of teams (i.e. more than 7). We can add the home team ground advantage criteria in this study. We can also obtain the cricket player's ranking. Efforts can be made for the team ranking of any game. Our Bayesian analysis is based upon non-informative prior

(Uniform), further analysis (ranking) can be done using informative prior for the parameter of the models.

References

Aslam, M., (2002). Bayesian analysis for paired comparisons models allowing ties and not allowing ties. *Pak. J. Statist.* Vol:18, 1,55-69

Aslam, M., (2007). Multidimensional Integration Using the Gibbs sampling . *Journal of South Asian Bulletin of Mathematics.* Vol: 31, 423-432

Aslam, M., (2003). An application of prior predictive distribution to elicit the prior density. *Journal of Statistical Theory and Applications.* Vol:2, 1,70-83

Aslm, M, (2005). Bayesian comparison of the paired comparison models allowing ties. *Journal of Statistical Theory and Application.* Vol: 4, No.2, 161-171

Annis, D.H., and Craig, B.A., (2005) Hybrid paired comparison analysis with application to the ranking of college football teams. *Journal of Quantitative Analysis in Sports*

Beaver, R. J., and Gokhale, D. V., (1975). A model to incorporate within-pair order effect in paired comparisons experiments. *Communications in statistics, 4, 923*-39

Berger, J. O., (1985*)*.Statistical decision theory and bayesian analysis, *Springer* (Ed), New York, 95-96 & 200.

Berger, J., and Wolpert, R., (1984). The likelihood principle.*Institue of Mathematical Statistics Monograph Serie. Hayward: California.*

Bolstad, W. M., (2004). Introduction to Bayesian Statistics, *New Zealand: John willey & sons, publication.*

Bock, R. Darrell., and L. V. Jones.,(1968). The measurement and prediction of judgement and choice. *San. Franciseo; Holden-Day, Inc.*

Bose, R. C., (1956). Paired comparison design for testing concordance between judges. *Biometrika,* Vol:43, 1/2, 113-121

Bradley, R. A., (1953a). Some statistical methods in taste testing and quality evaluation. *Biometrics* 9, 22-28

Bradley, R. A., (1953b). Rank Analysis of incomplete block design. *Biometrika* 39, 324-345

Bradley, R. A., (1954). Incomplete block rank analysis: On the appropriateness of the model for a method of paired comparisons. *Biometrics, Vol.*10,3, 375-390

Buchanan ., (1985). Measuring simple preferences. An approach to blind forced choice product testing. *Marketing Science*, Vol: 4. No 2

Davidson, R. R., (1969). On a relation between two representations of a model for paired comparison. *Biometrics,* Vol:25, No 3, 597-599

Davidson, R. R., (1970).On extending the Bradley-Terry model to accommodate ties in paired comparisons experiments. *Journal of American Statistical Association*, Vol:65, No.329, 317-328

Davidson, R. R., (1973). A Bayesian approach to paired comparison experimentation. Biometrika, Vol:60 , No 3, 477-487

Davidson, R. R., and Beaver, R.J., (1976). Discussion: Science, Statistics and paired comparison. *Biometrices.* Vol:32, No:2, 233-239

Davidson, R. R., (1971). A regression relationship for multivariate paired comparison *Biometrika*, Vol:58, No.3, 555-560

David, H. A., (1959). Tournaments and paired comparisons. *Biometrika*, Vol.46, No.1/2, 139-149

David, H. A., (1988). The method of paired comparisons, *2nd edn. London: Charles Griffin and Company.*

Draper, N. R., Hunter, W. G., Tierney, D. E. (1969). Which product is better? *Technometrices*, vol:11, No.2,309-320

Donald E.K Martin., (1999) Paired comparison models applied to the design of the major league baseball play-offs. *Journal of Applied Statistics*, Vol:26,69-80

Ford, L. R., (1957). Solution of a ranking problem from binary comparisons. *Amer. Math. Monthly,* 64, 28-33

Glickman, M. E., (1999). Parameter estimation in large dynamic paired comparisons experiment. *Appl. Statists*. 48, 3, 377-394.

Glickman, M. E., Stern, H.S., (1998).A state-space model for national football league scores. *Journal of American Statistical Association*. Vol.39, No.441, 25-35.

Glenn, W. A. and David, H.A., (1960). Ties in paired comparisons experiments using a modified Thurstone- Mosteller Model. *Biometrics*. Vol: 16, 86-109.

Gopalan, R., Berry, D. A.,(1998). Bayesian multiple comparison using Dirichlet process priors. *Journal of American Statistical Association*. Vol:93, No.443,1130-1139.

Gridgeman ., (1959). Paired comparison with and without ties. *Biometrics*, Vol: 15, No 3, 382-388

Grizzle, J. E., Starmer, C. F., and Koch, G.G., (1969). Analysis of categorical data by Liner models. *Biometrics*. Vol: 25, 489-504

Groeneveld, R. A., (1990). Ranking teams in a league with two divisions of *t* teams. *The American Statisticians,* vol:44, 4, 277-281

Hennery, R. J., (1992). An extension to the Thurstone- Mosteller model for chess. *The Statistician.* 41, 559-567.

Joe, H., (1990). Extended use of paired comparisons models with application to chess rankings. *Appl. Statists.*, 39 ,1, 85-93.

Kendall, M. G., Smith, B., (1940). On the method of paired comparisons. *Biometrika* 31, 324-330.

Kim, H. J., (2005). A Bayesian approach to paired comparisons ranking based on a graphical model. *Computational and Statistics and Data analysis, vol.48, 2,269-290.*

Kuk, Y. C. A., (1995). Modeling paired comparisons data with large Numbers of draws and large variability of draw percentages among players. *The Statistician,* Vol. 44, 4, 523-528

Leonard, T., (1977). An alternative Bayesian approach to the Bradley-Terry model for paired comparisons. *Biometrics* 33, 121-132.

Mosteller, F., (1951). Remark on the method of paired comparisons. *Psychometrika,* 16, 3-9

Rao, P. V., Kupper, Lawrence, L., (1967). Ties in paired comparisons experiments: A generalization of the Bradley-Terry model. *Journal of American statistical Association,* 62, 194-204.

Scheffe, (1952). An analysis of variance for paired comparisons. *JASA*. 47,381-395.

Smith, A. F. M., (1993) Bayesian computation via the gibbs sampler and related markov chain monto carlo methods. *Journal of Royal Statistical Society*. Vol:55, No.1, 3-23

Stern, H., (1990). A continuum of paired comparison models. *Biometrika* 77, 2, 265-273.

Thompson, W. A. Jr., and Singh, J. (1967). The use of limit theorems in paired comparison model building. *Psychometrica.* Vol:32, 255-264.

Thurstone, L. L., (1927). Law of comparative judgment. *Psychological Review,* 34, 273-286.

Zermelo, E., (1929). Die berechnung der turnier-ergebnisse Als Ein maximum problem der wahrscheinlichkeitsrechnang. *Math. Zeit.* Vol: 29, 436-496

APPENDIX-A

```
* POSTERIOR MODE FOR PARAMETERS OF RK MODEL

     (FOR Overall ODI MATCHES);

DATA DD;
INPUT Y1 Y2 Y3 Y4 Y5 Y6 Y7 Y8 Y9 N1 N2 N3 N4 N5 N6 N7 N12 N13 N14 N15 N16 N17
N23 N24 N25 N26 N27 N34 N35 N36 N37 N45 N46 N47 N56 N57
N67 N21 N31 N41 N51 N61 N71 N32 N42 N52 N62 N72 N43 N53 N63 N73 N54 N64 N74 N65
N75 N76 N;
CARDS;
0 0 0 0 0 0 0 0 0 308 170 255 220 181 174 161 40 47 54 73 43 51 31 32 31 22 23
68 49 68 25 40 54 31 35 29 18 31 31 31 30 21 38 14 20 18
23 13 44 30 42 34 39 42 28 28 29 19 62
;
PROC SYSNLIN DATA=DD METHOD=MARQUARDT;
Y1=(N1/T1)-N12/(T1+V*T2)-(N21*V/(T2+V*T1))-N13/(T1+V*T3)-(N31*V/(T3+V*T1))-
N14/(T1+V*T4)-(N41*V/(T4+V*T1))-
N15/(T1+V*T5)-(N51*V/(T5+V*T1))-N16/(T1+V*T6)-(N61*V/(T6+V*T1))-N17/(T1+V*T7)-
(N71*V/(T7+V*T1));
Y2=(N2/T2)-(N12*V/(T1+V*T2))-N21/(T2+V*T1)-(N32*V/(T3+V*T2))-N23/(T2+V*T3)-
(N42*V/(T4+V*T2))-N24/(T2+V*T4)
-(N52*V/(T5+V*T2))-N25/(T2+V*T5)-(N62*V/(T6+V*T2))-N26/(T2+V*T6)-
(N72*V/(T7+V*T2))-N27/(T2+V*T7);
Y3=(N3/T3)-(V*N13/(T1+V*T3))-N31/(T3+V*T1)-(V*N23/(T2+V*T3))-N32/(T3+V*T2)-
(V*N43/(T4+V*T3))-N34/(T3+V*T4)
-(V*N53/(T5+V*T3))-N35/(T3+V*T5)-(V*N63/(T6+V*T3))-N36/(T3+V*T6)-
(V*N73/(T7+V*T3))-N37/(T3+V*T7);
Y4=(N4/T4)-(V*N14/(T1+V*T4))-N41/(T4+V*T1)-(V*N24/(T2+V*T4))-N42/(T4+V*T2)-
(V*N34/(T3+V*T4))-N43/(T4+V*T3)
-(V*N54/(T5+V*T4))-N45/(T4+V*T5)-(V*N64/(T6+V*T4))-N46/(T4+V*T6)-
(V*N74/(T7+V*T4))-N47/(T4+V*T7);
Y5=(N5/T5)-(V*N15/(T1+V*T5))-N51/(T5+V*T1)-(V*N25/(T2+V*T5))-N52/(T5+V*T2)-
(V*N35/(T3+V*T5))-N53/(T5+V*T3)
-(V*N45/(T4+V*T5))-N54/(T5+V*T4)-(V*N65/(T6+V*T5))-N56/(T5+V*T6)-
(V*N75/(T7+V*T5))-N57/(T5+V*T7);
Y6=(N6/T6)-(V*N16/(T1+V*T6))-N61/(T6+V*T1)-(V*N26/(T2+V*T6))-N62/(T6+V*T2)-
(V*N36/(T3+V*T6))-N63/(T6+V*T3)
-(V*N46/(T4+V*T6))-N64/(T6+V*T4)-(V*N56/(T5+V*T6))-N65/(T6+V*T5)-
(V*N76/(T7+V*T6))-N67/(T6+V*T7);
Y7=(N7/T7)-(V*N17/(T1+V*T7))-N71/(T7+V*T1)-(V*N27/(T2+V*T7))-N72/(T7+V*T2)-
(V*N37/(T3+V*T7))-N73/(T7+V*T3)
-(V*N47/(T4+V*T7))-N74/(T7+V*T4)-(V*N57/(T5+V*T7))-N75/(T7+V*T5)-
(V*N67/(T6+V*T7))-N76/(T7+V*T6);
Y8=1-T1-T2-T3-T4-T5-T6-T7;
Y9=(2*N*V/(V**2-1))-(T2*N12/(T1+V*T2))-(T1*N21/(T2+V*T1))-(T3*N13/(T1+V*T3))-
(T1*N31/(T3+V*T1))-(T4*N14/(T1+V*T4))-(T1*N41/(T4+V*T1))
-(T5*N15/(T1+V*T5))-(T1*N51/(T5+V*T1))-(T6*N16/(T1+V*T6))-(T1*N61/(T6+V*T1))-
(T7*N17/(T1+V*T7))-(T1*N71/(T7+V*T1))
-(T3*N23/(T2+V*T3))-(T2*N32/(T3+V*T2))-(T4*N24/(T2+V*T4))-(T2*N42/(T4+V*T2))-
(T5*N25/(T2+V*T5))-(T2*N52/(T5+V*T2))
-(T6*N26/(T2+V*T6))-(T2*N62/(T6+V*T2))-(T7*N27/(T2+V*T7))-(T2*N72/(T7+V*T2))-
(T4*N34/(T3+V*T4))-(T3*N43/(T4+V*T3))
-(T5*N35/(T3+V*T5))-(T3*N53/(T5+V*T3))-(T6*N36/(T3+V*T6))-(T3*N63/(T6+V*T3))-
(T7*N37/(T3+V*T7))-(T3*N73/(T7+V*T3))
-(T5*N45/(T4+V*T5))-(T4*N54/(T5+V*T4))-(T6*N46/(T4+V*T6))-(T4*N64/(T6+V*T4))-
(T7*N47/(T4+V*T7))-(T4*N74/(T7+V*T4))
-(T6*N56/(T5+V*T6))-(T5*N65/(T6+V*T5))-(T7*N57/(T5+V*T7))-(T5*N75/(T7+V*T5))-
(T7*N67/(T6+V*T7))-(T6*N76/(T7+V*T6));
ENDO Y1 Y2 Y3 Y4 Y5 Y6 Y7 Y8 Y9;
EXO N1 N2 N3 N4 N5 N6 N7 N12 N13 N14 N15 N16 N17 N23 N24 N25 N26 N27 N34 N35
N36 N37 N45 N46 N47 N56 N57 N67 N21 N31 N41 N51 N61 N71 N32
N42 N52 N62 N72 N43 N53 N63 N73 N54 N64 N74 N65 N75 N76 N;
PARMS T1 0.2195 T2 0.1842 T3 0.1455 T4 0.1151 T5 0.1045 T6 0.1004 T7 0.1308 V
1.096;
RUN;
```

APPENDIX-B

```
/* 'C' Codes to find The Posterior Means for ODI from 1999 to 2003
by Quadrature Method */
# include <stdio.h>
# include <math.h>
# include <conio.h>
void main()
{
double
V,T1,T2,T3,T4,T5,T6,T7,fun,integ=0.0,integc=0.0,Vp=0.0,Tp1=0.0,Tp2=0.0,Tp3=0.0,T
p4=0.0,
Tp5=0.0,Tp6=0.0,Tp7,dl=0.045; // all variables separated by commas
//clrscr();
//getch();
printf("Start of Program...\n");
// Finding the Normalizing Constant
for (T1=dl;T1<=1.0-dl;T1+=dl)
for (T2=dl;T2<=1.0-T1-dl;T2+=dl)
for (T3=dl;T3<=1.0-T1-T2-dl;T3+=dl)
for (T4=dl;T4<=1.0-T1-T2-T3-dl;T4+=dl)
for (T5=dl;T5<=1.0-T1-T2-T3-T4-dl;T5+=dl)
for (T6=dl;T6<=1.0-T1-T2-T3-T4-T5-dl;T6+=dl)
for(V=1.0+dl;V<=10.0-dl;V+=dl)
{
T7=1.0-T1-T2-T3-T4-T5-T6;
if( T7<= 0.0) continue;

fun=pow((V*V-
1),16)*(pow(T1,79))*pow(T2,68)*pow(T3,61)*pow(T4,53)*pow(T5,42)*pow(T6,55)*p
ow(T7,27)
/(
(pow((T1+V*T2),14))*(pow((T1+V*T3),13))*(pow((T1+V*T4),15))*(pow((T1+V*T5),
11))*(pow((T1+V*T6),11))*(pow((T1+V*T7),15))*(pow((T2+V*T3),14))*
(pow((T2+V*T4),11))*(pow((T2+V*T5),18))*(pow((T2+V*T6),11))*(pow((T2+V*T7),
6))*(pow((T3+V*T4),10))*(pow((T3+V*T5),17))*(pow((T3+V*T6),15))*(pow((T3+V*
T7),7))*(pow((T4+V*T5),12))*(pow((T4+V*T6),12))*(pow((T4+V*T7),11))*
(pow((T5+V*T6),4))*(pow((T5+V*T7),3))*(pow((T6+V*T7),8))*(pow((T2+V*T1),8))*
(pow((T3+V*T1),6))*(pow((T4+V*T1),4))*
(pow((T5+V*T1),6))*(pow((T6+V*T1),4))*(pow((T7+V*T1),2))*(pow((T3+V*T2),6))*
(pow((T4+V*T2),9))*(pow((T5+V*T2),7))*(pow((T6+V*T2),9))*
(pow((T7+V*T2),4))*(pow((T4+V*T3),5))*(pow((T5+V*T3),6))*(pow((T6+V*T3),14))
*(pow((T7+V*T3),5))*(pow((T5+V*T4),16))*
(pow((T6+V*T4),9))*(pow((T7+V*T4),5))*(pow((T6+V*T5),11))*(pow((T7+V*T5),2))
*(pow((T7+V*T6),9)));
integ+=fun*pow(dl,7);
```

```
//printf("\nNC:\t%g",integ);
}
printf("\nNC:\t%g",integ);// getch();
for (T1=dl;T1<=1.0-dl;T1+=dl)
for (T2=dl;T2<=1.0-T1-dl;T2+=dl)
for (T3=dl;T3<=1.0-T1-T2-dl;T3+=dl)
for (T4=dl;T4<=1.0-T1-T2-T3-dl;T4+=dl)
for (T5=dl;T5<=1.0-T1-T2-T3-T4-dl;T5+=dl)
for (T6=dl;T6<=1.0-T1-T2-T3-T4-T5-dl;T6+=dl)
for(V=1.0+dl;V<=10.0-dl;V+=dl)
{
T7=1.0-T1-T2-T3-T4-T5-T6;
if( T7<=0.0) continue;
fun=(1.0/integ)*pow((V*V-
1),16)*(pow(T1,79))*pow(T2,68)*pow(T3,61)*pow(T4,53)*pow(T5,42)*pow(T6,55)*p
ow(T7,27)
/(
(pow((T1+V*T2),14))*(pow((T1+V*T3),13))*(pow((T1+V*T4),15))*(pow((T1+V*T5),
11))*(pow((T1+V*T6),11))*(pow((T1+V*T7),15))*(pow((T2+V*T3),14))*
(pow((T2+V*T4),11))*(pow((T2+V*T5),18))*(pow((T2+V*T6),11))*(pow((T2+V*T7),
6))*(pow((T3+V*T4),10))*(pow((T3+V*T5),17))*(pow((T3+V*T6),15))*(pow((T3+V*
T7),7))*(pow((T4+V*T5),12))*(pow((T4+V*T6),12))*(pow((T4+V*T7),11))*
(pow((T5+V*T6),4))*(pow((T5+V*T7),3))*(pow((T6+V*T7),8))*(pow((T2+V*T1),8))*
(pow((T3+V*T1),6))*(pow((T4+V*T1),4))*
(pow((T5+V*T1),6))*(pow((T6+V*T1),4))*(pow((T7+V*T1),2))*(pow((T3+V*T2),6))*
(pow((T4+V*T2),9))*(pow((T5+V*T2),7))*(pow((T6+V*T2),9))*
(pow((T7+V*T2),4))*(pow((T4+V*T3),5))*(pow((T5+V*T3),6))*(pow((T6+V*T3),14))
*(pow((T7+V*T3),5))*(pow((T5+V*T4),16))*
(pow((T6+V*T4),9))*(pow((T7+V*T4),5))*(pow((T6+V*T5),11))*(pow((T7+V*T5),2))
*(pow((T7+V*T6),9)));
integc+=fun*dl*dl*dl*dl*dl*dl*dl;
fun=(1.0/integ)*T1*pow((V*V-
1),16)*(pow(T1,79))*pow(T2,68)*pow(T3,61)*pow(T4,53)*pow(T5,42)*pow(T6,55)*p
ow(T7,27)
/(
(pow((T1+V*T2),14))*(pow((T1+V*T3),13))*(pow((T1+V*T4),15))*(pow((T1+V*T5),
11))*(pow((T1+V*T6),11))*(pow((T1+V*T7),15))*(pow((T2+V*T3),14))*
(pow((T2+V*T4),11))*(pow((T2+V*T5),18))*(pow((T2+V*T6),11))*(pow((T2+V*T7),
6))*(pow((T3+V*T4),10))*(pow((T3+V*T5),17))*(pow((T3+V*T6),15))*(pow((T3+V*
T7),7))*(pow((T4+V*T5),12))*(pow((T4+V*T6),12))*(pow((T4+V*T7),11))*
(pow((T5+V*T6),4))*(pow((T5+V*T7),3))*(pow((T6+V*T7),8))*(pow((T2+V*T1),8))*
(pow((T3+V*T1),6))*(pow((T4+V*T1),4))*
(pow((T5+V*T1),6))*(pow((T6+V*T1),4))*(pow((T7+V*T1),2))*(pow((T3+V*T2),6))*
(pow((T4+V*T2),9))*(pow((T5+V*T2),7))*(pow((T6+V*T2),9))*
(pow((T7+V*T2),4))*(pow((T4+V*T3),5))*(pow((T5+V*T3),6))*(pow((T6+V*T3),14))
*(pow((T7+V*T3),5))*(pow((T5+V*T4),16))*
```

```
(pow((T6+V*T4),9))*(pow((T7+V*T4),5))*(pow((T6+V*T5),11))*(pow((T7+V*T5),2))
*(pow((T7+V*T6),9)));
Tp1+=fun*dl*dl*dl*dl*dl*dl*dl;
fun=(1.0/integ)*T2*pow((V*V-
1),16)*(pow(T1,79))*pow(T2,68)*pow(T3,61)*pow(T4,53)*pow(T5,42)*pow(T6,55)*p
ow(T7,27)
/(
(pow((T1+V*T2),14))*(pow((T1+V*T3),13))*(pow((T1+V*T4),15))*(pow((T1+V*T5),
11))*(pow((T1+V*T6),11))*(pow((T1+V*T7),15))*(pow((T2+V*T3),14))*
(pow((T2+V*T4),11))*(pow((T2+V*T5),18))*(pow((T2+V*T6),11))*(pow((T2+V*T7),
6))*(pow((T3+V*T4),10))*(pow((T3+V*T5),17))*(pow((T3+V*T6),15))*(pow((T3+V*
T7),7))*(pow((T4+V*T5),12))*(pow((T4+V*T6),12))*(pow((T4+V*T7),11))*
(pow((T5+V*T6),4))*(pow((T5+V*T7),3))*(pow((T6+V*T7),8))*(pow((T2+V*T1),8))*
(pow((T3+V*T1),6))*(pow((T4+V*T1),4))*
(pow((T5+V*T1),6))*(pow((T6+V*T1),4))*(pow((T7+V*T1),2))*(pow((T3+V*T2),6))*
(pow((T4+V*T2),9))*(pow((T5+V*T2),7))*(pow((T6+V*T2),9))*
(pow((T7+V*T2),4))*(pow((T4+V*T3),5))*(pow((T5+V*T3),6))*(pow((T6+V*T3),14))
*(pow((T7+V*T3),5))*(pow((T5+V*T4),16))*
(pow((T6+V*T4),9))*(pow((T7+V*T4),5))*(pow((T6+V*T5),11))*(pow((T7+V*T5),2))
*(pow((T7+V*T6),9)));
Tp2+=fun*dl*dl*dl*dl*dl*dl*dl;
fun=(1.0/integ)*T3*pow((V*V-
1),16)*(pow(T1,79))*pow(T2,68)*pow(T3,61)*pow(T4,53)*pow(T5,42)*pow(T6,55)*p
ow(T7,27)
/(
(pow((T1+V*T2),14))*(pow((T1+V*T3),13))*(pow((T1+V*T4),15))*(pow((T1+V*T5),
11))*(pow((T1+V*T6),11))*(pow((T1+V*T7),15))*(pow((T2+V*T3),14))*
(pow((T2+V*T4),11))*(pow((T2+V*T5),18))*(pow((T2+V*T6),11))*(pow((T2+V*T7),
6))*(pow((T3+V*T4),10))*(pow((T3+V*T5),17))*(pow((T3+V*T6),15))*(pow((T3+V*
T7),7))*(pow((T4+V*T5),12))*(pow((T4+V*T6),12))*(pow((T4+V*T7),11))*
(pow((T5+V*T6),4))*(pow((T5+V*T7),3))*(pow((T6+V*T7),8))*(pow((T2+V*T1),8))*
(pow((T3+V*T1),6))*(pow((T4+V*T1),4))*
(pow((T5+V*T1),6))*(pow((T6+V*T1),4))*(pow((T7+V*T1),2))*(pow((T3+V*T2),6))*
(pow((T4+V*T2),9))*(pow((T5+V*T2),7))*(pow((T6+V*T2),9))*
(pow((T7+V*T2),4))*(pow((T4+V*T3),5))*(pow((T5+V*T3),6))*(pow((T6+V*T3),14))
*(pow((T7+V*T3),5))*(pow((T5+V*T4),16))*
(pow((T6+V*T4),9))*(pow((T7+V*T4),5))*(pow((T6+V*T5),11))*(pow((T7+V*T5),2))
*(pow((T7+V*T6),9)));
Tp3+=fun*dl*dl*dl*dl*dl*dl*dl;
```

APPENDIX-C

```
/* 'C' Codes to find The Posterior Predictive Probabilities for the RK Model for all ODI
by Quadrature Method */
# include <stdio.h>
# include <math.h>
# include <conio.h>
void main()
{
double
V,T1,T2,T3,T4,T5,T6,T7,fun,integ=0.0,integc=0.0,PPD12=0.0,PPD13=0.0,PPD14=0.0,P
PD15=0.0,PPD16=0.0,PPD17=0.0,
PPD23=0.0,PPD24=0.0,PPD25=0.0,PPD26=0.0,PPD27=0.0,PPD34=0.0,PPD35=0.0,PP
D36=0.0,PPD37=0.0,PPD45=0.0,
PPD46=0.0,PPD47=0.0,PPD56=0.0,PPD57=0.0,PPD67=0.0,dl=0.083; // all variables
separated by commas
//clrscr();
//getch();
printf("Start of Program...\n");
// Finding the Normalizing Constant
for (T1=dl;T1<=1.0-dl;T1+=dl)
for (T2=dl;T2<=1.0-T1-dl;T2+=dl)
for (T3=dl;T3<=1.0-T1-T2-dl;T3+=dl)
for (T4=dl;T4<=1.0-T1-T2-T3-dl;T4+=dl)
for (T5=dl;T5<=1.0-T1-T2-T3-T4-dl;T5+=dl)
for (T6=dl;T6<=1.0-T1-T2-T3-T4-T5-dl;T6+=dl)
for(V=1.0+dl;V<=10.0-dl;V+=dl)
{
T7=1.0-T1-T2-T3-T4-T5-T6;
if( T7<= 0.0) continue;

fun=pow((V*V-
1),12)*(pow(T1,62))*pow(T2,34)*pow(T3,51)*pow(T4,44)*pow(T5,36)*pow(T6,35)*p
ow(T7,32)
/(
(pow((T1+V*T2),8))*(pow((T1+V*T3),9))*(pow((T1+V*T4),11))*(pow((T1+V*T5),15
))*(pow((T1+V*T6),9))
*(pow((T1+V*T7),10))*(pow((T2+V*T3),6))*
(pow((T2+V*T4),6))*(pow((T2+V*T5),6))*(pow((T2+V*T6),4))*(pow((T2+V*T7),5))*
(pow((T3+V*T4),14))*
(pow((T3+V*T5),10))*(pow((T3+V*T6),14))*(pow((T3+V*T7),5))*(pow((T4+V*T5),8
))*(pow((T4+V*T6),11))*(pow((T4+V*T7),6))*
(pow((T5+V*T6),7))*(pow((T5+V*T7),6))*(pow((T6+V*T7),4))*(pow((T2+V*T1),6))*
(pow((T3+V*T1),6))*(pow((T4+V*T1),6))*
(pow((T5+V*T1),6))*(pow((T6+V*T1),4))*(pow((T7+V*T1),8))*(pow((T3+V*T2),3))*
(pow((T4+V*T2),4))*
```

```
(pow((T5+V*T2),4))*(pow((T6+V*T2),5))*
(pow((T7+V*T2),3))*(pow((T4+V*T3),9))*(pow((T5+V*T3),6))*(pow((T6+V*T3),9))*
(pow((T7+V*T3),7))*
(pow((T5+V*T4),8))*(pow((T6+V*T4),8))*(pow((T7+V*T4),6))*(pow((T6+V*T5),6))*
(pow((T7+V*T5),6))*
(pow((T7+V*T6),4)));
integ+=fun*pow(dl,7);
//printf("\nNC:\t%g",integ);
}
printf("\nNC:\t%g",integ);// getch();
for (T1=dl;T1<=1.0-dl;T1+=dl)
for (T2=dl;T2<=1.0-T1-dl;T2+=dl)
for (T3=dl;T3<=1.0-T1-T2-dl;T3+=dl)
for (T4=dl;T4<=1.0-T1-T2-T3-dl;T4+=dl)
for (T5=dl;T5<=1.0-T1-T2-T3-T4-dl;T5+=dl)
for (T6=dl;T6<=1.0-T1-T2-T3-T4-T5-dl;T6+=dl)
for(V=1.0+dl;V<=10.0-dl;V+=dl)
{
T7=1.0-T1-T2-T3-T4-T5-T6;
if( T7<=0.0) continue;
fun=(1.0/integ)*pow((V*V-
1),12)*(pow(T1,62))*pow(T2,34)*pow(T3,51)*pow(T4,44)*pow(T5,36)*pow(T6,35)*p
ow(T7,32)
/(
(pow((T1+V*T2),8))*(pow((T1+V*T3),9))*(pow((T1+V*T4),11))*(pow((T1+V*T5),15
))*(pow((T1+V*T6),9))
*(pow((T1+V*T7),10))*(pow((T2+V*T3),6))*
(pow((T2+V*T4),6))*(pow((T2+V*T5),6))*(pow((T2+V*T6),4))*(pow((T2+V*T7),5))*
(pow((T3+V*T4),14))*
(pow((T3+V*T5),10))*(pow((T3+V*T6),14))*(pow((T3+V*T7),5))*(pow((T4+V*T5),8
))*(pow((T4+V*T6),11))*(pow((T4+V*T7),6))*
(pow((T5+V*T6),7))*(pow((T5+V*T7),6))*(pow((T6+V*T7),4))*(pow((T2+V*T1),6))*
(pow((T3+V*T1),6))*(pow((T4+V*T1),6))*
(pow((T5+V*T1),6))*(pow((T6+V*T1),4))*(pow((T7+V*T1),8))*(pow((T3+V*T2),3))*
(pow((T4+V*T2),4))*
(pow((T5+V*T2),4))*(pow((T6+V*T2),5))*
(pow((T7+V*T2),3))*(pow((T4+V*T3),9))*(pow((T5+V*T3),6))*(pow((T6+V*T3),9))*
(pow((T7+V*T3),7))*
(pow((T5+V*T4),8))*(pow((T6+V*T4),8))*(pow((T7+V*T4),6))*(pow((T6+V*T5),6))*
(pow((T7+V*T5),6))*
(pow((T7+V*T6),4)));
integc+=fun*dl*dl*dl*dl*dl*dl*dl;

fun=(1.0/integ)*T1/(T1+V*T2)*pow((V*V-
1),12)*(pow(T1,62))*pow(T2,34)*pow(T3,51)*pow(T4,44)*pow(T5,36)*pow(T6,35)*p
ow(T7,32)
```

```
/(
(pow((T1+V*T2),8))*(pow((T1+V*T3),9))*(pow((T1+V*T4),11))*(pow((T1+V*T5),15
))*(pow((T1+V*T6),9))
*(pow((T1+V*T7),10))*(pow((T2+V*T3),6))*
(pow((T2+V*T4),6))*(pow((T2+V*T5),6))*(pow((T2+V*T6),4))*(pow((T2+V*T7),5))*
(pow((T3+V*T4),14))*
(pow((T3+V*T5),10))*(pow((T3+V*T6),14))*(pow((T3+V*T7),5))*(pow((T4+V*T5),8
))*(pow((T4+V*T6),11))*(pow((T4+V*T7),6))*
(pow((T5+V*T6),7))*(pow((T5+V*T7),6))*(pow((T6+V*T7),4))*(pow((T2+V*T1),6))*
(pow((T3+V*T1),6))*(pow((T4+V*T1),6))*
(pow((T5+V*T1),6))*(pow((T6+V*T1),4))*(pow((T7+V*T1),8))*(pow((T3+V*T2),3))*
(pow((T4+V*T2),4))*
(pow((T5+V*T2),4))*(pow((T6+V*T2),5))*
(pow((T7+V*T2),3))*(pow((T4+V*T3),9))*(pow((T5+V*T3),6))*(pow((T6+V*T3),9))*
(pow((T7+V*T3),7))*
(pow((T5+V*T4),8))*(pow((T6+V*T4),8))*(pow((T7+V*T4),6))*(pow((T6+V*T5),6))*
(pow((T7+V*T5),6))*
(pow((T7+V*T6),4)));
PPD12+=fun*dl*dl*dl*dl*dl*dl*dl;

fun=(1.0/integ)*T1/(T1+V*T3)*pow((V*V-
1),12)*(pow(T1,62))*pow(T2,34)*pow(T3,51)*pow(T4,44)*pow(T5,36)*pow(T6,35)*p
ow(T7,32)
/(
(pow((T1+V*T2),8))*(pow((T1+V*T3),9))*(pow((T1+V*T4),11))*(pow((T1+V*T5),15
))*(pow((T1+V*T6),9))
*(pow((T1+V*T7),10))*(pow((T2+V*T3),6))*
(pow((T2+V*T4),6))*(pow((T2+V*T5),6))*(pow((T2+V*T6),4))*(pow((T2+V*T7),5))*
(pow((T3+V*T4),14))*
(pow((T3+V*T5),10))*(pow((T3+V*T6),14))*(pow((T3+V*T7),5))*(pow((T4+V*T5),8
))*(pow((T4+V*T6),11))*(pow((T4+V*T7),6))*
(pow((T5+V*T6),7))*(pow((T5+V*T7),6))*(pow((T6+V*T7),4))*(pow((T2+V*T1),6))*
(pow((T3+V*T1),6))*(pow((T4+V*T1),6))*
(pow((T5+V*T1),6))*(pow((T6+V*T1),4))*(pow((T7+V*T1),8))*(pow((T3+V*T2),3))*
(pow((T4+V*T2),4))*
(pow((T5+V*T2),4))*(pow((T6+V*T2),5))*
(pow((T7+V*T2),3))*(pow((T4+V*T3),9))*(pow((T5+V*T3),6))*(pow((T6+V*T3),9))*
(pow((T7+V*T3),7))*
(pow((T5+V*T4),8))*(pow((T6+V*T4),8))*(pow((T7+V*T4),6))*(pow((T6+V*T5),6))*
(pow((T7+V*T5),6))*
(pow((T7+V*T6),4)));
PPD13+=fun*dl*dl*dl*dl*dl*dl*dl;
```

APPENDIX-D

```
/* 'C' Codes to find The Posterior probability P12 of RK model for Overall ODI matches
by Quadrature Method */
# include <stdio.h>
# include <math.h>
# include <conio.h>
void main()
{
double
S,P,V,T1,T2,T3,T4,T5,T6,T7,fun,integ=0.0,pho=0.0,dl=0.0875; // all variables separated
by commas
//clrscr();
//getch();
printf("\nStart of Program...");
// Finding the Normalizing Constant
for (P=dl;P<=1.0-dl;P+=dl)
for (S=P+dl;S<=(1.0+P)/2.0-dl;S+=dl)
for (T3=dl;T3<=1.0-2.0*S+P-dl;T3+=dl)
for (T4=dl;T4<=1.0-2.0*S+P-T3-dl;T4+=dl)
for (T5=dl;T5<=1.0-2.0*S+P-T3-T4-dl;T5+=dl)
for (T6=dl;T6<=1.0-2.0*S+P-T3-T4-T5-dl;T6+=dl)
for(V=1.0+dl;V<=5.0-dl;V+=dl)
{
T7=1.0-2.0*S+P-T3-T4-T5-T6;
if( T7<= 0.0) continue;
T1=S; T2=S-P;
fun=
pow((V*V-
1.0),12)*(pow(T1,62))*pow(T2,34)*pow(T3,51)*pow(T4,44)*pow(T5,36)*pow(T6,35)*
pow(T7,32)/(
(pow((T1+V*T2),8))*(pow((T1+V*T3),9))*(pow((T1+V*T4),11))*(pow((T1+V*T5),15
))*(pow((T1+V*T6),9))*
(pow((T1+V*T7),10))*(pow((T2+V*T3),6))*(pow((T2+V*T4),6))*(pow((T2+V*T5),6))
*(pow((T2+V*T6),4))*
(pow((T2+V*T7),5))*(pow((T3+V*T4),14))*(pow((T3+V*T5),10))*(pow((T3+V*T6),1
4))*(pow((T3+V*T7),5))*
(pow((T4+V*T5),8))*(pow((T4+V*T6),11))*(pow((T4+V*T7),6))*(pow((T5+V*T6),7))
*(pow((T5+V*T7),6))*
(pow((T6+V*T7),4))*(pow((T2+V*T1),6))*(pow((T3+V*T1),6))*(pow((T4+V*T1),6))*
(pow((T5+V*T1),6))*(pow((T6+V*T1),4))*
(pow((T7+V*T1),8))*(pow((T3+V*T2),3))*(pow((T4+V*T2),4))*(pow((T5+V*T2),4))*
(pow((T6+V*T2),5))*
(pow((T7+V*T2),3))*(pow((T4+V*T3),9))*(pow((T5+V*T3),6))*(pow((T6+V*T3),9))*
(pow((T7+V*T3),7))*
(pow((T5+V*T4),8))*(pow((T6+V*T4),8))*(pow((T7+V*T4),6))*(pow((T6+V*T5),6))*
```

```
(pow((T7+V*T5),6))*(pow((T7+V*T6),4)));
integ+=fun*pow(dl,7);
}

for (P=dl;P<=1.0-dl;P+=dl)
for (S=P+dl;S<=(1.0+P)/2.0-dl;S+=dl)
for (T3=dl;T3<=1.0-2.0*S+P-dl;T3+=dl)
for (T4=dl;T4<=1.0-2.0*S+P-T3-dl;T4+=dl)
for (T5=dl;T5<=1.0-2.0*S+P-T3-T4-dl;T5+=dl)
for (T6=dl;T6<=1.0-2.0*S+P-T3-T4-T5-dl;T6+=dl)
for(V=1.0+dl;V<=5.0-dl;V+=dl)
{
T7=1.0-2.0*S+P-T3-T4-T5-T6;
if( T7<= 0.0) continue;
T1=S; T2=S-P;
fun=1.0/integ*
pow((V*V-
1.0),12)*(pow(T1,62))*pow(T2,34)*pow(T3,51)*pow(T4,44)*pow(T5,36)*pow(T6,35)*
pow(T7,32)/(
(pow((T1+V*T2),8))*(pow((T1+V*T3),9))*(pow((T1+V*T4),11))*(pow((T1+V*T5),15
))*(pow((T1+V*T6),9))*
(pow((T1+V*T7),10))*(pow((T2+V*T3),6))*(pow((T2+V*T4),6))*(pow((T2+V*T5),6))
*(pow((T2+V*T6),4))*
(pow((T2+V*T7),5))*(pow((T3+V*T4),14))*(pow((T3+V*T5),10))*(pow((T3+V*T6),1
4))*(pow((T3+V*T7),5))*
(pow((T4+V*T5),8))*(pow((T4+V*T6),11))*(pow((T4+V*T7),6))*(pow((T5+V*T6),7))
*(pow((T5+V*T7),6))*
(pow((T6+V*T7),4))*(pow((T2+V*T1),6))*(pow((T3+V*T1),6))*(pow((T4+V*T1),6))*
(pow((T5+V*T1),6))*(pow((T6+V*T1),4))*
(pow((T7+V*T1),8))*(pow((T3+V*T2),3))*(pow((T4+V*T2),4))*(pow((T5+V*T2),4))*
(pow((T6+V*T2),5))*
(pow((T7+V*T2),3))*(pow((T4+V*T3),9))*(pow((T5+V*T3),6))*(pow((T6+V*T3),9))*
(pow((T7+V*T3),7))*
(pow((T5+V*T4),8))*(pow((T6+V*T4),8))*(pow((T7+V*T4),6))*(pow((T6+V*T5),6))*
(pow((T7+V*T5),6))*(pow((T7+V*T6),4)));
pho+=fun*pow(dl,7);
}
printf("\nInteg:%g",integ);
printf("\nIncremenyt:\t %g, P(Ho):\t%g, P(H1):\t %g"dl,pho,1.0-pho);
printf("\nProgram ended, Press Enter to exit...");
getch();
}
```

APPENDIX-E

```
                OPTION NONOTES;
   *GIBBS SAMPLING FOR POSTERIOR MEAN OF RAO-KUPPER MODEL m=7;
DATA DD;
INPUT N1 N2 N3 N4 N5 N6 N7 N12 N13 N14 N15 N16 N17 N23 N24 N25 N26 N27 N34 N35 N36 N37
N45 N46 N47
N56 N57 N67 N21 N31 N41 N51 N61 N71 N32 N42 N52 N62 N72 N43 N53 N63 N73 N54 N64 N74 N65
N75 N76 N DA;
CARDS;
        61.6 34.0 51.0 44.0 36.2 34.8 32.2 8.0 9.4 10.8 14.6 8.6 10.2 6.2 6.4 6.2 4.4 4.6
13.6 9.8 13.6 5.0 8.0 10.8 6.2 7.0 5.8 3.6
        6.2 6.2 6.2 6.0 4.2 7.6 2.8 4.0 3.6 4.6 2.6 8.8 6.0 8.4 6.8 7.8 8.4 5.6 5.6 5.8
3.8 12.4 0.01
;
*PROC PRINT DATA=DD; RUN;

DATA DT1; SET DD; T2=0.3; T3=0.2; T4=0.15; T5=0.1; T6=0.1; T7=0.1; V=1.23; DA=0.01;
QQ=1;
DO T1=DA TO 1-T2-T3-T4-T5-T6-T7-DA BY 0.01;
PTH=(((V**2-1)**N)*(T1**N1)*(T2**N2)*(T3**N3)*(T4**N4)*(T5**N5)*(T6**N6)*(T7**N7))/
(((T1+V*T2)**N12)*((T1+V*T3)**N13)*((T1+V*T4)**N14)*((T1+V*T5)**N15)*((T1+V*T6)**N16)*((T
1+V*T7)**N17)*
((T2+V*T3)**N23)*((T2+V*T4)**N24)*((T2+V*T5)**N25)*((T2+V*T6)**N26)*((T2+V*T7)**N27)*
((T3+V*T4)**N34)*((T3+V*T5)**N35)*((T3+V*T6)**N36)*((T3+V*T7)**N37)*
((T4+V*T5)**N45)*((T4+V*T6)**N46)*((T4+V*T7)**N47)*((T5+V*T6)**N56)*((T5+V*T7)**N57)*((T6
+V*T7)**N67)*((T2+V*T1)**N21)*
((T3+V*T1)**N31)*((T4+V*T1)**N41)*((T5+V*T1)**N51)*((T6+V*T1)**N61)*((T7+V*T1)**N71)*
((T3+V*T2)**N32)*((T4+V*T2)**N42)*((T5+V*T2)**N52)*((T6+V*T2)**N62)*((T7+V*T2)**N72)*
((T4+V*T3)**N43)*((T5+V*T3)**N53)*((T6+V*T3)**N63)*((T7+V*T3)**N73)*
((T5+V*T4)**N54)*((T6+V*T4)**N64)*((T7+V*T4)**N74)*
((T6+V*T5)**N65)*((T7+V*T5)**N75)*((T7+V*T6)**N76));
DFT1=PTH*DA; AFT1=DFT1**(2/3);
OUTPUT; END; *KEEP T2 T3 T4 T5 T6 T7 V  DFT1 AFT1 QQ;
*RUN;
PROC MEANS DATA=DT1 MAX NOPRINT; VAR AFT1; OUTPUT OUT=DDM MAX=MDT1;
DATA DD1; SET DDM (KEEP=MDT1); QQ=1;
DATA DD2; MERGE DD1 DT1; BY QQ;
ZZ=MDT1-AFT1;
IF ZZ=0; MT1=T1; *KEEP MT1 QQ;
DATA DD3; MERGE DT1 DD2; BY QQ;
YY=T1-MT1; FBB=DFT1**(1/3);
DGM=YY*FBB; *KEEP DGM;
*PROC PRINT DATA=DD3; *RUN;
PROC MEANS DATA=DD3 MIN NOPRINT; VAR DGM; OUTPUT OUT=BBN MIN=MINB;
PROC MEANS DATA=DD3 MAX NOPRINT; VAR DGM; OUTPUT OUT=BBP MAX=MAXB;
DATA DDR; MERGE BBN BBP DD2;
DO UNTIL(0<T1<1-T2-T3-T4-T5-T6-T7);
DO UNTIL(UUU<FRT1);
R1=RANUNI(0);    R2=RANUNI(0);
UU=MDT1*R1; VV=MINB+(MAXB-MINB)*R2;
UUU=UU**(3/2); T1=VV/SQRT(UU)+MT1;
FRT1=(((V**2-1)**N)*(T1**N1)*(T2**N2)*(T3**N3)*(T4**N4)*(T5**N5)*(T6**N6)*(T7**N7))/
(((T1+V*T2)**N12)*((T1+V*T3)**N13)*((T1+V*T4)**N14)*((T1+V*T5)**N15)*((T1+V*T6)**N16)*((T
1+V*T7)**N17)*
((T2+V*T3)**N23)*((T2+V*T4)**N24)*((T2+V*T5)**N25)*((T2+V*T6)**N26)*((T2+V*T7)**N27)*
((T3+V*T4)**N34)*((T3+V*T5)**N35)*((T3+V*T6)**N36)*((T3+V*T7)**N37)*
((T4+V*T5)**N45)*((T4+V*T6)**N46)*((T4+V*T7)**N47)*((T5+V*T6)**N56)*((T5+V*T7)**N57)*((T6
+V*T7)**N67)*((T2+V*T1)**N21)*
((T3+V*T1)**N31)*((T4+V*T1)**N41)*((T5+V*T1)**N51)*((T6+V*T1)**N61)*((T7+V*T1)**N71)*
((T3+V*T2)**N32)*((T4+V*T2)**N42)*((T5+V*T2)**N52)*((T6+V*T2)**N62)*((T7+V*T2)**N72)*
((T4+V*T3)**N43)*((T5+V*T3)**N53)*((T6+V*T3)**N63)*((T7+V*T3)**N73)*
((T5+V*T4)**N54)*((T6+V*T4)**N64)*((T7+V*T4)**N74)*
((T6+V*T5)**N65)*((T7+V*T5)**N75)*((T7+V*T6)**N76));
END; END; KEEP T1;
*PROC PRINT DATA=DDR; RUN;


DATA DT2; MERGE DT1 DDR;T3=0.2; T4=0.15; T5=0.1; T6=0.1; T7=0.05; V=1.23; QQ=1;
DO T2=DA TO 1-T1-T3-T4-T5-T6-T7-0.01 BY 0.01;
PTH=(((V**2-1)**N)*(T1**N1)*(T2**N2)*(T3**N3)*(T4**N4)*(T5**N5)*(T6**N6)*(T7**N7))/
(((T1+V*T2)**N12)*((T1+V*T3)**N13)*((T1+V*T4)**N14)*((T1+V*T5)**N15)*((T1+V*T6)**N16)*((T
1+V*T7)**N17)*
((T2+V*T3)**N23)*((T2+V*T4)**N24)*((T2+V*T5)**N25)*((T2+V*T6)**N26)*((T2+V*T7)**N27)*
((T3+V*T4)**N34)*((T3+V*T5)**N35)*((T3+V*T6)**N36)*((T3+V*T7)**N37)*
((T4+V*T5)**N45)*((T4+V*T6)**N46)*((T4+V*T7)**N47)*((T5+V*T6)**N56)*((T5+V*T7)**N57)*((T6
+V*T7)**N67)*((T2+V*T1)**N21)*
((T3+V*T1)**N31)*((T4+V*T1)**N41)*((T5+V*T1)**N51)*((T6+V*T1)**N61)*((T7+V*T1)**N71)*
((T3+V*T2)**N32)*((T4+V*T2)**N42)*((T5+V*T2)**N52)*((T6+V*T2)**N62)*((T7+V*T2)**N72)*
```

```
((T4+V*T3)**N43)*((T5+V*T3)**N53)*((T6+V*T3)**N63)*((T7+V*T3)**N73)*
((T5+V*T4)**N54)*((T6+V*T4)**N64)*((T7+V*T4)**N74)*
((T6+V*T5)**N65)*((T7+V*T5)**N75)*((T7+V*T6)**N76));
DFT2=PTH*DA; AFT2=DFT2**(2/3);
OUTPUT; END; *KEEP T2 T3 T4 T5 T6 T7 V DFT2 AFT2 QQ;
*RUN;
PROC MEANS DATA=DT2 MAX NOPRINT; VAR AFT2; OUTPUT OUT=DDM MAX=MDT2;
DATA DD1; SET DDM (KEEP=MDT2); QQ=1;
DATA DD2; MERGE DD1 DT2; BY QQ;
ZZ=MDT2-AFT2;
IF ZZ=0; MT2=T2; *KEEP MT2 QQ;
DATA DD3; MERGE DT2 DD2; BY QQ;
YY=T2-MT2; FBB=DFT2**(1/3);
DGM=YY*FBB; *KEEP DGM;
*PROC PRINT DATA=DD3; *RUN;
PROC MEANS DATA=DD3 MIN NOPRINT; VAR DGM; OUTPUT OUT=BBN MIN=MINB;
PROC MEANS DATA=DD3 MAX NOPRINT; VAR DGM; OUTPUT OUT=BBP MAX=MAXB;
DATA DDR1; MERGE BBN BBP DD2 DDR;
DO UNTIL(0<T2<1-T1-T3-T4-T5-T6-T7);
DO UNTIL(UUU<FRT2);
R1=RANUNI(0);    R2=RANUNI(0);
UU=MDT2*R1; VV=MINB+(MAXB-MINB)*R2;
UUU=UU**(3/2); T2=VV/SQRT(UU)+MT2;
FRT2=(((V**2-1)**N)*(T1**N1)*(T2**N2)*(T3**N3)*(T4**N4)*(T5**N5)*(T6**N6)*(T7**N7))/
(((T1+V*T2)**N12)*((T1+V*T3)**N13)*((T1+V*T4)**N14)*((T1+V*T5)**N15)*((T1+V*T6)**N16)*((T
1+V*T7)**N17)*
((T2+V*T3)**N23)*((T2+V*T4)**N24)*((T2+V*T5)**N25)*((T2+V*T6)**N26)*((T2+V*T7)**N27)*
((T3+V*T4)**N34)*((T3+V*T5)**N35)*((T3+V*T6)**N36)*((T3+V*T7)**N37)*
((T4+V*T5)**N45)*((T4+V*T6)**N46)*((T4+V*T7)**N47)*((T5+V*T6)**N56)*((T5+V*T7)**N57)*((T6
+V*T7)**N67)*((T2+V*T1)**N21)*
((T3+V*T1)**N31)*((T4+V*T1)**N41)*((T5+V*T1)**N51)*((T6+V*T1)**N61)*((T7+V*T1)**N71)*
((T3+V*T2)**N32)*((T4+V*T2)**N42)*((T5+V*T2)**N52)*((T6+V*T2)**N62)*((T7+V*T2)**N72)*
((T4+V*T3)**N43)*((T5+V*T3)**N53)*((T6+V*T3)**N63)*((T7+V*T3)**N73)*
((T5+V*T4)**N54)*((T6+V*T4)**N64)*((T7+V*T4)**N74)*
((T6+V*T5)**N65)*((T7+V*T5)**N75)*((T7+V*T6)**N76));
END; END; KEEP T2;
*PROC PRINT DATA=DDR1; RUN;
DATA DT3; MERGE DT1 DDR DDR1; T4=0.15; T5=0.1; T6=0.1; T7=0.05; V=1.23; QQ=1;
DO T3=DA TO 1-T1-T2-T4-T5-T6-T7-0.01 BY 0.01;
PTH=(((V**2-1)**N)*(T1**N1)*(T2**N2)*(T3**N3)*(T4**N4)*(T5**N5)*(T6**N6)*(T7**N7))/
(((T1+V*T2)**N12)*((T1+V*T3)**N13)*((T1+V*T4)**N14)*((T1+V*T5)**N15)*((T1+V*T6)**N16)*((T
1+V*T7)**N17)*
((T2+V*T3)**N23)*((T2+V*T4)**N24)*((T2+V*T5)**N25)*((T2+V*T6)**N26)*((T2+V*T7)**N27)*
((T3+V*T4)**N34)*((T3+V*T5)**N35)*((T3+V*T6)**N36)*((T3+V*T7)**N37)*
((T4+V*T5)**N45)*((T4+V*T6)**N46)*((T4+V*T7)**N47)*((T5+V*T6)**N56)*((T5+V*T7)**N57)*((T6
+V*T7)**N67)*((T2+V*T1)**N21)*
((T3+V*T1)**N31)*((T4+V*T1)**N41)*((T5+V*T1)**N51)*((T6+V*T1)**N61)*((T7+V*T1)**N71)*
((T3+V*T2)**N32)*((T4+V*T2)**N42)*((T5+V*T2)**N52)*((T6+V*T2)**N62)*((T7+V*T2)**N72)*
((T4+V*T3)**N43)*((T5+V*T3)**N53)*((T6+V*T3)**N63)*((T7+V*T3)**N73)*
((T5+V*T4)**N54)*((T6+V*T4)**N64)*((T7+V*T4)**N74)*
((T6+V*T5)**N65)*((T7+V*T5)**N75)*((T7+V*T6)**N76));
DFT3=PTH*DA; AFT3=DFT3**(2/3);
OUTPUT; END; *KEEP T2 T3 T4 T5 T6 T7 V DFT3 AFT3 QQ;
*RUN;
PROC MEANS DATA=DT3 MAX NOPRINT; VAR AFT3; OUTPUT OUT=DDM MAX=MDT3;
DATA DD1; SET DDM (KEEP=MDT3); QQ=1;
DATA DD2; MERGE DD1 DT3; BY QQ;
ZZ=MDT3-AFT3;
IF ZZ=0; MT3=T3; *KEEP MT3 QQ;
DATA DD3; MERGE DT3 DD2; BY QQ;
YY=T3-MT3; FBB=DFT3**(1/3);
DGM=YY*FBB; *KEEP DGM;
*PROC PRINT DATA=DD3; *RUN;
PROC MEANS DATA=DD3 MIN NOPRINT; VAR DGM; OUTPUT OUT=BBN MIN=MINB;
PROC MEANS DATA=DD3 MAX NOPRINT; VAR DGM; OUTPUT OUT=BBP MAX=MAXB;
DATA DDR2; MERGE BBN BBP DD2 DDR DDR1;
DO UNTIL(0<T3<1-T1-T2-T4-T5-T6-T7);
DO UNTIL(UUU<FRT3);
R1=RANUNI(0);    R2=RANUNI(0);
UU=MDT3*R1; VV=MINB+(MAXB-MINB)*R2;
UUU=UU**(3/2); T3=VV/SQRT(UU)+MT3;
FRT3=(((V**2-1)**N)*(T1**N1)*(T2**N2)*(T3**N3)*(T4**N4)*(T5**N5)*(T6**N6)*(T7**N7))/
(((T1+V*T2)**N12)*((T1+V*T3)**N13)*((T1+V*T4)**N14)*((T1+V*T5)**N15)*((T1+V*T6)**N16)*((T
1+V*T7)**N17)*
((T2+V*T3)**N23)*((T2+V*T4)**N24)*((T2+V*T5)**N25)*((T2+V*T6)**N26)*((T2+V*T7)**N27)*
((T3+V*T4)**N34)*((T3+V*T5)**N35)*((T3+V*T6)**N36)*((T3+V*T7)**N37)*
((T4+V*T5)**N45)*((T4+V*T6)**N46)*((T4+V*T7)**N47)*((T5+V*T6)**N56)*((T5+V*T7)**N57)*((T6
+V*T7)**N67)*((T2+V*T1)**N21)*
((T3+V*T1)**N31)*((T4+V*T1)**N41)*((T5+V*T1)**N51)*((T6+V*T1)**N61)*((T7+V*T1)**N71)*
((T3+V*T2)**N32)*((T4+V*T2)**N42)*((T5+V*T2)**N52)*((T6+V*T2)**N62)*((T7+V*T2)**N72)*
((T4+V*T3)**N43)*((T5+V*T3)**N53)*((T6+V*T3)**N63)*((T7+V*T3)**N73)*
```

```
((T5+V*T4)**N54)*((T6+V*T4)**N64)*((T7+V*T4)**N74)*
((T6+V*T5)**N65)*((T7+V*T5)**N75)*((T7+V*T6)**N76));
END; END; KEEP T3;
*PROC PRINT DATA=DDR2; RUN;
DATA DT4; MERGE DT1 DDR DDR1 DDR2; T5=0.1; T6=0.1; T7=0.05; V=1.23; QQ=1;
DO T4=DA TO 1-T1-T2-T3-T5-T6-T7-0.01 BY 0.01;
PTH=(((V**2-1)**N)*(T1**N1)*(T2**N2)*(T3**N3)*(T4**N4)*(T5**N5)*(T6**N6)*(T7**N7))/
(((T1+V*T2)**N12)*((T1+V*T3)**N13)*((T1+V*T4)**N14)*((T1+V*T5)**N15)*((T1+V*T6)**N16)*((T
1+V*T7)**N17)*
((T2+V*T3)**N23)*((T2+V*T4)**N24)*((T2+V*T5)**N25)*((T2+V*T6)**N26)*((T2+V*T7)**N27)*
((T3+V*T4)**N34)*((T3+V*T5)**N35)*((T3+V*T6)**N36)*((T3+V*T7)**N37)*
((T4+V*T5)**N45)*((T4+V*T6)**N46)*((T4+V*T7)**N47)*((T5+V*T6)**N56)*((T5+V*T7)**N57)*((T6
+V*T7)**N67)*((T2+V*T1)**N21)*
((T3+V*T1)**N31)*((T4+V*T1)**N41)*((T5+V*T1)**N51)*((T6+V*T1)**N61)*((T7+V*T1)**N71)*
((T3+V*T2)**N32)*((T4+V*T2)**N42)*((T5+V*T2)**N52)*((T6+V*T2)**N62)*((T7+V*T2)**N72)*
((T4+V*T3)**N43)*((T5+V*T3)**N53)*((T6+V*T3)**N63)*((T7+V*T3)**N73)*
((T5+V*T4)**N54)*((T6+V*T4)**N64)*((T7+V*T4)**N74)*
((T6+V*T5)**N65)*((T7+V*T5)**N75)*((T7+V*T6)**N76));
DFT4=PTH*DA; AFT4=DFT4**(2/3);
OUTPUT; END; *KEEP T2 T3 T4 T5 T6 T7 V DFT4 AFT4 QQ;
*RUN;
PROC MEANS DATA=DT4 MAX NOPRINT; VAR AFT4; OUTPUT OUT=DDM MAX=MDT4;
DATA DD1; SET DDM (KEEP=MDT4); QQ=1;
DATA DD2; MERGE DD1 DT4; BY QQ;
ZZ=MDT4-AFT4;
IF ZZ=0; MT4=T4; *KEEP MT4 QQ;
DATA DD3; MERGE DT4 DD2; BY QQ;
YY=T4-MT4; FBB=DFT4**(1/3);
DGM=YY*FBB; *KEEP DGM;
*PROC PRINT DATA=DD3; *RUN;
PROC MEANS DATA=DD3 MIN NOPRINT; VAR DGM; OUTPUT OUT=BBN MIN=MINB;
PROC MEANS DATA=DD3 MAX NOPRINT; VAR DGM; OUTPUT OUT=BBP MAX=MAXB;
DATA DDR3; MERGE BBN BBP DD2 DDR DDR1 DDR2;
DO UNTIL(0<T4<1-T1-T2-T3-T5-T6-T7);
DO UNTIL(UUU<FRT4);
R1=RANUNI(0);   R2=RANUNI(0);
UU=MDT4*R1; VV=MINB+(MAXB-MINB)*R2;
UUU=UU**(3/2); T4=VV/SQRT(UU)+MT4;
FRT4=(((V**2-1)**N)*(T1**N1)*(T2**N2)*(T3**N3)*(T4**N4)*(T5**N5)*(T6**N6)*(T7**N7))/
(((T1+V*T2)**N12)*((T1+V*T3)**N13)*((T1+V*T4)**N14)*((T1+V*T5)**N15)*((T1+V*T6)**N16)*((T
1+V*T7)**N17)*
((T2+V*T3)**N23)*((T2+V*T4)**N24)*((T2+V*T5)**N25)*((T2+V*T6)**N26)*((T2+V*T7)**N27)*
((T3+V*T4)**N34)*((T3+V*T5)**N35)*((T3+V*T6)**N36)*((T3+V*T7)**N37)*
((T4+V*T5)**N45)*((T4+V*T6)**N46)*((T4+V*T7)**N47)*((T5+V*T6)**N56)*((T5+V*T7)**N57)*((T6
+V*T7)**N67)*((T2+V*T1)**N21)*
((T3+V*T1)**N31)*((T4+V*T1)**N41)*((T5+V*T1)**N51)*((T6+V*T1)**N61)*((T7+V*T1)**N71)*
((T3+V*T2)**N32)*((T4+V*T2)**N42)*((T5+V*T2)**N52)*((T6+V*T2)**N62)*((T7+V*T2)**N72)*
((T4+V*T3)**N43)*((T5+V*T3)**N53)*((T6+V*T3)**N63)*((T7+V*T3)**N73)*
((T5+V*T4)**N54)*((T6+V*T4)**N64)*((T7+V*T4)**N74)*
((T6+V*T5)**N65)*((T7+V*T5)**N75)*((T7+V*T6)**N76));
END; END; KEEP T4;
*PROC PRINT DATA=DDR3; RUN;
DATA DT5; MERGE DT1 DDR DDR1 DDR2 DDR3; T6=0.1; T7=0.05; V=1.23; QQ=1;
DO T5=DA TO 1-T1-T2-T3-T4-T6-T7-0.01 BY 0.01;
PTH=(((V**2-1)**N)*(T1**N1)*(T2**N2)*(T3**N3)*(T4**N4)*(T5**N5)*(T6**N6)*(T7**N7))/
(((T1+V*T2)**N12)*((T1+V*T3)**N13)*((T1+V*T4)**N14)*((T1+V*T5)**N15)*((T1+V*T6)**N16)*((T
1+V*T7)**N17)*
((T2+V*T3)**N23)*((T2+V*T4)**N24)*((T2+V*T5)**N25)*((T2+V*T6)**N26)*((T2+V*T7)**N27)*
((T3+V*T4)**N34)*((T3+V*T5)**N35)*((T3+V*T6)**N36)*((T3+V*T7)**N37)*
((T4+V*T5)**N45)*((T4+V*T6)**N46)*((T4+V*T7)**N47)*((T5+V*T6)**N56)*((T5+V*T7)**N57)*((T6
+V*T7)**N67)*((T2+V*T1)**N21)*
((T3+V*T1)**N31)*((T4+V*T1)**N41)*((T5+V*T1)**N51)*((T6+V*T1)**N61)*((T7+V*T1)**N71)*
((T3+V*T2)**N32)*((T4+V*T2)**N42)*((T5+V*T2)**N52)*((T6+V*T2)**N62)*((T7+V*T2)**N72)*
((T4+V*T3)**N43)*((T5+V*T3)**N53)*((T6+V*T3)**N63)*((T7+V*T3)**N73)*
((T5+V*T4)**N54)*((T6+V*T4)**N64)*((T7+V*T4)**N74)*
((T6+V*T5)**N65)*((T7+V*T5)**N75)*((T7+V*T6)**N76));
DFT5=PTH*DA; AFT5=DFT5**(2/3);
OUTPUT; END; *KEEP T2 T3 T4 T5 T6 T7 V DFT5 AFT5 QQ;
*RUN;
PROC MEANS DATA=DT5 MAX NOPRINT; VAR AFT5; OUTPUT OUT=DDM MAX=MDT5;
DATA DD1; SET DDM (KEEP=MDT5); QQ=1;
DATA DD2; MERGE DD1 DT5; BY QQ;
ZZ=MDT5-AFT5;
IF ZZ=0; MT5=T5; *KEEP MT5 QQ;
DATA DD3; MERGE DT5 DD2; BY QQ;
YY=T5-MT5; FBB=DFT5**(1/3);
DGM=YY*FBB; *KEEP DGM;
*PROC PRINT DATA=DD3; *RUN;
PROC MEANS DATA=DD3 MIN NOPRINT; VAR DGM; OUTPUT OUT=BBN MIN=MINB;
PROC MEANS DATA=DD3 MAX NOPRINT; VAR DGM; OUTPUT OUT=BBP MAX=MAXB;
DATA DDR4; MERGE BBN BBP DD2 DDR DDR1 DDR2 DDR3;
```

```
DO UNTIL(0<T5<1-T1-T2-T3-T4-T6-T7);
DO UNTIL(UUU<FRT5);
R1=RANUNI(0);   R2=RANUNI(0);
UU=MDT5*R1; VV=MINB+(MAXB-MINB)*R2;
UUU=UU**(3/2); T5=VV/SQRT(UU)+MT5;
FRT5=(((V**2-1)**N)*(T1**N1)*(T2**N2)*(T3**N3)*(T4**N4)*(T5**N5)*(T6**N6)*(T7**N7))/
(((T1+V*T2)**N12)*((T1+V*T3)**N13)*((T1+V*T4)**N14)*((T1+V*T5)**N15)*((T1+V*T6)**N16)*((T
1+V*T7)**N17)*
((T2+V*T3)**N23)*((T2+V*T4)**N24)*((T2+V*T5)**N25)*((T2+V*T6)**N26)*((T2+V*T7)**N27)*
((T3+V*T4)**N34)*((T3+V*T5)**N35)*((T3+V*T6)**N36)*((T3+V*T7)**N37)*
((T4+V*T5)**N45)*((T4+V*T6)**N46)*((T4+V*T7)**N47)*((T5+V*T6)**N56)*((T5+V*T7)**N57)*((T6
+V*T7)**N67)*((T2+V*T1)**N21)*
((T3+V*T1)**N31)*((T4+V*T1)**N41)*((T5+V*T1)**N51)*((T6+V*T1)**N61)*((T7+V*T1)**N71)*
((T3+V*T2)**N32)*((T4+V*T2)**N42)*((T5+V*T2)**N52)*((T6+V*T2)**N62)*((T7+V*T2)**N72)*
((T4+V*T3)**N43)*((T5+V*T3)**N53)*((T6+V*T3)**N63)*((T7+V*T3)**N73)*
((T5+V*T4)**N54)*((T6+V*T4)**N64)*((T7+V*T4)**N74)*
((T6+V*T5)**N65)*((T7+V*T5)**N75)*((T7+V*T6)**N76));
END; END; KEEP T5;
*PROC PRINT DATA=DDR4; RUN;
DATA DT6; MERGE DT1 DDR DDR1 DDR2 DDR3 DDR4; T7=0.05; V=1.23; QQ=1;
DO T6=DA TO 1-T1-T2-T3-T4-T5-T7-0.01 BY 0.01;
PTH=(((V**2-1)**N)*(T1**N1)*(T2**N2)*(T3**N3)*(T4**N4)*(T5**N5)*(T6**N6)*(T7**N7))/
(((T1+V*T2)**N12)*((T1+V*T3)**N13)*((T1+V*T4)**N14)*((T1+V*T5)**N15)*((T1+V*T6)**N16)*((T
1+V*T7)**N17)*
((T2+V*T3)**N23)*((T2+V*T4)**N24)*((T2+V*T5)**N25)*((T2+V*T6)**N26)*((T2+V*T7)**N27)*
((T3+V*T4)**N34)*((T3+V*T5)**N35)*((T3+V*T6)**N36)*((T3+V*T7)**N37)*
((T4+V*T5)**N45)*((T4+V*T6)**N46)*((T4+V*T7)**N47)*((T5+V*T6)**N56)*((T5+V*T7)**N57)*((T6
+V*T7)**N67)*((T2+V*T1)**N21)*
((T3+V*T1)**N31)*((T4+V*T1)**N41)*((T5+V*T1)**N51)*((T6+V*T1)**N61)*((T7+V*T1)**N71)*
((T3+V*T2)**N32)*((T4+V*T2)**N42)*((T5+V*T2)**N52)*((T6+V*T2)**N62)*((T7+V*T2)**N72)*
((T4+V*T3)**N43)*((T5+V*T3)**N53)*((T6+V*T3)**N63)*((T7+V*T3)**N73)*
((T5+V*T4)**N54)*((T6+V*T4)**N64)*((T7+V*T4)**N74)*
((T6+V*T5)**N65)*((T7+V*T5)**N75)*((T7+V*T6)**N76));
DFT6=PTH*DA; AFT6=DFT6**(2/3);
OUTPUT; END; *KEEP T2 T3 T4 T5 T6 T7 V DFT6 AFT6 QQ;
*RUN;
PROC MEANS DATA=DT6 MAX NOPRINT; VAR AFT6; OUTPUT OUT=DDM MAX=MDT6;
DATA DD1; SET DDM (KEEP=MDT6); QQ=1;
DATA DD2; MERGE DD1 DT6; BY QQ;
ZZ=MDT6-AFT6;
IF ZZ=0; MT6=T6; *KEEP MT6 QQ;
DATA DD3; MERGE DT6 DD2; BY QQ;
YY=T6-MT6; FBB=DFT6**(1/3);
DGM=YY*FBB; *KEEP DGM;
*PROC PRINT DATA=DD3; *RUN;
PROC MEANS DATA=DD3 MIN NOPRINT; VAR DGM; OUTPUT OUT=BBN MIN=MINB;
PROC MEANS DATA=DD3 MAX NOPRINT; VAR DGM; OUTPUT OUT=BBP MAX=MAXB;
DATA DDR5; MERGE BBN BBP DD2 DDR DDR1 DDR2 DDR3 DDR4;
DO UNTIL(0<T6<1-T1-T2-T3-T4-T5-T7);
DO UNTIL(UUU<FRT6);
R1=RANUNI(0);   R2=RANUNI(0);
UU=MDT6*R1; VV=MINB+(MAXB-MINB)*R2;
UUU=UU**(3/2); T6=VV/SQRT(UU)+MT6;
FRT6=(((V**2-1)**N)*(T1**N1)*(T2**N2)*(T3**N3)*(T4**N4)*(T5**N5)*(T6**N6)*(T7**N7))/
(((T1+V*T2)**N12)*((T1+V*T3)**N13)*((T1+V*T4)**N14)*((T1+V*T5)**N15)*((T1+V*T6)**N16)*((T
1+V*T7)**N17)*
((T2+V*T3)**N23)*((T2+V*T4)**N24)*((T2+V*T5)**N25)*((T2+V*T6)**N26)*((T2+V*T7)**N27)*
((T3+V*T4)**N34)*((T3+V*T5)**N35)*((T3+V*T6)**N36)*((T3+V*T7)**N37)*
((T4+V*T5)**N45)*((T4+V*T6)**N46)*((T4+V*T7)**N47)*((T5+V*T6)**N56)*((T5+V*T7)**N57)*((T6
+V*T7)**N67)*((T2+V*T1)**N21)*
((T3+V*T1)**N31)*((T4+V*T1)**N41)*((T5+V*T1)**N51)*((T6+V*T1)**N61)*((T7+V*T1)**N71)*
((T3+V*T2)**N32)*((T4+V*T2)**N42)*((T5+V*T2)**N52)*((T6+V*T2)**N62)*((T7+V*T2)**N72)*
((T4+V*T3)**N43)*((T5+V*T3)**N53)*((T6+V*T3)**N63)*((T7+V*T3)**N73)*
((T5+V*T4)**N54)*((T6+V*T4)**N64)*((T7+V*T4)**N74)*
((T6+V*T5)**N65)*((T7+V*T5)**N75)*((T7+V*T6)**N76));
END; END; KEEP T6;
*PROC PRINT DATA=DDR5; RUN;
DATA DT7; MERGE DT1 DDR DDR1 DDR2 DDR3 DDR4 DDR5; V=1.23; QQ=1;
DO T7=DA TO 1-T1-T2-T3-T4-T5-T6-0.01 BY 0.01;
PTH=(((V**2-1)**N)*(T1**N1)*(T2**N2)*(T3**N3)*(T4**N4)*(T5**N5)*(T6**N6)*(T7**N7))/
(((T1+V*T2)**N12)*((T1+V*T3)**N13)*((T1+V*T4)**N14)*((T1+V*T5)**N15)*((T1+V*T6)**N16)*((T
1+V*T7)**N17)*
((T2+V*T3)**N23)*((T2+V*T4)**N24)*((T2+V*T5)**N25)*((T2+V*T6)**N26)*((T2+V*T7)**N27)*
((T3+V*T4)**N34)*((T3+V*T5)**N35)*((T3+V*T6)**N36)*((T3+V*T7)**N37)*
((T4+V*T5)**N45)*((T4+V*T6)**N46)*((T4+V*T7)**N47)*((T5+V*T6)**N56)*((T5+V*T7)**N57)*((T6
+V*T7)**N67)*((T2+V*T1)**N21)*
((T3+V*T1)**N31)*((T4+V*T1)**N41)*((T5+V*T1)**N51)*((T6+V*T1)**N61)*((T7+V*T1)**N71)*
((T3+V*T2)**N32)*((T4+V*T2)**N42)*((T5+V*T2)**N52)*((T6+V*T2)**N62)*((T7+V*T2)**N72)*
((T4+V*T3)**N43)*((T5+V*T3)**N53)*((T6+V*T3)**N63)*((T7+V*T3)**N73)*
((T5+V*T4)**N54)*((T6+V*T4)**N64)*((T7+V*T4)**N74)*
((T6+V*T5)**N65)*((T7+V*T5)**N75)*((T7+V*T6)**N76));
```

```
DFT7=PTH*DA; AFT7=DFT7**(2/3);
OUTPUT; END; *KEEP T2 T3 T4 T5 T6 T7 V DFT7 AFT7 QQ;
*RUN;
PROC MEANS DATA=DT7 MAX NOPRINT; VAR AFT7; OUTPUT OUT=DDM MAX=MDT7;
DATA DD1; SET DDM (KEEP=MDT7); QQ=1;
DATA DD2; MERGE DD1 DT7; BY QQ;
ZZ=MDT7-AFT7;
IF ZZ=0; MT7=T7; *KEEP MT7 QQ;
DATA DD3; MERGE DT7 DD2; BY QQ;
YY=T7-MT7; FBB=DFT7**(1/3);
DGM=YY*FBB; *KEEP DGM;
*PROC PRINT DATA=DD3; *RUN;
PROC MEANS DATA=DD3 MIN NOPRINT; VAR DGM; OUTPUT OUT=BBN MIN=MINB;
PROC MEANS DATA=DD3 MAX NOPRINT; VAR DGM; OUTPUT OUT=BBP MAX=MAXB;
DATA DDR6; MERGE BBN BBP DD2 DDR DDR1 DDR2 DDR3 DDR4 DDR5;
DO UNTIL(0<T7<1-T1-T2-T3-T4-T5-T6);
DO UNTIL(UUU<FRT7);
R1=RANUNI(0);   R2=RANUNI(0);
UU=MDT7*R1; VV=MINB+(MAXB-MINB)*R2;
UUU=UU**(3/2); T7=VV/SQRT(UU)+MT7;
FRT7=(((V**2-1)**N)*(T1**N1)*(T2**N2)*(T3**N3)*(T4**N4)*(T5**N5)*(T6**N6)*(T7**N7))/
(((T1+V*T2)**N12)*((T1+V*T3)**N13)*((T1+V*T4)**N14)*((T1+V*T5)**N15)*((T1+V*T6)**N16)*((T
1+V*T7)**N17)*
((T2+V*T3)**N23)*((T2+V*T4)**N24)*((T2+V*T5)**N25)*((T2+V*T6)**N26)*((T2+V*T7)**N27)*
((T3+V*T4)**N34)*((T3+V*T5)**N35)*((T3+V*T6)**N36)*((T3+V*T7)**N37)*
((T4+V*T5)**N45)*((T4+V*T6)**N46)*((T4+V*T7)**N47)*((T5+V*T6)**N56)*((T5+V*T7)**N57)*((T6
+V*T7)**N67)*((T2+V*T1)**N21)*
((T3+V*T1)**N31)*((T4+V*T1)**N41)*((T5+V*T1)**N51)*((T6+V*T1)**N61)*((T7+V*T1)**N71)*
((T3+V*T2)**N32)*((T4+V*T2)**N42)*((T5+V*T2)**N52)*((T6+V*T2)**N62)*((T7+V*T2)**N72)*
((T4+V*T3)**N43)*((T5+V*T3)**N53)*((T6+V*T3)**N63)*((T7+V*T3)**N73)*
((T5+V*T4)**N54)*((T6+V*T4)**N64)*((T7+V*T4)**N74)*
((T6+V*T5)**N65)*((T7+V*T5)**N75)*((T7+V*T6)**N76));
END; END; KEEP T7;
*PROC PRINT DATA=DDR6; RUN;
DATA DT8; MERGE  DT1 DDR DDR1 DDR2 DDR3 DDR4 DDR5 DDR6; QQ=1;
DO V=1-0.01 TO 4-0.01 BY 0.01;
PTH=(((V**2-1)**N)*(T1**N1)*(T2**N2)*(T3**N3)*(T4**N4)*(T5**N5)*(T6**N6)*(T7**N7))/
(((T1+V*T2)**N12)*((T1+V*T3)**N13)*((T1+V*T4)**N14)*((T1+V*T5)**N15)*((T1+V*T6)**N16)*((T
1+V*T7)**N17)*
((T2+V*T3)**N23)*((T2+V*T4)**N24)*((T2+V*T5)**N25)*((T2+V*T6)**N26)*((T2+V*T7)**N27)*
((T3+V*T4)**N34)*((T3+V*T5)**N35)*((T3+V*T6)**N36)*((T3+V*T7)**N37)*
((T4+V*T5)**N45)*((T4+V*T6)**N46)*((T4+V*T7)**N47)*((T5+V*T6)**N56)*((T5+V*T7)**N57)*((T6
+V*T7)**N67)*((T2+V*T1)**N21)*
((T3+V*T1)**N31)*((T4+V*T1)**N41)*((T5+V*T1)**N51)*((T6+V*T1)**N61)*((T7+V*T1)**N71)*
((T3+V*T2)**N32)*((T4+V*T2)**N42)*((T5+V*T2)**N52)*((T6+V*T2)**N62)*((T7+V*T2)**N72)*
((T4+V*T3)**N43)*((T5+V*T3)**N53)*((T6+V*T3)**N63)*((T7+V*T3)**N73)*
((T5+V*T4)**N54)*((T6+V*T4)**N64)*((T7+V*T4)**N74)*
((T6+V*T5)**N65)*((T7+V*T5)**N75)*((T7+V*T6)**N76));
DFT8=PTH*DA; AFT8=DFT8**(2/3);
OUTPUT; END; *KEEP T1 T2 T3 T4 T5 T6 T7 V DFT8 AFT8 QQ;
*RUN;
PROC MEANS DATA=DT8 MAX NOPRINT; VAR AFT8; OUTPUT OUT=DDM MAX=MDT8;
DATA DD1; SET DDM (KEEP=MDT8); QQ=1;
DATA DD2; MERGE DD1 DT8; BY QQ;
ZZ=MDT8-AFT8;
IF ZZ=0; MT8=V; *KEEP MT8 QQ;
DATA DD3; MERGE DT8 DD2; BY QQ;
YY=V-MT8; FBB=DFT8**(1/3);
DGM=YY*FBB; *KEEP DGM;
*PROC PRINT DATA=DD3; *RUN;
PROC MEANS DATA=DD3 MIN NOPRINT; VAR DGM; OUTPUT OUT=BBN MIN=MINB;
PROC MEANS DATA=DD3 MAX NOPRINT; VAR DGM; OUTPUT OUT=BBP MAX=MAXB;
DATA DDR7; MERGE BBN BBP DD2 DDR DDR1 DDR2 DDR3 DDR4 DDR5 DDR6;
DO UNTIL(1<V<4);
DO UNTIL(UUU<FRT8);
R1=RANUNI(0);   R2=RANUNI(0);
UU=MDT8*R1; VV=MINB+(MAXB-MINB)*R2;
UUU=UU**(3/2); V=VV/SQRT(UU)+MT8;
FRT8=(((V**2-1)**N)*(T1**N1)*(T2**N2)*(T3**N3)*(T4**N4)*(T5**N5)*(T6**N6)*(T7**N7))/
(((T1+V*T2)**N12)*((T1+V*T3)**N13)*((T1+V*T4)**N14)*((T1+V*T5)**N15)*((T1+V*T6)**N16)*((T
1+V*T7)**N17)*
((T2+V*T3)**N23)*((T2+V*T4)**N24)*((T2+V*T5)**N25)*((T2+V*T6)**N26)*((T2+V*T7)**N27)*
((T3+V*T4)**N34)*((T3+V*T5)**N35)*((T3+V*T6)**N36)*((T3+V*T7)**N37)*
((T4+V*T5)**N45)*((T4+V*T6)**N46)*((T4+V*T7)**N47)*((T5+V*T6)**N56)*((T5+V*T7)**N57)*((T6
+V*T7)**N67)*((T2+V*T1)**N21)*
((T3+V*T1)**N31)*((T4+V*T1)**N41)*((T5+V*T1)**N51)*((T6+V*T1)**N61)*((T7+V*T1)**N71)*
((T3+V*T2)**N32)*((T4+V*T2)**N42)*((T5+V*T2)**N52)*((T6+V*T2)**N62)*((T7+V*T2)**N72)*
((T4+V*T3)**N43)*((T5+V*T3)**N53)*((T6+V*T3)**N63)*((T7+V*T3)**N73)*
((T5+V*T4)**N54)*((T6+V*T4)**N64)*((T7+V*T4)**N74)*
((T6+V*T5)**N65)*((T7+V*T5)**N75)*((T7+V*T6)**N76));
END; END; KEEP V;*PROC PRINT DATA=DDR7; RUN;
```