

BAYESIAN ANALYSIS OF THE VAN BAAREN MODEL FOR PAIRED COMPARISON

Saima Altaf ^{*}, Muhammad Aslam[†], Muhammad Aslam[‡]

Received 23:09:2010 : Accepted 10:04:2012

Abstract

The technique of paired comparison is being commonly studied these days because of its attractive applications for the comparison of several objects, simultaneously. This technique permits the ranking of the objects by means of a score, which reflects the merit of the items on a linear scale. The present study is concerned with the Bayesian analysis of a paired comparison model, namely the van Baaren model VI using the informative and the conjugate priors. For this purpose, an inclusive elicitation technique to evaluate the hyperparameters of the prior distributions has also been elaborated. The joint posterior distribution for the parameters of the model, their marginal distributions and their inferences are obtained via programming in the SAS package. The model is also tested for its appropriateness.

Keywords: Bayesian hypothesis testing; Conjugate prior; Informative prior; Posterior distribution; Predictive probability.

2000 AMS Classification: 62C10, 62C12.

1. Introduction

When the objects that can be scored on the same scale, are compared subjectively, they are ranked on the basis of the scores. In some cases, especially when more than two objects are being compared simultaneously, it is not possible to assign the score to every object on the same scale. In such circumstances, the technique of paired comparison comes to rescue. In this method, the treatments are presented in pairs to one or more judges who in the simplest situation, choose one from the pair or simply just have no preference. Wherever sensory testing is involved, this method has its frequent applications. It is used in taste testing, in professional and intercollegiate sports competitions, market

^{*}Department of Statistics, PMAS University of Arid Agriculture, Rawalpindi, Pakistan
Email: nodupk@yahoo.com

[†]Department of Statistics, Quaid-i-Azam University, Islamabad, Pakistan Email:
aslamsdqu@yahoo.com

[‡]Department of Statistics, Bahauddin Zakariya University, Multan, Pakistan Email:
aslamasadi@bzu.edu.pk

research, voting systems, product comparisons performed by the consumers, multidimensional scaling in personal ratings and generally where the study of choice behavior is concerned. This technique has also worked surprisingly well for the environmental and ecosystem issues (see e.g., Neuman and Watson, 1993). The paired comparison techniques have also been used in the road safety problems and in the quantification of motor vehicle driver's crash risks. Li and Kim (2000) apply the extended Bradley-Terry model for paired comparison to estimate the motor vehicle crash risks using only the crash data.

In the past two centuries, there is a dramatic growth in research and applications of the Bayesian methods. Bayesian methods which are widely accepted and used now a days, provide a theory of inference which enables us to narrate the results of observation with hypothetical predictions. The great steps, which the Bayesian statistics have taken in the recent years, have led the statisticians to focus their attention on the Bayesian analysis of different paired comparison models. Many of the statisticians have studied this technique in detail with varying perceptions due to its sensible and convenient nature and have performed Bayesian analyses of many of the paired comparison models. These statisticians include Aslam (2002, 2003, and 2005), Davidson and Solomon (1973), Glickman (2008), Kim (2005), Kim and Kim (2004), Merrick et al. (2005) and Szwed et al. (2006).

van Baaren (1978) has presented six different extensions of the Bradley-Terry paired comparison model. These extensions differ only in the ways the ties are treated. van Baaren model VI is one of the famous paired comparison models used for the comparison of several objects pairwise, simultaneously. The model takes account of the no preference category as well as order effect factor when two treatments are being compared. The model is practical in the situations where the characteristics can only be observed subjectively instead of being measured and also in different sports competitions like chess tournaments where the result of each game is quantified by stating that either it has ended in a win for one of the players or in a draw. In this model the probability of the tie depends not only on the ratio of the preference parameters but also on their actual level. In the chess tournaments, when there is a tendency among the stronger players to mutually agree to a draw or tie at a relatively early stage of the game, the model can be used. In such a case the number of ties between the objects with high preference parameters is increased and so is the estimated probability of a draw.

In the current study, the Bayesian analysis of van Baaren model VI is presented. Section 2 is about the notations and likelihood of the model. Sections 3 and 4 deal with the Bayesian analysis of the model using the informative and conjugate priors, respectively. Section 5 is reserved to test the appropriateness of the model through the χ^2 statistic while Section 6 is for conclusion.

2. The van Baaren Model VI for Paired Comparison

van Baaren (1978) proposes and compares six extensions of the Bradley-Terry (1952) paired comparison model. According to van Baaren, all the six extensions differ only in the way the ties are treated if the order effect parameter is kept constant. In model I, the probability of the tie is the same for all the preference parameters whose sum is constant. This property makes the model far from being realistic. For models II and III, the probability of a tie is the same for all the pairs of preference parameters for which the ratio, the larger over the smaller say is constant. In models IV, V and VI, the probability of a tie depends, in addition to the ratio of the preference parameters, on their actual level also. In addition to this, models I, IV and V are considered to fit better when the comparison are being made by visual means. On the other hand, model VI is considered to fit best in case of tournaments and matches such as chess tournaments. Furthermore,

this model is simpler and easy to handle as compared to models III, IV and V. We choose model VI for the Bayesian analysis because of its handy and useful approach for sensory paired comparison experiments. According to the proposed model VI, the probability that the preference of treatment T_i over treatment T_j when T_i is presented first is denoted by $\psi_{ij}(1)$ and is defined as:

$$(2.1) \quad \psi_{ij}(1) = \frac{\gamma\theta_i}{\gamma\theta_i + \theta_j + v\theta_i\theta_j}$$

where $\gamma(> 0)$ is the multiplicative order effect parameter, $v(> 0)$ is the tie parameter and $0 < \theta_i < 1$. The preference probability of treatment T_j when T_i is presented first is denoted by $\psi_{ij}(2)$ and is defined as:

$$(2.2) \quad \psi_{ij}(2) = \frac{\theta_j}{\gamma\theta_i + \theta_j + v\theta_i\theta_j}, 0 < \theta_i < 1, \gamma, v > 0.$$

The probability for no preference, which is proportional to the product of both the treatment parameters, is given by:

$$(2.3) \quad P(i \approx j|i, j) = \psi_{ij}(0) = \frac{v\theta_i\theta_j}{\gamma\theta_i + \theta_j + v\theta_i\theta_j}, 0 < \theta_i < 1, \gamma, v > 0.$$

We define notations for the van Baaren model VI;

- $w_{ijk}(1) = 1$ or 0 , accordingly as the treatment T_i is preferred to the treatment T_j when the treatment T_i is presented first in the k 'th repetition of the comparison.
- $w_{ijk}(2) = 1$ or 0 , accordingly as the treatment T_j is preferred to the treatment T_i when the treatment T_i is presented first in the k 'th repetition of the comparison.
- $w_{jik}(1) = 1$ or 0 , accordingly as the treatment T_j is preferred to the treatment T_i when the treatment T_j is presented first in the k 'th repetition of the comparison.
- $w_{jik}(2) = 1$ or 0 , accordingly as the treatment T_i is preferred to the treatment T_j when the treatment T_j is presented first in the k 'th repetition of the comparison.
- $t_{ijk} = 1$ or 0 , accordingly as the treatment T_i is tied with the treatment T_j when T_i is presented first in the k 'th repetition.
- $t_{jik} = 1$ or 0 , accordingly as the treatment T_i is tied with the treatment T_j when T_j is presented first in the k 'th repetition.

The total number of independent comparisons for the pair (i, j) is given by r_{ij} . Let $w_{ij}(1)$ is the number of preferences for the object presented first, $w_{ij}(2)$ be the number of preferences for the object presented second (object j) and t_{ij} are the number of no preferences, $r_{ij} = w_{ij}(1) + w_{ij}(2) + t_{ij}$ and $r_{ji} = w_{ji}(1) + w_{ji}(2) + t_{ji}$, when the order of presentation is (j, i) .

The probability of the observed result in the k^{th} repetition of the pair of treatments (T_i, T_j) is:

$$(2.4) \quad P_{ijk} = \frac{(\gamma\theta_i)^{w_{ijk}(1)}\theta_j^{w_{ijk}(2)}(\gamma\theta_j)^{w_{jik}(1)}\theta_i^{w_{jik}(2)}(v\theta_i\theta_j)^{t_{ijk}+t_{jik}}}{(\gamma\theta_i + \theta_j + v\theta_i\theta_j)^{r_{ijk}}(\theta_i + \gamma\theta_j + v\theta_i\theta_j)^{r_{jik}}}$$

To study and analyze the model, Bayesian analysis of the model has been carried out using the informative and conjugate priors.

3. Bayesian Analysis of the Model using the Informative Prior

The Bayesian approach allows the use of objective data or subjective opinion in specifying a prior distribution. Practical implementations of subjective Bayesian approach provide the opportunity for new and workable methods of prior distribution assessment. An informative prior expresses proper, specific and definite information about a variable.

This prior information quantified in terms of the prior distribution depends on the hyperparameters which are the parameters specified and elicited with the help of available prior information.

Elicitation is the process of extracting experts' knowledge about some parameter of interest, or the probability of some future event and also the quantification of this prior information accurately, which then supplements the given data. Kadane and Wolfson (1998) emphasize the importance of expert opinion in the elicitation of the prior distribution. The hyperparameters for the prior distributions of the van Baaren model VI are also elicited using the experts' opinion. Berger (1985) gives a description of numerous different methods for the elicitation of prior distribution. For different sampling models, different methods for specification of opinions have been developed. According to Aslam (2003), the method of assessment is to compare the predictive distribution with experts' assessment about this distribution and then to choose the hyperparameters that make the assessment agree closely with the member of the family. He discusses three important methods to elicit the hyperparameters. According to the second method, the prior predictive distribution is used for the elicitation of the hyperparameters which is compared with the experts' judgment about this distribution and then the hyperparameters are chosen in such a way so as to make the judgment agree closely as possible with the given distribution. The hyperparameters are elicited via elicitation of confidence levels (a confidence level is a probability for a given interval). Before eliciting the hyperparameters, the specification of the prior distributions of the parameters of the model is required. This depends upon the supports of the parameters. Therefore, Dirichlet distribution is supposed to be the prior distribution for the treatments parameters and the gamma distribution is assumed to be the prior distribution for both the parameters γ and v . Let $\theta = (\theta_1, \dots, \theta_m)$ be the vector of m parameters which sums up to unity for identification i.e., $\sum_{i=1}^m \theta_i = 1$, and considering that the parameters are independent, the prior distribution of $\theta_1, \theta_2, \dots, \theta_m, \gamma$ and v may be the Dirichlet-gamma-gamma distribution:

$$(3.1) \quad p(\theta_1, \theta_2, \dots, \theta_m, \gamma, v) = \frac{b_1^{c_1} c_1^{c_1} \gamma^{c_1-1} v^{b_1-1} e^{-c_2 \gamma} e^{-b_2 v}}{B(a_1, a_2, \dots, a_m) \Gamma(b_1) \Gamma(c_1)} \prod_{i=1}^m \theta_i^{a_i-1}$$

$$\theta_i \geq 0, i = 1, 2, \dots, m, \sum_{i=1}^m \theta_i = 1, \gamma, v > 0,$$

where $a_1, a_2, \dots, a_m, b_1, b_2, c_1$, and c_2 are the hyperparameters.

The likelihood function of the van Baaren model VI is given as:

$$(3.2) \quad l(\mathbf{x}; \theta_1, \dots, \theta_m, \gamma, v) = \prod_{i \neq j=1}^m \prod_{j=1}^m P_{ijk}$$

$$= \prod_{i \neq j}^m \prod_j \frac{r_{ij}!}{w_{ij}(1)! w_{ij}(2)! t_{ij}!} \left[\frac{(\theta_i^{g_i} \gamma^K v^T)}{(\gamma \theta_i + \theta_j + v \theta_i \theta_j)^{r_{ij}}} \right]$$

$$0 < \theta_i < 1; i = 1, 2, \dots, m, \gamma, v > 0,$$

where $r_{ij} = w_{ij}(1) + w_{ij}(2) + t_{ij}$. The total number of preferences for the treatment presented first is $K = \sum_{i=1}^m \sum_{j \neq i}^m w_{ij}(1)$, the total number of ties is $T = \sum_{i=1}^m \sum_{j \neq i}^m t_{ij}$ and $g_i = w_i + t_i$. where $w_i = \sum_{j=1}^m \{w_{ij}(1) + w_{ji}(2)\}$ is the total number of wins and $t_i = \sum_{j=1}^m \{t_{ij} + t_{ji}\}$ is the total number of ties for the i^{th} treatment.

The joint posterior distribution of $\theta_1, \theta_2, \dots, \theta_m, \gamma$, and v becomes:

$$(3.3) \quad p(\theta_1, \theta_2, \dots, \theta_m, \gamma, v | \mathbf{x}) \propto \prod_{i \neq j=1}^m \prod_{j=1}^m \frac{\theta_i^{g_i+a_i-1} \gamma^{K+c_1-1} v^{T+b_1-1} e^{-c_2 \gamma} e^{-b_2 v}}{(\gamma \theta_i + \theta_j + v \theta_i \theta_j)^{r_{ij}}},$$

$$\theta_i \geq 0, i = 1, 2, \dots, m, \sum_{i=1}^m \theta_i = 1, \gamma, v > 0,$$

The marginal posterior distribution for θ_1 when m treatments are being compared, is:

$$p(\theta_1|\mathbf{x}) = \frac{\theta_1^{g_1+a_1-1}}{q} \int_{\theta_2=0}^{1-\theta_1} \dots \int_{\theta_{m-1}=0}^{1-\theta_1 \dots -\theta_{m-2}} \int_{\gamma=0}^{\infty} \int_{v=0}^{\infty} \prod_{i \neq j=1}^{m-1} \prod_j \frac{\theta_i^{g_i+a_i-1} \gamma^{K+c_1-1} v^{T+b_1-1} e^{-c_2\gamma} e^{-b_2v}}{(\gamma\theta_i + \theta_j + v\theta_i\theta_j)^{r_{ij}}} dv d\gamma d\theta_{m-1} \dots d\theta_2,$$

$$0 < \theta_1 < 1, \gamma, v > 0,$$

where q is the normalizing constant. Similarly, the other marginal posterior distributions can be derived.

3.1. Elicitation of the Hyperparameters. As mentioned earlier that the method of eliciting the hyperparameters is via the elicitation of confidence levels which is the probability for a given credible interval. These levels of the given prior predictive distribution are elicited for particular credible intervals of $\{w_{ij}(1), t_{ij}$ and $w_{ji}(1), t_{ji}\}$. To elicit the hyperparameters via confidence levels, the following function (3.5) is minimized. The values, for which this function is minimized, are considered to be the values of the elicited hyperparameters.

$$(3.5) \quad \zeta(a_1, \dots, a_m, b_1, b_2, c_1, c_2) = \min_{a_1, \dots, a_m, b_1, b_2, c_1, c_2} \sum_{h=1}^k |(CCL)_h - (ECL)_h|,$$

where k is the number of credible intervals considered in the elicitation, CCL is the confidence level characterized by the hyperparameters and ECL is the elicited confidence level.

3.2. The Prior Predictive Distribution of the Model. For elicitation of the hyperparameters, the same function (3.5) is minimized. The prior predictive distribution $\{p(w_{ij}(1), t_{ij}, w_{ji}(1), t_{ji})\}$ for the number of times the treatment T_i is preferred to and tied up with the treatment T_j , when a pair of the treatments (T_i, T_j) are being compared under both the orders of presentation with the r_{ij} and r_{ji} numbers of times is:

$$p(w_{ij}(1), t_{ij}, w_{ji}(1), t_{ji}) = Q \int_{\theta_i=0}^1 \int_{\gamma=0}^{\infty} \int_{v=0}^{\infty} \frac{\theta_i^{g_i+a_i-1} (1-\theta_i)^{g_j+a_j-1} \gamma^{K+c_1-1} v^{T+b_1-1} e^{-c_2\gamma} e^{-b_2v}}{\{\theta_i + \gamma(1-\theta_i) + v\sqrt{\theta_i(1-\theta_i)}\}^{r_{ij}} \{\gamma\theta_i + (1-\theta_i) + v\sqrt{\theta_i(1-\theta_i)}\}^{r_{ji}}} dv d\gamma d\theta_i,$$

$$w_{ij}(1), t_{ij} = 0, 1, \dots, r_{ij}, w_{ji}(2), t_{ji} = 0, 1 \dots, r_{ji},$$

where $Q = \frac{r_{ij}! r_{ji}! B(a_i, a_j) b_2^{b_1} c_2^{c_1}}{w_{ij}(1)! w_{ij}(2)! t_{ij}! w_{ji}(1)! w_{ji}(2)! t_{ji}! \Gamma(b_1) \Gamma(c_1)}$, $w_{ij}(2) = r_{ij} - w_{ij}(1) - t_{ij}$, a_i, a_j, b_1, b_2, c_1 and c_2 are the hyperparameters.

For the elicitation of the hyperparameters, r_{ij} 's are taken as the independent responses for the pair (i, j) and r_{ji} as the independent responses for the pair (j, i) . To obtain the values of the hyperparameters, we assume a balanced design that the number of comparisons for each pair, when both the orders of presentation are being considered, is equal. The confidence levels (ECL) are assumed for each pair for different assumed credible intervals of $w_{ij}(1)$ and t_{ij} and $w_{ji}(1)$ and t_{ji} . The function(3.5) is minimized for these credible intervals to make the difference between ECL and CCL as small as possible.

For the elicitation of the hyperparameters for $m = 3$, a program is run in the SAS package (programs can be obtained from the author) with the assumed credible intervals and levels for equal number of comparison for each pair of treatments i.e. ($r_{ij} = r_{ji} = 30$); ($i \neq j = 1, 2, 3$), for different intervals of $w_{ij}(1)$, and t_{ij} , ($i \neq j = 1, 2, 3$) for the order (i, j) and $w_{ji}(1)$ and t_{ji} , ($i \neq j = 1, 2, 3$) for the order (j, i) .

For example, let $r_{12} = 30$, then $w_{12}(1) = 11$ to 15 (that the treatment T_1 is preferred to the treatment T_2 from 11 to 15 times when the treatment T_1 is presented first) and $t_{12} = 2$ to 4 (that both the treatments end up in a tie from 2 to 4 times when the treatment T_1 is presented first).

Similarly, for the pair $(2, 1)$, $r_{21} = 30$, $w_{21}(1) = 14$ to 15 (that the treatment T_2 is preferred to the treatment T_1 from 14 to 15 times when the treatment T_2 is presented first) and $t_{21} = 2$ to 3 (that both the treatments end up in a tie from 2 to 3 times when T_2 is presented first).

The assumed confidence level[§] for the credible intervals of $w_{12}(1)$ and t_{12} and $w_{21}(1)$ and t_{21} for both the order $(1, 2)$ and $(2, 1)$ will be:

$$(ECL)_1 = 0.06$$

Similarly, the confidence level assumed for the credible intervals of $w_{13}(1)$ and t_{13} and $w_{31}(1)$ and t_{31} can be given as:

$$(ECL)_2 = 0.01.$$

The following numbers of credible intervals are considered for the minimization of the function given in (3.5) for these confidence levels:

$$\begin{aligned} \sum_{w_{12}(1)=11}^{15} \sum_{t_{12}=2}^4 \sum_{w_{21}(1)=14}^{15} \sum_{t_{21}=2}^3 p(w_{12}(1), t_{12}, w_{21}(1), t_{21}) &= 0.06 = (ECL)_1, \\ \sum_{w_{13}(1)=14}^{15} \sum_{t_{13}=1}^2 \sum_{w_{31}(1)=13}^{14} \sum_{t_{31}=1}^3 p(w_{13}(1), t_{13}, w_{31}(1), t_{31}) &= 0.01 = (ECL)_2, \\ \sum_{w_{23}(1)=17}^{18} \sum_{t_{23}=1}^2 \sum_{w_{32}(1)=9}^{11} \sum_{t_{32}=1}^2 p(w_{23}(1), t_{23}, w_{32}(1), t_{32}) &= 0.02 = (ECL)_3, \end{aligned}$$

where $\{p(w_{ij}(1), t_{ij}, w_{ji}(1), t_{ji})\}$ is the prior predictive distribution and 0.06, 0.01 and 0.02 are the 1st, 2nd and 3rd elicited levels. Having run the program, the minimum value of the function (3.5) is obtained to be 0.0899 for the elicited values of the hyperparameters $a_1, a_2, a_3, b_1, b_2, c_1$ and c_2 which are 10.56, 10.56, 9.52, 3.98, 10.54, 5.51 and 2.54, respectively.

Having the hyperparameters elicited, now the joint posterior distribution and the marginal posterior distributions for all the parameters can be derived from (3.3). For the complete Bayesian analysis of the model using the informative prior, we need data to carry out the effective procedures. The data for the analysis is taken from Davidson and Beaver (1977) given in Table 3.1 which is of packaged food mixes, with different number of comparisons for each pair.

The posterior means of $\theta_1, \theta_2, \theta_3\gamma$ and v are obtained via a quadrature method and are found to be 0.3361, 0.4197, 0.2442, 2.2993 and 1.0661, respectively. Posterior mode is the most probable value of the parameters. The posterior modes have been obtained in the SAS package by sorting out the highest posterior density with a help of a SAS

[§]The confidence levels are assumed to be very small as we are considering many intervals at a time.

TABLE 1. Responses for the Preference Testing Experiment for $m = 3$

Pairs(i, j)	r_{ij}	$w_{ij}(1)$	$w_{ij}(2)$	t_{ij}
(1,2)	42	23	11	8
(2,1)	43	29	6	8
(1,3)	43	27	11	5
(3,1)	42	22	14	6
(2,3)	41	34	6	1
(3,2)	42	23	16	3

program. The values of the modes of the parameters are attained to be 0.3387, 0.4201, 0.2421, 2.3075 and 1.0769, respectively.

The results (posterior means and modes) imply that the treatment T_2 is better than the other two treatments. The treatment T_1 is the next preferred one and T_3 is the least favored one.

To test the hypotheses: $H_{ij} : \theta_i > \theta_j$ and $H_{ij}^c : \theta_j \geq \theta_i, i \neq j = 1, 2, 3$, the posterior probability for $H_{ij}, p_{ij} = p(\theta_i > \theta_j)$ is attained using a transformation: $\varphi = \theta_i - \theta_j$ and $\xi = \theta_i$ so the posterior probability p_{ij} is:

$$(3.7) \quad p_{ij} = p(\varphi > 0|\mathbf{x}) = \int_{\varphi=0}^1 \int_{\xi=\varphi}^{(1+\varphi)/2} \int_{\gamma=0}^{\infty} \int_{v=0}^{\infty} p(\varphi, \xi, \gamma, v|\mathbf{x}) dv d\gamma d\xi d\varphi.$$

The program is run in the SAS package and the posterior probabilities p_{ij} and q_{ij} are obtained and given in the Table 3.2.

TABLE 2. Posterior Probabilities using the informative (D-G-G) Prior

Hypotheses	p_{ij}	Hypotheses	q_{ij}
$H_{12} : \theta_1 > \theta_2$	0.0586	$H_{12}^c : \theta_2 \geq \theta_1$	0.9432
$H_{13} : \theta_1 > \theta_3$	0.9129	$H_{13}^c : \theta_3 > \theta_1$	0.0871
$H_{23} : \theta_2 > \theta_3$	0.9970	$H_{23}^c : \theta_3 > \theta_2$	0.0030

To test the above mentioned hypotheses, we apply the rule given by Aslam (1996). Let $s = \min(p_{ij}, q_{ij})$. If p_{ij} is small then H_{ij}^c is accepted with high probability. If q_{ij} is small then H_{ij} is accepted with high probability. If ' s ' is small, we can reject one hypothesis otherwise if $s > 0.1$ then the evidence is inconclusive. Thus, using the same criteria, we test the hypotheses. The hypotheses H_{12}^c and H_{23} are accepted showing the preference of the treatment T_2 . For H_{13} , the evidence is conclusive. H_{13} is accepted making evident the treatment T_1 is considered to be preferred to T_3 . The ranking of the treatments obtained via testing of the hypotheses is in complete agreement with that obtained through the posterior estimates.

The predictive probability is the probability that the treatment T_1 is preferred to the treatment T_2 in the future single comparison of these two treatments. For the van Baaren model VI, the predictive probability for the preference of T_1 , when T_1 is presented first after applying the constraint $\theta_3 = 1 - \theta_1 - \theta_2$, is:

$$(3.8) \quad P_{12}(1) = \int_{\theta_1=0}^1 \int_{\theta_2=0}^{1-\theta_1} \int_{\gamma=0}^{\infty} \int_{v=0}^{\infty} \psi_{12}(1) p(\theta_1, \theta_2, \gamma, v|\mathbf{x}) dv d\gamma d\theta_2 d\theta_1,$$

$$\theta_i \geq 0, i = 1, 2, 3, \sum_{i=1}^3 \theta_i = 1, \gamma, v > 0,$$

where $\psi_{12}(1)$ is the preference probability of the treatment T_1 upon the treatment T_2 given in (2.1) and $p(\theta_1, \theta_2, \gamma, v|\mathbf{x}) = \prod_{i \neq j=1}^2 \prod_{j=1}^2 \frac{\theta_i^{g_i} \gamma^K v^T}{q(\gamma\theta_i + \theta_j + v\theta_i\theta_j)^{r_{ij}}}$ is the posterior distribution.

The predictive probability that the treatment T_2 is preferred to T_1 is:

$$(3.9) \quad P_{12}(2) = \int_{\theta_1=0}^1 \int_{\theta_2=0}^{1-\theta_1} \int_{\gamma=0}^{\infty} \int_{v=0}^{\infty} \psi_{12}(2)p(\theta_1, \theta_2, \gamma, v|\mathbf{x})dv d\gamma d\theta_2 d\theta_1,$$

$$\theta_i \geq 0, i = 1, 2, 3, \sum_{i=1}^3 \theta_i = 1, \gamma, v > 0,$$

where $\psi_{12}(2)$ is the preference probability given in (2.2).

The predictive probability for no preference is:

$$(3.10) \quad P_{12}(0) = \int_{\theta_1=0}^1 \int_{\theta_2=0}^{1-\theta_1} \int_{\gamma=0}^{\infty} \int_{v=0}^{\infty} \psi_{12}(0)p(\theta_1, \theta_2, \gamma, v|\mathbf{x})dv d\gamma d\theta_2 d\theta_1,$$

$$\theta_i \geq 0, i = 1, 2, 3, \sum_{i=1}^3 \theta_i = 1, \gamma, v > 0$$

where $\psi_{12}(0)$ is the preference probability given in (2.3).

The predictive probabilities using the informative prior are attained with the help of programming in the SAS package.

TABLE 3. The Predictive Probabilities using the informative (D-G-G) Prior

Pairs (i, j)	(1,2)	(2,1)	(1,3)	(3,1)	(2,3)	(3,2)
$P_{ij}(1)$	0.5727	0.6621	0.6967	0.5670	0.7291	0.5120
$P_{ij}(2)$	0.3161	0.2349	0.2245	0.3448	0.1884	0.3886
$P_{ij}(0)$	0.1112	0.1030	0.0789	0.0881	0.0825	0.0994

The effect of order of presentation is quite overwhelming. The treatment which is presented first has more probability of being preferred while compared pairwise. The predictive probabilities of no preference among all the pairwise comparisons of the treatments are not more than 0.12.

4. Bayesian Analysis using the Conjugate Prior

Another informative prior is assumed to be the conjugate prior. The analysis of the van Baaren model VI for three treatments is presented below using the said prior.

4.1. The Choice of Conjugate Prior for van Baaren model VI. Clarke (1996) mentions that the use of the conjugate prior is possible for the likelihoods in an exponential form. When the likelihood function belongs to the exponential family then a conjugate prior can be taken according to Gelman et al. (2003). The likelihood function of the van Baaren model, given in (3.2), belongs to the exponential family as is obvious:

if we take $\prod_{i \neq j} \prod_j \frac{(\theta_i^{g_i} \gamma^{c_K} v^{T+c_o})}{(\gamma \theta_i + \theta_j + v \theta_i \theta_j)^{r_{ij}}}$ as $g(\theta_i, \nu, \gamma)$ and $\prod_{i \neq j} \prod_j \frac{r_{ij}!}{w_{ij}(1)! w_{ij}(2)! t_{ij}!}$ as $h(\mathbf{x})$ then, $\wp_1(\mathbf{x}) = g_i = \sum_j \{w_{ij}(1) + w_{ji}(2) + t_{ij} + t_{ji}\}$, $\wp_2(\mathbf{x}) = T = \sum_{i=1}^m \sum_{j \neq i}^m t_{ij}$, similarly $\wp_3(\mathbf{x}) = K = \sum_{i=1}^m \sum_{j \neq i}^m w_{ij}(1)$ and $\wp_4(\mathbf{x}) = r_{ij} = w_{ij}(1) + w_{ij}(2) + t_{ij}$.

Hence, according to Clarke (1996) and Gelman et al. (2003), conjugate prior of the model can be taken. The conjugate prior of the parameters $\theta_1, \theta_2, \dots, \theta_m, \gamma$ and v for the model can be assumed as:

$$(4.1) \quad p(\theta_1, \dots, \theta_m, \gamma, v) = \prod_{i(\neq j)=1}^m \prod_{j=1}^m \frac{\theta_i^{c_i} v^{c_o} \gamma^{c_K}}{(\gamma \theta_i + \theta_j + v \theta_i \theta_j)^{c'}};$$

$\sum_{i=1}^m \theta_i = 1$, for $i = 1, \dots, m, \gamma, v > 0$, where c_i, c_o, c_K and c' for $i = 1, \dots, m$ are the hyperparameters.

The joint posterior distribution of the parameters $\theta_1, \theta_2, \dots, \theta_m, \gamma$ and v obtained by combining the likelihood given in (3.2) with the conjugate prior of the model is:

$$(4.2) \quad p(\theta_1, \theta_2, \dots, \theta_m, \gamma, v) \propto \prod_{i(\neq j)=1}^m \prod_{j=1}^m \frac{\theta_i^{g_i+c_i} \gamma^{K+c_K} v^{T+c_o}}{(\gamma \theta_i + \theta_j + v \theta_i \theta_j)^{r_{ij}+c'}}$$

$$\theta_i \geq 0, i = 1, 2, \dots, m, \sum_{i=1}^m \theta_i = 1, \gamma, v > 0$$

The marginal posterior distribution for θ_1 is derived as:

$$(4.3) \quad p(\theta_1 | \mathbf{x}) = \frac{\theta_1^{g_1+c_1}}{q} \int_{\theta_2=0}^{1-\theta_1} \dots \int_{\theta_{m-1}=0}^{1-\theta_2 \dots - \theta_{m-2}} \int_{\gamma=0}^{\infty} \int_{v=0}^{\infty} \prod_{i \neq j=1}^{m-1} \prod_j^{m-1} \frac{\theta_i^{g_i+c_i} \gamma^{K+c_K} v^{T+c_o}}{(\gamma \theta_i + \theta_j + v \theta_i \theta_j)^{r_{ij}+c'}} dv d\gamma d\theta_{m-1} \dots d\theta_2;$$

$$0 \leq \theta_1 \leq 1, \gamma, v > 0,$$

where q is the normalizing constant.

4.2. Elicitation of the Hyperparameters. The hyperparameters are elicited via the elicitation of the confidence levels as given in Sub-section 3.1. The following function is minimized for the elicitation:

$$(4.4) \quad \zeta(c_1, \dots, c_m, c_o, c_K, c') = \min_{c_1, \dots, c_m, c_o, c_K, c'} \sum_{h=1}^k |(CCL)_h - (ECL)_h|.$$

Let the prior predictive distribution $\{p(w_{ij}(1), t_{ij}, w_{ji}(1), t_{ji})\}$ for the number of times the treatment T_i is preferred to and tied up with the treatment T_j when a pair of the treatments (T_i, T_j) is being compared with the number of times r_{ij} and r_{ji} be:

$$(4.5) \quad p(w_{ij}(1), t_{ij}, w_{ji}(1), t_{ji}) = q \int_{\theta_i=0}^1 \int_{v=0}^{\infty} \int_{\gamma=0}^{\infty} \frac{\theta_i^{g_i+c_i} (1-\theta_i)^{g_j+c_j} \gamma^{K+c_K} v^{T+c_o}}{\{\gamma \theta_i + (1-\theta_i) + v \theta_i (1-\theta_i)\}^{(r_{ij}+c')}} \frac{1}{\{\theta_i + \gamma(1-\theta_i) + v \theta_i (1-\theta_i)\}^{(r_{ji}+c')}} d\gamma dv d\theta_i$$

$$w_{ij}(1), t_{ij} = 0, 1, \dots, r_{ij}, w_{ji}(1), t_{ji} = 0, 1, \dots, r_{ji},$$

where $q = \frac{r_{ij}! r_{ji}!}{w_{ij}(1)! w_{ij}(2)! t_{ij}! w_{ji}(1)! w_{ji}(2)! t_{ji}!}$.

The hyperparameters c_1, c_2, c_3, c_o, c_k and c' are obtained assuming the same confidence levels and credible intervals as used in the elicitation of the hyperparameters of informative (D-G-G) prior in Section 3. The values of the hyperparameters are elicited

with the help of an elicitation program and are evaluated to be 0.25, 0.25, 0.01, 6.85, 2.85 and 4.26, respectively at the minimum value 0.0593 of the function (4.4).

For the Bayesian analysis of van Baaren model VI for $m = 3$ using the conjugate prior, the joint posterior and marginal posterior distributions of $\theta_1, \theta_2, \theta_3, v$ and γ are derived from (4.2). The data required to carry out the analysis is the same as used in the previous analysis given in Table 3.1.

To find the posterior means of the parameters $\theta_1, \theta_2, \theta_3, \gamma$ and v using the conjugate prior, the quadrature method is used. The means are obtained to be 0.3235, 0.4595, 0.2170, 2.2138 and 1.6928, respectively.

The posterior modes have been found out to be 0.3327, 0.4521, 0.2152, 2.2057 and 1.5970, respectively. Both the posterior estimates are observed to be comparatively closer to those obtained using the informative (D-G-G) prior. Both the posterior estimates (means and modes) show same ranking of the treatments as is revealed by the posterior estimates acquired using D-G-G prior.

The hypotheses compared for the case of conjugate prior are: $H_{ij} : \theta_i > \theta_j$ and $H_{ij}^c : \theta_j \geq \theta_i, i \neq j = 1, 2, 3$. The posterior probabilities of the hypotheses are given in Table 4.1. The hypotheses H_{12}^c, H_{13} and H_{23} are accepted. The test decisions are the same as are obtained using the informative (D-G-G) prior.

TABLE 4. Posterior probabilities using the conjugate prior

Hypotheses	p_{ij}	Hypotheses	q_{ij}
$H_{12} : \theta_1 > \theta_2$	0.0411	$H_{12}^c : \theta_2 \geq \theta_1$	0.9590
$H_{13} : \theta_1 > \theta_3$	0.9593	$H_{13}^c : \theta_3 > \theta_1$	0.0407
$H_{23} : \theta_2 > \theta_3$	0.9996	$H_{23}^c : \theta_3 > \theta_2$	0.0004

We find the predictive probabilities using the conjugate prior following the same procedure as described in the previous analysis. The probabilities are shown in Table 4.2:

TABLE 5. The Predictive Probabilities using the conjugate Prior

Pairs (i, j)	(1,2)	(2,1)	(1,3)	(3,1)	(2,3)	(3,2)
$P_{ij}(1)$	0.4996	0.6364	0.6779	0.5182	0.7224	0.4313
$P_{ij}(2)$	0.3267	0.2076	0.2106	0.3550	0.1587	0.4189
$P_{ij}(0)$	0.1737	0.1560	0.1115	0.1267	0.1189	0.1498

The values of the predictive probabilities obtained make it quite obvious that they differ from those obtained by the informative (D-G-G) prior. The difference is not overwhelming but it may be because of the values of elicited hyperparameters of the conjugate prior.

5. Appropriateness of the Model

The hypotheses to test the appropriateness of the model are:

H_o : The model is considered to be true for any value of $\theta = \theta_o$.

H^c : The model is considered not to be true for any value of θ .

To test the appropriateness of the model as is done by Aslam (1996) in case of three treatments, observed number of preferences is compared with the expected number of

TABLE 6. Observed and Expected Number of Preferences

Pairs(i, j)	$w_{ij}(1)$	$\hat{w}_{ij}(1)$	$w_{ij}(2)$	$\hat{w}_{ij}(2)$	t_{ij}	\hat{t}_{ij}
(1,2)	23	24.17	11	13.13	8	4.70
(2,1)	29	28.59	6	9.96	8	4.45
(1,3)	27	30.08	11	9.51	5	3.41
(3,1)	22	23.94	14	14.33	6	3.73
(2,3)	34	30.00	6	7.6	1	3.40
(3,2)	23	21.63	16	16.16	3	4.20

preferences. The χ^2 statistic is used to test the goodness of fit of the model for paired comparison. The χ^2 Statistic is:

$$(5.1) \quad \chi^2 = \sum_{i \neq j}^m \left\{ \frac{(w_{ij}(1) - \hat{w}_{ij}(1))^2}{\hat{w}_{ij}(1)} + \frac{(w_{ij}(2) - \hat{w}_{ij}(2))^2}{\hat{w}_{ij}(2)} + \frac{(t_{ij} - \hat{t}_{ij})^2}{\hat{t}_{ij}} + \frac{(w_{ji}(1) - \hat{w}_{ji}(1))^2}{\hat{w}_{ji}(1)} + \frac{(w_{ji}(2) - \hat{w}_{ji}(2))^2}{\hat{w}_{ji}(2)} + \frac{(t_{ji} - \hat{t}_{ji})^2}{\hat{t}_{ji}} \right\},$$

with $2m(m - 1) - (m + 1)$ degree of freedom taken from Davidson and Beaver (1977). Where $\hat{w}_{ij}(1)$ and $\hat{w}_{ij}(2)$ are the expected number of times T_i and T_j are preferred, respectively and \hat{t}_{ij} is the expected number of times T_i and T_j end up in a tie when T_i is presented first. Similarly, $\hat{w}_{ji}(1), \hat{w}_{ji}(2)$ and \hat{t}_{ji} are described when T_j is presented first.

The value of the χ^2 statistic is obtained as 12.96 and the p-value is 0.39 which interprets that the model is suitable for the data.

6. Conclusion

Bayesian analysis of a paired comparison model, the van Baaren model VI, has proved the model to be appropriate to fit for a paired comparison data when order effect factor is being considered. Bayesian analysis has been carried out using the informative and conjugate priors. The posterior estimates (means and modes) attained using both the priors have depicted similar values. While comparing the treatment parameters via Bayesian testing of hypotheses, it is evident that the treatment T_2 is the most favoured one and T_3 is the least preferred one. The preference of any of the treatment over any other one in a future single comparison cannot be predicted because of the overwhelming order effect. The model seems to be useful for the situations when the order of presentation influences the preferences of treatments. Hence, the said model is interpreted as suitable and practical especially for the cases of sensory paired comparison testing. The study depicts that as both of the priors have shown nearly same results so any one of them may be used for Bayesian analysis.

References

- [1] Aslam, M. *Bayesian analysis for paired comparisons data*. Ph.D. Dissertation, Dept. of Mathematics. University of Wales. Aberystwyth, 1996.
- [2] Aslam, M. Bayesian analysis for paired comparisons models allowing ties and not allowing ties. *Pakistan Journal of Statistics* **18**, 53-69, 2002.
- [3] Aslam, M. An application of prior predictive distribution to elicit the prior density. *Journal of Statistical Theory and Applications* **2**, 70-83, 2003.
- [4] Aslam, M. Bayesian comparison of the paired comparison models allowing ties. *Journal of Statistical Theory and Applications* **4**, 161-171, 2005.

- [5] Berger, J. O. *Statistical decision theory and Bayesian analysis*, 2nd edn. (New York: Springer-Verlag, 1985).
- [6] Bradley, R. A., and Terry, M.E. Rank analysis of incomplete block design: I. The method of paired comparisons. *Biometrika* **39**, 324-345, 1952.
- [7] Clarke, B. Implications of reference priors for prior information and for sample size. *Journal of the American Statistical Association* **91**, 173-184, 1996.
- [8] Davidson, R.R. and Beaver, R.J. On extending the Bradley-Terry model to incorporate within pair order effects. *Biometrics* **33**, 693-702, 1977.
- [9] Davidson, R.R. and Solomon, D.L. A Bayesian approach to paired comparison experimentation. *Biometrika* **60**, 477-87, 1973.
- [10] Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. *Bayesian data analysis*, 2ndedn. (London: Chapman and Hall, 2003).
- [11] Glickman, M. E. Bayesian Locally Optimal Design of Knockout Tournaments. *Journal of Statistical Planning and Inference* **138**, 2117 – 2127, 2008.
- [12] Kadane, J. B. and Wolfson, L. J. Experiences in elicitation. *The Statistician* **47**, 3-19, 1998.
- [13] Kim, D. H. and Kim, H. J. A Bayesian approach to paired comparison of several products of Poisson rates. Proceedings of the Autumn Conference, Korean Statistical Society, 2004.
- [14] Kim, H. J. A Bayesian approach to paired comparison. Rankings based on a graphical model. *Computational Statistics & Data Analysis* **48**, 269-290, 2005.
- [15] Li, L. and Kim, K. Estimating driver's crash risks based on extended Bradley-Terry model: An induced exposure method. *Journal of Royal Statistical Society A* **163**, 227-240, 2000.
- [16] Merrick, J. R. W., van Dorp, J. R. and Singh, A. Analysis of correlated expert judgments from extended pairwise comparisons. *Decision Analysis* **2**, 17-29, 2005.
- [17] Neuman, K. and Watson, B. G. Application of paired comparison methodology in measuring Canadian's forest values. *The Public Opinion Quarterly* **57**, 437-464, 1993.
- [18] Szved, P., van Dorp, J. R., Merrick, J.R.W., Mazzuchi, T.A. and Singh, A. A Bayesian paired comparison approach for relative accident probability assessment with covariate information. *European Journal of Operations Research* **169**, 157-177, 2006.
- [19] van Baaren. On a class of extensions to the Bradley-Terry model in paired comparisons. *Statistica Neerlandica* **32**, 57-66, 1978.