



A note on the statistical analysis of point judgment matrices

MG Kabera*

LM Haines†

Received: 24 July 2012; Revised: 13 February 2013; Accepted: 21 February 2013

Dedication to Emeritus Professor Theodor Stewart

Professor Theo Stewart has an encyclopaedic knowledge of a great many things, ranging from operational research, decision theory and statistics through to LaTeX and single malt whiskey. He shares this knowledge generously and unconditionally with colleagues and students alike. Theo is an excellent teacher, an outstanding researcher and a boundless source of institutional wisdom, in short the epitome of a fine and dedicated academic. It has been a great pleasure for the second author to have known Theo for many years and to have worked with him in the same department over the last eight years. We salute him on this, his 70th birthday year.

Abstract

The Analytic Hierarchy Process is a multicriteria decision making technique developed by Saaty in the 1970s. The core of the approach is the pairwise comparison of objects according to a single criterion using a 9-point ratio scale and the estimation of weights associated with these objects based on the resultant judgment matrix. In the present paper some statistical approaches to extracting the weights of objects from a judgment matrix are reviewed and new ideas which are rooted in the traditional method of paired comparisons are introduced.

Key words: Analytic Hierarchy Process, least squares, logistic distribution, Bradley-Terry model, Kullback-Leibler distance.

1 Introduction

The Analytic Hierarchy Process is a multicriteria decision making technique developed by Saaty in the 1970s which has received considerable attention in the mathematical and statistical literature [11, 18]. The core of the approach is the pairwise comparison of objects according to a single criterion using a 9-point ratio scale and the estimation of weights associated with these objects based on the resultant judgment matrix. Saaty suggested that these estimates be taken to be proportional to the eigenvector corresponding to the largest right eigenvalue of the judgment matrix [17]. However, his idea has been

*Department of Statistics, University of Johannesburg, PO Box 524, Auckland Park 2006, South Africa.

†Corresponding author: LM Haines, Department of Statistical Sciences, University of Cape Town, email: linda.haines@uct.ac.za

the subject of much criticism, in particular because it has a deterministic rather than a statistical basis [10].

The aim of this paper is to review the key statistical approaches to extracting the weights of objects from a judgment matrix and, in so doing, to relate the embedded models to the traditional linear models used in the method of paired comparisons [5]. The remainder of the paper is organized as follows. §2 comprises a brief account of the method of paired comparisons and a formal specification of a judgment matrix. Statistical approaches to the analysis of judgment matrices are then introduced within the context of paired comparisons, with those which are distribution-based discussed in §3 and those based more directly on the method of paired comparisons in §4. An illustrative example is presented in §5 and conclusions and pointers for further research are given in §6.

2 Preliminaries

A brief introduction to paired comparisons and judgment matrices within the AHP context is presented in the following subsections.

2.1 The linear model for paired comparisons

In the traditional setting for paired comparisons, d decision makers are invited to compare n objects pairwise with respect to a single criterion and to state, quite simply, their preference for each pair with no ties permitted. The results can be assembled in a matrix $D = \{d_{ij}\}$ where d_{ij} is the number of decision makers who prefer object A_i to object A_j for $i \neq j, i, j = 1, \dots, n$ and thus $d_{ij} + d_{ji} = d$. A full account of the analysis of such data is given in the seminal book by David [5] and a brief summary is presented here.

The perceived underlying merit of object A_i is taken to be a continuous random variable Y_i with mean μ_i and the probability that A_i is preferred to A_j can be introduced as

$$\pi_{ij} = P(A_i \text{ preferred to } A_j) = P(Y_i > Y_j), \quad i, j = 1, 2, \dots, n, \quad i \neq j.$$

By defining $Z_i - Z_j = (Y_i - \mu_i) - (Y_j - \mu_j)$ and invoking subtle distributional arguments, it can be shown that

$$\pi_{ij} = P(Y_i > Y_j) = P(Z_i - Z_j > -(\mu_i - \mu_j)) = P(Z_i - Z_j < \mu_i - \mu_j) = H(\mu_i - \mu_j), \quad (1)$$

where $H(\cdot)$ is the distribution function of $Z_i - Z_j$. Note that the merits of the objects A_i , with $i = 1, 2, \dots, n$, are chosen on a linear scale and, since the origin of the linear scale is arbitrary, it is usual to impose an additional constraint on the μ_i such as $\sum_{i=1}^n \mu_i = 0$ or $\mu_n = 0$. It also follows immediately from result (1) that $\pi_{ij} = \frac{1}{2}$ if and only if $\mu_i = \mu_j$, that $\pi_{ij} > \frac{1}{2}$ if and only if $\mu_i > \mu_j$ and that $\pi_{ij} < \frac{1}{2}$ if and only if $\mu_i < \mu_j$ for $i, j = 1, 2, \dots, n$ and $i \neq j$.

The preference probabilities π_{ij} , for $i, j = 1, 2, \dots, n$, and $i \neq j$ given in equation (1) are specified by the choice of distribution function H and two such choices are of particular interest. Specifically, in the Thurstone-Mosteller model [16, 22], the variables Y_i are taken

to be normally distributed and it follows straightforwardly that

$$\pi_{ij} = P(Y_i > Y_j) = \Phi\left(\frac{\mu_i - \mu_j}{\sigma}\right), \quad (2)$$

where $\Phi(\cdot)$ is the cumulative density function of the standard normal distribution and σ^2 is the variance of the difference $Y_i - Y_j$. In the more widely used Bradley-Terry model [3], the variables Y_i are taken to follow independent extreme value distributions and the differences $Y_i - Y_j$ therefore follow logistic distributions [6]. It can then be shown that

$$\pi_{ij} = P(Y_i > Y_j) = \frac{\pi_i}{\pi_i + \pi_j}, \quad (3)$$

where the parameters π_i have an immediate and natural interpretation as weights or probabilities associated with the objects A_i , $i = 1, 2, \dots, n$, respectively.

2.2 Judgment matrices

Suppose now that n objects are again to be compared according to a particular criterion but that a single decision maker expresses his or her relative preferences on a ratio scale. Saaty [18] suggested choosing relative preferences on an ordinal scale of integers from 1 to 9, together with their reciprocals. However, other scales can be used. For example Becker and co-authors [2] use a 5-point Likert scale. More formally, let a_{ij} denote the relative preference of object A_i when compared with object A_j for $i, j = 1, 2, \dots, n$. The relative preferences can then be assembled in a positive reciprocal matrix, termed a point judgment matrix, of the form $A = \{a_{ij}\}$ where $a_{ij} > 0$, $a_{ji} = 1/a_{ij}$ and $a_{ii} = 1$ for $i, j = 1, 2, \dots, n$. The entries of A can be regarded as the ratios of weights associated with the objects in the pairwise comparisons. Once the pairwise judgments have been elicited, the crucial question is how to determine the weights associated with the objects.

3 Distributional approaches

In this section three distributional approaches are now discussed.

3.1 Logarithmic least squares and the normal distribution

Crawford and Williams [4] introduced what is arguably the first statistical approach to the analysis of point judgment matrices, termed the logarithmic least squares method (LLSM), in 1985. More detailed work on the LLSM was developed later by Kabera [13], Kabera and Haines [14] and Laininen and Hämäläinen [15]. However, the LLSM can be derived using the same arguments as those invoked in deriving the Thurstone-Mosteller model for paired comparisons and this insight is now discussed.

The relative preferences a_{ij} can be expressed as

$$a_{ij} = \frac{w_i}{w_j} e_{ij}^*, \quad 1 \leq i < j \leq n, \quad (4)$$

where w_i and w_j are the unknown weights associated with the objects A_i and A_j respectively. Moreover, $w_i > 0$ for $i = 1, 2, \dots, n$ and $\sum_{i=1}^n w_i = 1$ and e_{ij}^* is a positive error term which captures the inconsistency in the judgments. By invoking a logarithmic transformation, model (4) can be re-expressed as the linear model with no intercept

$$y_{ij} = \beta_i - \beta_j + e_{ij}, \quad 1 \leq i < j \leq n,$$

where $y_{ij} = \ln a_{ij}$, $\beta_i = \ln w_i$, $\beta_j = \ln w_j$ and $e_{ij} = \ln e_{ij}^*$. Assuming that the differences y_{ij} are observed values of random variables $Y_{ij} = Y_i - Y_j$ taken to be normally distributed as in the Thurstone-Mosteller model, that is $Y_{ij} \sim N(\beta_i - \beta_j, \sigma^2)$, simple arguments similar to those developed in Kabera and Haines [14] yield the following results. Pairwise differences of the parameters, $\beta_i - \beta_j$, are estimable and it is straightforward to show that the least squares or, equivalently, the maximum likelihood estimates (MLEs) are given by

$$\tilde{\beta}_i - \tilde{\beta}_j = \frac{1}{n} \left(2y_{ij} + \sum_{\substack{k=1 \\ k \neq i, j}}^n y_{ik} - \sum_{\substack{k=1 \\ k \neq i, j}}^n y_{jk} \right) \quad 1 \leq i < j \leq n.$$

Estimates of the weights associated with the objects can be formulated as

$$\hat{w}_i = \frac{\exp \tilde{\beta}_i}{\sum_{j=1}^n \exp \tilde{\beta}_j} = \frac{1}{\sum_{j=1}^n \exp \{-(\tilde{\beta}_i - \tilde{\beta}_j)\}}, \quad i = 1, 2, \dots, n.$$

Since these weights depend only on the unique estimates of $\beta_i - \beta_j$, they are unique. Approximate variances and covariances of the estimates \hat{w}_i , for $i = 1, \dots, n$, can be obtained by observing that $\text{Var}(\tilde{\beta}_i - \tilde{\beta}_j) = 2\sigma^2/n$ and by invoking the delta method to give

$$\text{Var}(\hat{w}_i) \approx \frac{\sigma^2}{n} w_i^2 (1 - 2w_i + \sum_{k=1}^n w_k^2), \quad i = 1, 2, \dots, n \quad (5)$$

and

$$\text{Cov}(\hat{w}_i, \hat{w}_j) \approx \frac{\sigma^2}{n} w_i w_j (-w_i - w_j + \sum_{k=1}^n w_k^2), \quad 1 \leq i < j \leq n \quad (6)$$

respectively. Estimates of these variances and covariances can immediately be obtained by the ‘‘plug-in’’ principle, that is by replacing the weights with their estimates and σ^2 with the residual variance s^2 .

3.2 The logistic distribution

The approach based on the logistic distribution for analyzing judgment matrices in the AHP was introduced by Haines and Litvine [9] in the context of interval judgments, and was fully developed for point judgment matrices by Kabera [13] and Kabera and Haines [14]. The derivation of weights associated with objects compared pairwise and the corresponding covariance structure can be deduced directly from the Bradley-Terry model reviewed in §2 as follows.

Suppose that, in a pairwise comparison of n objects, the merit of object A_i is assumed to be represented by an outcome of a random variable Y_i which follows an extreme value distribution with probability density function (pdf) $f(y_i) = e^{-(y_i-\theta_i)} \exp[-e^{-(y_i-\theta_i)}]$, with y_i taking a value on the real line and the unknown parameter $\theta_i > 0$ for $i = 1, 2, \dots, n$. Then, if Y_i and Y_j are independent, the random variable $Y_{ij} = Y_i - Y_j$, for $i, j = 1, 2, \dots, n$ and $i \neq j$, follows a logistic distribution with location parameter $\theta_i - \theta_j$ and scale parameter 1 [6]. For the Bradley-Terry model, the probability that object A_i is preferred to object A_j for $1 \leq i < j \leq n$ is simply

$$\pi_{ij} = P(Y_i > Y_j) = \frac{\pi_i}{\pi_i + \pi_j} = \frac{e^{\theta_i}}{e^{\theta_i} + e^{\theta_j}},$$

that is

$$\text{logit}(\pi_{ij}) = \ln \frac{\pi_{ij}}{1 - \pi_{ij}} = \theta_i - \theta_j$$

and the logistic distribution itself is subsumed in this development [3, 5]. In contrast, suppose that the entry a_{ij} in the judgment matrix A is a realization of a random variable A_{ij} which represents the strength of object A_i relative to object A_j and is related to the variable Y_{ij} as $Y_{ij} = \ln(A_{ij})$. In other words, A_{ij} is assumed to be log-logistic. Then the likelihood for the parameters $\boldsymbol{\theta} = (\theta_1, \dots, \theta_n)$ associated with the judgment matrix A is given by

$$L(\boldsymbol{\theta}) = \prod_{\substack{i < j \\ i, j = 1, n}}^n \frac{e^{-[y_{ij} - (\theta_i - \theta_j)]}}{[1 + e^{-[y_{ij} - (\theta_i - \theta_j)]}]^2}$$

where $y_{ij} = \ln(a_{ij})$ and the MLEs of $\boldsymbol{\theta}$ can be readily obtained subject to a single constraint, for example $\theta_n = 0$, in order to ensure identifiability. Estimates of the weights follow as

$$\hat{w}_i = \frac{e^{\hat{\theta}_i}}{\sum_{j=1}^n e^{\hat{\theta}_j}}, \quad i = 1, 2, \dots, n,$$

where $\hat{\theta}_i$ is the MLE of θ_i , for $i = 1, \dots, n$. The asymptotic variances and covariances of the estimates of weights \hat{w}_i can be obtained using the delta method and are the same as those for the LLSM approach but with σ^2 replaced by 3 in expressions (5) and (6) respectively [14].

3.3 Discretization

An approach which recognizes that judgments represented on a ratio scale are in fact discrete was introduced by Kabera [13] and Kabera and Haines [14]. The method is termed “discretization” and is a natural extension to the approaches based on the normal and logistic distributions described above. Specifically, the entries in a judgment matrix from the AHP are taken on an integer and reciprocal integer scale. Thus, strictly, these entries are realizations of a discrete random variable A_{ij}^* and $Y_{ij}^* = \ln A_{ij}^*$ is necessarily discretely distributed for $1 \leq i < j \leq n$. In discussions so far, the random variable

$Y_{ij} = \ln(A_{ij})$ has been assumed to be continuous. It is thus more correct to regard the event $Y_{ij}^* = \ln a_{ij}$, as equivalent to the event

$$Y_{ij} \in \left(\frac{\ln(a_{ij} - 1) + \ln a_{ij}}{2}, \frac{\ln a_{ij} + \ln(a_{ij} + 1)}{2} \right).$$

Let $\alpha = \ln \sqrt{(a_{ij} - 1)a_{ij}}$ and $\beta = \ln \sqrt{a_{ij}(a_{ij} + 1)}$. Then it follows that

$$P(A_{ij}^* = a_{ij}) = P(Y_{ij}^* = \ln a_{ij}) = \int_{\alpha}^{\beta} h(y_{ij}) dy_{ij}, \quad (7)$$

where $h(y_{ij})$ is the pdf of y_{ij} for $1 \leq i < j \leq n$. Estimates for the unknown parameters can then be obtained by maximizing the likelihood $\prod_{i < j} P(Y_{ij}^* = \ln a_{ij})$ or, equivalently, the log-likelihood subject to appropriate constraints and the variance-covariance matrix of the resultant estimates can be approximated by the inverse of the observed Fisher information matrix. Note that integrals of the form (7) can be evaluated explicitly in the case of the logistic distribution since

$$\begin{aligned} P(m_{ij} < Y_{ij} < M_{ij}) &= \int_{m_{ij}}^{M_{ij}} \frac{e^{-[y_{ij} - (\theta_i - \theta_j)]}}{(1 + e^{-[y_{ij} - (\theta_i - \theta_j)]})^2} dy_{ij} \\ &= \frac{e^{-[M_{ij} - (\theta_i - \theta_j)]} - e^{-[m_{ij} - (\theta_i - \theta_j)]}}{(1 + e^{-[m_{ij} - (\theta_i - \theta_j)]})(1 + e^{-[M_{ij} - (\theta_i - \theta_j)]})} \end{aligned}$$

but must be calculated numerically for the normal distribution.

4 Direct approaches

The question as to whether the traditional paired comparison models can be used more directly in modelling judgment matrices of the form $A = \{a_{ij}\}$ than the models discussed in §3 now arises and is addressed in the following subsections.

4.1 Genest-M'Lan approach

Genest and M'Lan [7] suggested that a_{ij} , the relative preference for object A_i relative to object A_j recorded on a ratio scale such as the Saaty scale, can be interpreted as reflecting the fact that the objects have been compared n_{ij} times where

$$n_{ij} = \max\{a_{ij}, a_{ji}\} + 1 \text{ for } 1 \leq i < j \leq n \quad (8)$$

and thus that A_i is preferred to A_j x_{ij} times where x_{ij} is determined by $a_{ij} = x_{ij}/(n_{ij} - x_{ij})$. Thus $x_{ij} = n_{ij}a_{ij}/(1 + a_{ij})$ or, more succinctly, $x_{ij} = \max\{a_{ij}, 1\}$, for $1 \leq i < j \leq n$. With this interpretation, the pairwise comparison of objects A_i and A_j can be regarded as following a binomial distribution with the number of trials n_{ij} taken to be independent, the number of successes given by x_{ij} and the probability of success π_{ij} taken to follow the Bradley-Terry model for paired comparisons.

This approach is indeed an interesting one but the following concerns should be noted. First the interpretation of the term n_{ij} defined in expression (8) as the number of comparisons of objects A_i and A_j for $1 \leq i < j \leq n$, implies that pairs of objects are not examined the same number of times. Second, and as an example, suppose that $a_{12} = 4$. This can be interpreted as indicating that object A_1 is preferred to object A_2 in 4 out of 5, or 40 out of 50, or 400 out of 500 comparisons or any number of preferences and trials preserving the proportion 4 : 5 in favour of object A_1 . These concerns indicate an arbitrariness in the selection of the number of pairwise comparisons of the objects. Genest and M'LAN [7] recognized these problems and suggested that the number of comparisons between all pairs of objects be taken to be equal and, specifically, to be the least common multiplier (LCM) of $\max\{a_{ij}, a_{ji}\} + 1$. Thus for the Saaty scale the LCM of 2, 3, ..., 10, is 2520. However, the number of comparisons still remains essentially arbitrary. An example illustrating the above concerns can be found in Kabera [13]. In conclusion therefore the interpretation of the preferences in a point judgment matrix as emanating from a binomial model requires great caution.

4.2 A Distance approach

An alternative approach to that of Genest and M'LAN [7] based on the method of paired comparisons was developed in the thesis of Kabera [13] and is now introduced here. Specifically, the entry a_{ij} in a judgment matrix A is interpreted as an odds ratio for the preference probability and thus as

$$a_{ij} = \frac{p_{ij}}{1 - p_{ij}}$$

so that

$$p_{ij} = \frac{a_{ij}}{a_{ij} + 1} \quad 1 \leq i < j \leq n, \quad (9)$$

where p_{ij} is the observed probability that object A_i is preferred to object A_j for $1 \leq i < j \leq n$. Note that the relative preference a_{ij} is taken to be on a ratio scale, for example (but not necessarily) the Saaty scale. Estimates of the parameters of models describing the preference probabilities π_{ij} can be obtained as those values for which the observed probabilities p_{ij} are as close as possible, in some sense, to the true values π_{ij} , $1 \leq i < j \leq n$. Two measures of "closeness," one based on least squares and the other on the Kullback-Leibler distance, are now considered and the ideas reinforced by invoking the Bradley-Terry model (3).

4.2.1 Least squares

Consider the least squares approach of David [5] which involves minimizing the sum of squares

$$S = \sum_{\substack{j \neq i \\ j=1}}^n [H^{-1}(p_{ij}) - H^{-1}(\pi_{ij})]^2 = \sum_{\substack{j \neq i \\ j=1}}^n [H^{-1}(p_{ij}) - (\mu_i - \mu_j)]^2$$

where the observed probabilities p_{ij} are given by equation (9) and $H(\cdot)$ is given by (2). For the Bradley-Terry model (3), $\mu_i = \ln(\pi_i)$ and

$$H^{-1}(p_{ij}) = \ln \frac{p_{ij}}{1 - p_{ij}} = \ln a_{ij} \quad 1 \leq i < j \leq n,$$

so that estimates of the parameters are obtained by minimizing the expression

$$\sum_{\substack{j \neq i \\ j=1}}^n \left(\ln a_{ij} - \ln \frac{\pi_i}{\pi_j} \right)^2.$$

Note immediately that this approach gives the same results as the logarithmic least squares method discussed in §3.1 (see also [13] and [15]), although the underlying “philosophy” is very different. There is therefore no need to pursue this method further.

4.2.2 The Kullback-Leibler distance

The Kullback-Leibler distance between the probabilities p_{ij} and π_{ij} for $1 \leq i < j \leq n$, with respect to the distribution specified by the p_{ij} is given by

$$D = \sum_{\substack{i < j \\ i, j=1}}^n \left[p_{ij} \ln \frac{p_{ij}}{\pi_{ij}} + (1 - p_{ij}) \ln \frac{1 - p_{ij}}{1 - \pi_{ij}} \right]. \quad (10)$$

Consider now minimizing D and hence, since the probabilities p_{ij} are observed, maximizing the expression

$$D^* = \sum_{\substack{i < j \\ i, j=1}}^n [p_{ij} \ln \pi_{ij} + (1 - p_{ij}) \ln(1 - \pi_{ij})] = \sum_{\substack{i < j \\ i, j=1}}^n \left[p_{ij} \ln \frac{\pi_{ij}}{1 - \pi_{ij}} + \ln(1 - \pi_{ij}) \right] \quad (11)$$

with respect to the parameters on which the true probabilities π_{ij} depend. Maximizing expression (11) is equivalent to maximizing the log-likelihood for a binomial distribution with numbers of trials all equal and with probabilities of success π_{ij} . In particular note that the model proposed by Genest and M’Lan [7] gives the same parameter estimates as those obtained by maximizing (11) provided that all pairs of objects are assumed to be compared the same number of times. The present approach therefore gives some support to their methodology.

Consider now the case where the probability π_{ij} that object A_i is preferred to object A_j follows the Bradley-Terry model (3), that is $\pi_{ij} = \pi_i / (\pi_i + \pi_j)$ for $i, j = 1, 2, \dots, n$ and $i \neq j$. Then equation (11) can be written as

$$D^* = \sum_{\substack{i < j \\ i, j=1}}^n \left[p_{ij} \ln \frac{\pi_i}{\pi_j} + \ln \frac{\pi_j}{\pi_i + \pi_j} \right]. \quad (12)$$

Estimates of the weights which minimize the Kullback-Leibler distance (10) can be readily obtained by developing an iterative procedure which preserves the constraints that $\pi_i >$

0 and $\sum_{i=1}^n \pi_i = 1$ and which is similar to that developed for the method of paired comparisons [5]. Specifically, using (12) and solving for $\partial D^*/\partial \pi_i = 0$ gives

$$\frac{1}{\widehat{\pi}_i} \sum_{\substack{i < j \\ i, j = 1}}^n p_{ij} = \sum_{\substack{i < j \\ i, j = 1}}^n (\widehat{\pi}_i + \widehat{\pi}_j)^{-1} \quad i = 1, 2, \dots, n,$$

where $\widehat{\pi}_i$ is an estimate of π_i . This system of equations then forms the basis for an iterative scheme for finding the estimates $\widehat{\pi}_i, i = 1, 2, \dots, n$. The calculations require the following steps.

- (1) Choose starting values for the $\widehat{\pi}_i$, such as $\widehat{\pi}_1^{(0)} = \widehat{\pi}_2^{(0)} = \dots = \widehat{\pi}_n^{(0)} = \frac{1}{n}$.
- (2) Take the k^{th} iterative solution as

$$\widehat{\pi}_i^{(k)} = \frac{\sum_{\substack{i < j \\ i, j = 1}}^n p_{ij}}{\sum_{\substack{i < j \\ i, j = 1}}^n (\widehat{\pi}_i^{(k-1)} + \widehat{\pi}_j^{(k)})^{-1} + \sum_{\substack{i < j \\ i, j = 1}}^n (\widehat{\pi}_i^{(k-1)} + \widehat{\pi}_j^{(k-1)})^{-1}}, \quad i = 1, 2, \dots, n$$

- (3) Continue the process until $\widehat{\pi}_i^{(k+1)}$ is deemed to be sufficiently close to $\widehat{\pi}_i^{(k)}$, say at iteration m , and then the final estimates are taken to be

$$\widehat{\pi}_i = \frac{\widehat{\pi}_i^{(m)}}{\sum_{j=1}^n \widehat{\pi}_j^{(m)}}, \quad i = 1, 2, \dots, n.$$

Alternatively the weights can be estimated by invoking a constrained optimization routine.

The Kullback-Leibler distance is not a likelihood function. Thus standard errors for the parameter estimates cannot be obtained from the asymptotic results of likelihood theory. However, the jackknife technique can be used to estimate the weights and to provide attendant standard errors. Specifically consider estimating the weights by omitting the comparison of objects A_i and A_j as the vector $\widehat{\pi}_{(i,j)}$ where the subscripts indicate the removed pairs for $i, j = 1, 2, \dots, n$. There are $\binom{n}{2}$ such estimates and these form the basis for finding an overall estimate $\widehat{\pi}$ and the associated standard errors.

5 An example

The aim of this example is to fix ideas by comparing judgment weights and associated standard errors obtained using the Kullback-Leibler approach with those obtained using methods discussed earlier in this paper and in [14]. Thus, Saaty [18, p. 38] considered the following example in which four identical chairs, labelled 1, 2, 3 and 4, were placed in a line at distances of 9, 15, 21 and 28 yards from a light source respectively. Chairs were compared according to brightness and the following judgment matrix assembled:

$$A = \begin{bmatrix} 1 & 4 & 6 & 7 \\ 1/4 & 1 & 3 & 4 \\ 1/6 & 1/3 & 1 & 2 \\ 1/7 & 1/4 & 1/2 & 1 \end{bmatrix}.$$

The weights and their associated standard errors using the LLSM and the approaches based on the logistic distribution, discretization using the logistic distribution and the Kullback-Leibler distance are summarized in Table 1. For the LLSM, the coefficient of determination obtained from the analysis of variance (ANOVA) table was $R^2 = 96.8\%$, indicating an excellent fit of the model to the data and more particularly indicating that the matrix A is close to being consistent [8]. In addition, in terms of brightness, chair 1 was found to be significantly different to the others, chair 2 was significantly different from chair 4 but not chair 3, and chairs 3 and 4 were not significantly different from each other at the 5% level of significance.

Chairs	LLSM		Logistic		Discretization		Kullback-Leibler	
	Weights	Se	Weights	Se	Weights	Se	Weights	Se
1	0.614	0.052	0.614	0.245	0.613	0.203	0.632	0.043
2	0.225	0.041	0.225	0.194	0.225	0.160	0.218	0.034
3	0.099	0.020	0.099	0.096	0.098	0.078	0.094	0.012
4	0.062	0.013	0.062	0.062	0.064	0.052	0.056	0.007

Table 1: Estimates of the weights and associated standard errors for chair brightness.

It is clear that the distribution-based methods give practically the same weights, while the Kullback-Leibler distance approach gives slightly different weights, but that the ranking of chairs with respect to brightness is the same for all the four methods. It is also clear that the standard errors associated with the estimates of the weights for the LLSM and the Kullback-Leibler approaches are smaller than those obtained from the logistic and discretization approaches. This result is not surprising since it was demonstrated in §3 that the standard errors associated with the estimates of the weights from the logistic approach are $\sqrt{3/\sigma^2}$ larger than those obtained from the LLSM. In the present case this multiplier is approximately $\sqrt{3/s^2} = 4.7$, where s^2 is the estimate of error obtained from the ANOVA table for the LLSM, in accord with the results in Table 1. Also, it is clear that if the judgment matrix is consistent, then the jackknife estimates will have zero error. Since the matrix A is close to being consistent, this observation is reflected in the small standard errors for the Kullback-Leibler approach given in Table 1.

6 Conclusions

In this paper the connection between the traditional method of paired comparisons and a range of statistically based approaches to extracting weights for judgment matrices embedded in the Analytical Hierarchy Process is introduced and explored. The results from the example presented here, and indeed from other examples [13], suggest that the various statistical methods for deriving such weights produce estimates which are similar but attendant standard errors which are somewhat different. The question then arises as to which method is to be preferred. The approach based on the logistic distribution does not incorporate a scale parameter, while the standard errors for the estimates of the weights obtained from the Kullback-Leibler method tend to be understandably small. In contrast, the logarithmic least squares method includes a scale parameter in its model formulation and provides sound statistically based standard errors for the estimates of the weights. On

balance therefore, the logarithmic least squares method, or the discretized version thereof, is recommended for deriving weights from judgment matrices and, furthermore, is readily implemented in practice.

There is scope for further research into statistical approaches for analyzing judgment matrices. In particular statistically based methods address rank reversal since standard errors are associated with estimates of the weights and thus the rankings are not stated with certainty. However, the weights are constrained to lie in a simplex and it would be useful to develop confidence regions for the weights rather than simply reporting the standard errors associated with their estimates. More generally, Stern [21] introduced a generalization of the traditional paired comparison model by introducing the notion that the random variables which describe the perceived values of objects follow a Gamma distribution with the same shape parameter but with different scale parameters. This approach is particularly interesting in that it relates to a Poisson process which can be construed as describing how a decision maker quantifies his or her preferences and could well be pursued within the context of extracting weights for judgment matrices.

Finally it should be emphasized that approaches to the analysis of judgment matrices within the context of the AHP, not all statistically based, are the subject of on-going research. Some key developments and new methodologies are presented in the papers by Bajwa [1], Jones and Mardle [12], Srdjevic [19], Srdjevic and Srdjevic [20] and Wang, Parkan and Luo [23, 24].

Acknowledgments

The authors would like to thank the University of KwaZulu-Natal, the Medical Research Council, the University of Cape Town, the University of Johannesburg and the National Research Foundation of South Africa for financial support. The authors are also grateful to an anonymous referee for many extremely valuable suggestions.

References

- [1] BAJWA G, CHOO EU & WEDLEY WC, 2008, *Effectiveness analysis of deriving priority vectors from reciprocal pairwise comparison matrices*, Asia-Pacific Journal of Operational Research, **25(3)**, pp. 279–299.
- [2] BECKER PJ, WOLVAARDT JS, AND HENNESSY A & MAREE C, 2009, *A composite score for a measuring instrument utilising re-scaled Likert values and item weights from matrices of pairwise ratios*, Health SA Gesondheid: Journal of Interdisciplinary Health Sciences, **14(1)**, pp. 29–32.
- [3] BRADLEY RA & TERRY ME, 1952, *Rank analysis of incomplete block designs. I. The method of paired comparisons*, Biometrika, **39(3/4)**, pp. 324–345.
- [4] CRAWFORD G & WILLIAMS C, 1985, *A note on the analysis of subjective judgment matrices*, Journal of Mathematical Psychology, **29(4)**, pp. 387–405.
- [5] DAVID HA, 1988, *The Method of Paired Comparisons*, Second Edition, Charles Griffin, London.
- [6] DAVIDSON RR, 1969, *On a relationship between two representations of a model for paired comparisons*, Biometrics, **25(3)**, pp. 597–599.
- [7] GENEST C & M'LAN CE, 1999, *Deriving priorities from the Bradley-Terry Model*, Mathematical and Computer Modelling, **29(4)**, pp. 87–102.

- [8] GENEST C & RIVEST LP, 1994, *A statistical look at Saaty's Method of estimating pairwise preferences expressed on a ratio scale*, Journal of Mathematical Psychology, **38(4)**, pp. 477–496.
- [9] HAINES LM & LITVINE IM, 1999, *A statistical analysis of judgment matrices in the Analytic Hierarchy Process*, Proceedings of the 14th International Workshop in Statistical Modelling, Graz.
- [10] HOLDER RD, 1990, *Some comments on the Analytic Hierarchy Process*, The Journal of the Operational Research Society, **41(11)**, pp. 1073–1076.
- [11] ISHIZAKA A & LABIB A, 2011, *Review of the main developments in the analytic hierarchy process*, Expert Systems with Applications, **38(11)**, pp. 14336–14345.
- [12] JONES DF & MARDLE SJ, 2004, *A distance-metric methodology for the derivation of weights from a pairwise comparison matrix*, Journal of the Operational Research Society, **55(8)**, pp. 869–875.
- [13] KABERA MG, 2002, *A Statistical Approach to the Analytic Hierarchy Process*, MSc Thesis, University of Natal, South Africa.
- [14] KABERA MG & HAINES LM, 2010, *A statistical approach in the Analytic Hierarchy Process*, Proceedings of the 52nd Annual Conference of the South African Statistical Association, North-West University, Potchefstroom, pp. 28–35.
- [15] LAININEN P & HÄMÄLÄINEN RP, 2003, *Analyzing AHP-matrices by regression*, European Journal of Operational Research, **148(3)**, pp. 514–524.
- [16] MOSTELLER F, 1977, *Remarks on the method of paired comparisons: I. The least-squares solution assuming equal standard deviations and equal correlations*, Psychometrika, **16(1)**, pp. 3–9.
- [17] SAATY TL, 1977, *A scaling method for priorities in hierarchical structures*, Journal of Mathematical Psychology, **15(3)**, pp. 234–281.
- [18] SAATY TL, 1980, *The Analytic Hierarchy Process*, McGraw-Hill, New York (NY).
- [19] SRDJEVIC B, 2005, *Combining different prioritization methods in the analytic hierarchy process synthesis*, Computers and Operations Research, **32(7)**, pp. 1897–1919.
- [20] SRDJEVIC B & SRDJEVIC Z, 2011, *Bi-criteria evolution strategy in estimating weights from the AHP ratio-scale matrices*, Applied Mathematics and Computation, **218(4)**, pp. 1254–1266.
- [21] STERN H, 1990, *A continuum of paired comparisons models*, Biometrika, **77(2)**, pp. 265–273.
- [22] THURSTONE LL, 1927, *A law of comparative judgment*, Psychological Review, **34(4)**, pp. 273–289.
- [23] WANG Y-M, PARKAN C & LUO Y, 2007, *Priority estimation in the AHP through maximization of correlation coefficient*, Applied Mathematical Modelling, **31(12)**, pp. 2711–2718.
- [24] WANG Y-M, PARKAN C & LUO Y, 2008, *A linear programming method for generating the most favorable weights from a pairwise comparison matrix*, Computers and Operations Research, **35(12)**, pp. 3918–3930.