DigitalCommons @WayneState

Journal of Modern Applied Statistical Methods

Volume 12 | Issue 1 Article 12

5-1-2013

Priorities in Thurstone Scaling and Steady-State Probabilities in Markov Stochastic Modeling

Stan Lipovetsky
GfK Custom Research North America, Minneapolis, MN, stan.lipovetsky@gfk.com

Journal of Modern Applied Statistical Methods
May 2013, Vol. 12, No. 1, 65-77

1538 – 9472/13/$95.00

Priorities in Thurstone Scaling and Steady-State Probabilities in Markov Stochastic Modeling

Stan Lipovetsky
GfK Custom Research North America,
Minneapolis, MN

Thurstone scaling is widely used in marketing and advertising research where various methods of applied psychology are utilized. This article considers several analytical tools useful for positioning a set of items on a Thurstone scale via regression modeling and Markov stochastic processing in the form of Chapman-Kolmogorov equations. These approaches produce interval and ratio scales of preferences and enrich the possibilities of paired comparison estimation applied for solving practical problems of prioritization and probability of choice modeling.

Key words: Thurstone scale, regression estimation, Bradley-Terry model, Markov model, Chapman-Kolmogorov equations, steady-states probability.

Introduction

Thurstone scaling is a method of priority evaluation among items by the frequency of their empirical pairwise preferences (Thurstone 1927, 1959; Thurstone & Jones, 1957). This technique is widely used in fields of applied psychology, particularly, in marketing and advertising research (Edwards, 1957; Torgerson, 1958; Bock & Jones, 1968; Green & Tull, 1978; Conklin & Lipovetsky, 1999, 2004a, 2004b; Lipovetsky, 2007a, 2007b). Thurstone scaling transforms ranked or paired comparison data into a scale that is used for displaying the results of a ranking procedure. Statistical properties of Thurstone multiple comparisons were considered by Mosteller (1951) and Daniels (1950) and this technique is also known as Thurstone-Mosteller-Daniels (TMD) model (David, 1988; Stern, 1990; Ennis & Johnson, 1993). Connections of TMD with other methods of multiple paired comparisons, particularly, with the Analytic Hierarchy Process (AHP), are considered in several studies (Zinnes & MacKay, 1989; MakKay, Bowen & Zinnes, 1996; Lipovetsky & Conklin, 2001, 2002).

Positioning the items on a Thurstone scale consists in taking the proportions of respondents who prefer one item over each of the others, finding the corresponded percentiles (z-scores) of the cumulative normal distribution and averaging them. In practice it is convenient to rescale the obtained scores so that the best and the worst performing items will have scores 1 and 0, respectively. A Thurstone scale is typically constructed from ranked data when it is determined how often one item ranked ahead of another one, thus, the data could be reduced to or collected as paired comparisons and their frequencies. The paired approach to analysis also means that it is not required for every respondent to have ranked or compared every item; however, because the result is a relative scale, it is important that the pairwise comparisons be balanced. Thurstone scales can also be created from rating data, although this approach can produce a large number of ties that make the TMD unstable.

Stan Lipovetsky, Ph.D., is Senior Research Director at the GfK Research Center for Excellence, Marketing Sciences. He has numerous publications in multivariate statistics, multiple criteria decision making, econometrics, microeconomics and marketing research. He is a member of the editorial boards of the International Journal of Operations and Quantitative Management, the Journal of Electronic Modeling and the Journal of Model Assisted Statistics and Applications. Email him at: stan.lipovetsky@gfk.com.

This article considers several possibilities of priority estimation by pairwise data. One is evaluation of the TMD scale and the statistical significance of the obtained levels. For this purpose TMD is as presented a regression model by a special design of dummy variables. In constructing such a regression the standard errors and t-statistics for the levels of the compared items are obtained simultaneously in order to estimate precision and statistical significance of the differences among the items. The Thurstone model defines a scale of differences; standardizing to zero-one range corresponds to the interval scale. Together with the TMD model, the Bradley-Terry-Luce (BTL) model is also considered for pair comparison (Bradley & Terry, 1952; Luce (1959); Luce & Suppes, 1965; Lipovetsky, 2008) that corresponds to applying the logistic as opposed to the normal probability function.

Another possibility for pair comparisons evaluation may be found in stochastic Markov chain modeling via Chapman-Kolmogorov equations for discrete states and continuous time of transitions probabilities (Bellman, 1960; Hillier & Lieberman, 1974; Bar-Niv & Lipovetsky, 1995; Lipovetsky, 2005, 2006). This approach uses pair comparison data for intensity of transitions among the states (items) for constructing Chapman-Kolmogorov system of differential equations and solving for the dynamic as well as for the eventually reached steady-state probabilities. Although in the Thurstone model only differences are meaningful, the Markov states approach elaborates a ratio scale of probabilities to choose each of the items in comparison. Thus, the Thurstone and Markov models correspond to relative and absolute preference estimates.

Thurstone Scale as a Regression Model

The TMD general model is defined due by Thurstone's law of comparative judgment. According to Thurstone, a psychological characteristic x_i (where i = 1, 2, …, m denotes different characteristics) can be presented as a random normal variable $x_i = N(v_i\ ,\ \sigma_i\ )$ with a mean value v_i and standard deviation σ_i . The problem consists in estimating v_i values as the positions of the stimulus on the total psychological scale. The random variable of the difference $y = x_i - x_j$ between two psychological values (stimulus) has probability density function

$$p(y) = \frac{1}{\sqrt{2\pi}\bar{\sigma}} \exp\left(- \frac{\left(y - (v_i - v_j)\right)^2}{2\bar{\sigma}^2} \right), \tag{1}$$

where the standard deviation for the difference of two stimulus is

$$\bar{\sigma} = (\sigma_i^2 + \sigma_j^2 - 2\sigma_i \sigma_j r_{ij})^{1/2} \tag{2}$$

with r_{ij} denoting the correlation between i^{th} and j^{th} variables. The cumulative probability is then:

$$p_{ij} = \Phi(v_i - v_j) = \frac{1}{\sqrt{2\pi}} \int_{-(v_i - v_j)/\bar{\sigma}}^{\infty} e^{-y^2/2} dy\ . \tag{3}$$

The Case-V of TMD, which is the most widely used of these models, corresponds to equal standard deviations $\bar{\sigma}$ (2) for all paired differences of stimulus (it is fulfilled by the assumption of equal variances of the independent/uncorrelated variables).

If the values p_{ij} at the left-hand side of (3) are given, then the corresponded quantiles, or z-values can be defined as

$$z_{ij} = \Phi^{-1}(p_{ij}) = v_i - v_j\ . \tag{4}$$

In (4) there are more equations for the pairs ($i > j$, j = 1, …, m−1) than m values v_i themselves. For estimation of v_i values Mosteller (1951) suggested to use the Least Squares (LS) objective:

$$LS = \sum_{i \neq j}^{m} \left(z_{ij} - (v_i - v_j)\right)^2 \rightarrow \min\ . \tag{5}$$

The objective (5) is homogeneous by the parameters v_i of estimation, therefore, it needs a normalizing condition:

$$\sum_{i=1}^{m} v_i = 0\ . \tag{6}$$

The first order condition $dLS / d\ v_i = 0$ for minimizing (5) yields the estimate

$$v_i = \frac{1}{m}\sum_{j=1}^{m} z_{ij} + \frac{1}{m}\sum_{j=1}^{m} v_j = \frac{1}{m}\sum_{j=1}^{m} z_{ij} \ , \tag{7}$$

where the relation (6) is accounted. So the position on the psychological scale for each i^{th} item equals the mean value of *z*-scores of comparison of this item with the others.

In practical TMD modeling (for example, in comparison of a product's flavors or brands) as opposed to probabilities p_{ij} (3), sample estimates of frequencies corresponding to the observed proportions of cases with item *j* preferred to item *i* are available. These frequencies are usually presented in a matrix

$$F = \begin{pmatrix} f_{11} & f_{12} & \cdots & f_{1m} \\ f_{21} & f_{22} & \cdots & f_{2m} \\ \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots \\ f_{m1} & f_{m2} & \cdots & f_{mm} \end{pmatrix} \tag{8}$$

where each element f_{ij} corresponds to the preference of the item *j* over the item *i*. If in the pair comparison n_i respondents preferred the i^{th} item and n_j respondents preferred the j^{th} one, then ij^{th} and ji^{th} elements of frequency matrix (8) are

$$f_{ij} = \frac{n_j}{n_i + n_j}, \quad f_{ji} = \frac{n_i}{n_i + n_j}. \tag{9}$$

In a general case the frequencies can be obtained by a different number of respondents in each pair comparison. The elements in (8) are positive and satisfy the relation of symmetry

$$f_{ij} + f_{ji} = 1 \ , \tag{10}$$

thus, the diagonal elements are $f_{ii} = 0.5$. The quantiles z_{ij} of normal distribution (4) obtained for the elements of matrix (8) are

$$z_{ij} = \Phi^{-1}(f_{ij}) . \tag{11}$$

The values $z_{ij} = -z_{ji}$ and $z_{ii} = 0$, thus they define the elements of a skew-symmetric matrix *Z*. Due to the definition of the matrix (8) with elements (9), where each element f_{ij} corresponds to the prevalence of the item *j* over the item *i*, the means in the columns of matrix *Z* correspond to the estimates (7) obtained by the empirical frequencies (8), that is,

$$v_j = \frac{1}{m}\sum_{i=1}^{m} z_{ij} = \frac{1}{m}\sum_{i=1}^{m} \Phi^{-1}(f_{ij}) . \tag{12}$$

The total of these means equals zero, so condition (6) is satisfied. Thus, the averaged z-values (12) are used as positions of items on the Thurston scale of preferences. These values are usually reduced to the standard zero-one scale of preferences by the transformation

$$\tilde{v}_j = \frac{v_j - \min(v)}{\max(v) - \min(v)} . \tag{13}$$

Using (13) a Thurstone scale is reduced to an interval scale.

Consider the Bradley-Terry-Luce (BTL) model. The BTL model for pair comparisons defines probability that an item *a* is preferred to an item *b* as a share *p(a,b)=v(a)/(v(a)+v(b))* where *v* denotes a utility function. Using a logarithmic scale *A = ln(v(a))* and *B = ln(v(b))* results in *p(a,b) = exp(A)/(exp(A)+exp(B)) =1/(1+exp(-(A-B)))* that is a logistic probability function. The standardized logistic cumulative probability

$$p = \frac{1}{1 + \exp(-\gamma z)} \tag{14}$$

practically coincides with the standardized cumulative normal probability (3) when the parameter $\gamma = \pi / \sqrt{3} \approx 1.81$; this choice defines the logistic probability density function with a unit variance (see Long, 1997, chapter 3). Thus, for a simpler estimation the logistic as opposed to normal probability can be used when the *z*-value defined from (14) equals

$$z = \frac{1}{\gamma} \ln \frac{p}{1-p} . \qquad (15)$$

Using in (12) z (15) defined by the empirical frequencies in (8)-(9) results in the values:

$$\begin{aligned} v_j &= \frac{1}{m} \sum_{i=1}^{m} z_{ij} \\ &= \frac{1}{\gamma m} \sum_{i=1}^{m} ln \frac{f_{ij}}{1-f_{ij}} \qquad (16) \\ &= \frac{\sqrt{3}}{\pi} \, ln \left(\prod_{i=1}^{m} \frac{f_{ij}}{1-f_{ij}} \right)^{1/m} . \end{aligned}$$

Therefore, the Thurstonian logistic scale for a j^{th} item in comparison is proportional to the logarithm of the geometric mean of odds of the frequency in the j^{th} column of matrix (8). This solution is convenient for analytical consideration.

Returning to LS objective (5) that yields solution (12) for a matrix of paired comparison (8) notice that it corresponds to minimizing deviations for the linear regression:

$$\begin{aligned} z_{ij} &= v_j - v_i + \varepsilon_{ij} \\ &= a_1 u_{ij,1} + a_2 u_{ij,2} + \ldots + a_m u_{ij,m} + \varepsilon_{ij} \end{aligned} , \qquad (17)$$

where each ij^{th} value of the dependent variable z_{ij} is represented by its theoretical model v_j-v_i and random noise ε_{ij} . The theoretical model consists of a set of m dummy variables $u_{ij,1}$, …, $u_{ij,m}$ combined with the regression coefficients a_1, …, a_m . If all z_{ij} values are stacked into one vector of m^2 observations by the dependent variable and all vectors of the variables $u_{ij,1}$,…, $u_{ij,m}$ are arranged as a design matrix U of m^2 by m order, then in the row of matrix U defined by any ij^{th} pair of indices the only non-zero elements are in the j^{th} and i^{th} columns, and they equal 1 and –1, respectively. Therefore the dummy variables can be defined as:

$$u_{ij,k} = \begin{cases} +1, & if \quad k = j \\ -1, & if \quad k = i \end{cases} \qquad (18)$$

where k = 1, …, m corresponds to different dummies.

For linear regression model (17) with the predictors (18) the least squares objective is

$$\begin{aligned} LS &= \sum_{i,j=1}^{m} \left(z_{ij} - a_1 u_{ij,1} - a_2 u_{ij,2} - \ldots - a_m u_{ij,m} \right)^2 . \\ &= \rightarrow min \end{aligned} \qquad (19)$$

The coefficients a_k are the estimates of the Thurstone scale levels v_k. The totals in each row of the design matrix (18) equal zero, thus, this matrix has the rank m−1 and in regression modeling it is only necessary to use m−1 dummy variables. One of the coefficients, for example, a_m=0, can be fixed to construct the regression by other m−1 variables, and then renormalize all m coefficients by condition (6).

Numerically the coefficients of regression (17) or (19) coincide with the explicit solution (12), the regression approach, however, yields much richer results. To name some of them, besides the coefficients themselves, their standard errors and t-statistics, the coefficient of multiple determination as a characteristic of the quality of the approximation, deviations in each point of observation, etc., are obtained. The statistical difference between the Thurston scale levels can be checked, or the minimum distance found, between the significantly different levels.

In applied research with a large number of items, the pair comparison is usually arranged by experimental design when each respondent compares several items (not all) from a total set. In this case the frequencies (9) can correspond to different numbers of respondents in each paired comparison. Suppose, there are $n_{ij} = n_i + n_j$ counts in the ij^{th} pair of the items, so the variance of the proportion in this comparison equals

$$\sigma_{ij}^2 = \frac{f_{ij}(1-f_{ij})}{n_{ij}-1} . \qquad (20)$$

The standardized normal probability density function can then be written as

$$dp = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2} \right) dz , \qquad (21)$$

thus, due to the rule of error propagation the relation for the variances is

$$\sigma^2(z) = 2\pi \exp\left(z^2\right) \sigma^2(p) . \qquad (22)$$

Taking (20) for the variance of empirical frequency as the estimate for the variance $\sigma^2(p)$ in (22) results in the variance for z-values in each ij^{th} pair comparison. In place of (19) the Gauss-Markov weighted least squares objective can then be used

$$LS = \sum_{i,j=1}^{m} w_{ij} \left(z_{ij} - a_1 u_{ij,1} - a_2 u_{ij,2} - \ldots - a_m u_{ij,m} \right)^2 \\ = \rightarrow min \qquad (23)$$

with the weights of the observations defined as follows:

$$w_{ij} = \frac{1}{\sigma^2(z_{ij})} \\ = \frac{1}{2\pi} exp\left(-z_{ij}^2\right) \frac{n_{ij} - 1}{f_{ij}(1 - f_{ij})} . \qquad (24)$$

The weighted regression (23)-(24) can be constructed without difficulty and used as a weighted estimation for the TMD model.

Stochastic Modeling by Chapman-Kolmogorov Equations.

Return to frequency matrix (8)-(9) and consider other possibilities to estimate preferences among compared items. In the approach developed in (Lipovetsky & Conklin, 2002) the analytic hierarchy process (AHP) matrix of pairwise ratios was transformed to a share matrix with the elements of the kind (9), and a specific eigenproblem was designed for evaluating the priorities among the items. The results of that work applied to a Thurstone matrix (8) (that corresponds to the transposition of a transformed AHP matrix) can be presented in the eigenproblem:

$$[F' + diag(F'e)]\, \alpha = \lambda \alpha , \qquad (25)$$

where prime denotes transposition of the matrix F (8), $diag(F'e)$ is a diagonal matrix of the totals in the columns of matrix F, and e denotes a uniform vector of the m^{th} order. Solving (25) for the maximum eigenvalue λ yields the estimate α for the priority vector.

The matrix at the left-hand side (25) is proportional to a transposed stochastic matrix. It means that totals in the columns of this matrix equal the following vector:

$$[F' + diag(F'e)]'\, e = Fe + F'e = m\, e , \qquad (26)$$

where the property (10) is used, so each element of the vector (26) equals m. Dividing (25) by this term the eigenproblem is represented as

$$\left(\frac{1}{m} F' + \frac{1}{m} diag(F'e) \right) \alpha = \mu \alpha . \qquad (27)$$

Totals in every column of matrix (27) equal one. A positive matrix with such property is a transposed stochastic matrix. Such matrices have a maximum eigenvalue equal to one. Due to the Perron-Frobenius theory for positive matrices the principal eigenvector always exists, is unique and has all positive elements; the desired properties of the priority vector are thus ensured.

Consider the eigenproblem (27) from the point of view of Markov chain modeling – one of the most widely used tools in theoretical and applied statistics. A discrete state and continuous time model are presented via a system of Chapman-Kolmogorov differential equations used for a stochastic process of transitions among the states. These well-known (especially in queuing theory) equations express change in probability to be found in any of m states as a linear combination of these probabilities with the coefficients of the transition intensities.

Any pair of elements f_{ij} and f_{ji} (9) of Thurstone matrix (8) can be interpreted in terms of frequency to prefer one of the items over the other. Thus, each element f_{ij} can be used to describe the preference of the j^{th} item over the i^{th} item that corresponds to transition to the preferred state j from the state i with the intensity of transition f_{ij}. The frequency matrix F (8) can be presented as a connected oriented graph with m nodes of states (items) and two edges between each pair of nodes – the one going to state j from state i corresponds to transition intensity f_{ij}, and the other going from state j to state i corresponds to transition intensity f_{ji}. An example of such graph for three states is presented in Figure 1.

The system of Chapman-Kolmogorov equations can thus be written as:

$$\frac{dp_k}{dt} = \sum_{i \neq k}^{m} f_{ik} p_i - \sum_{j \neq k}^{m} f_{kj} p_j, \quad k = 1,...,m, \tag{28}$$

where p_i denotes probability of belonging to each of the states. Items with positive signs at the right-hand side (28) define influx to each state from all the others and those with negative signs define departure from a state to all the other states. If $0.5p_k$ is added into each sum in each k^{th} equation (28) this system can be represented in a matrix form:

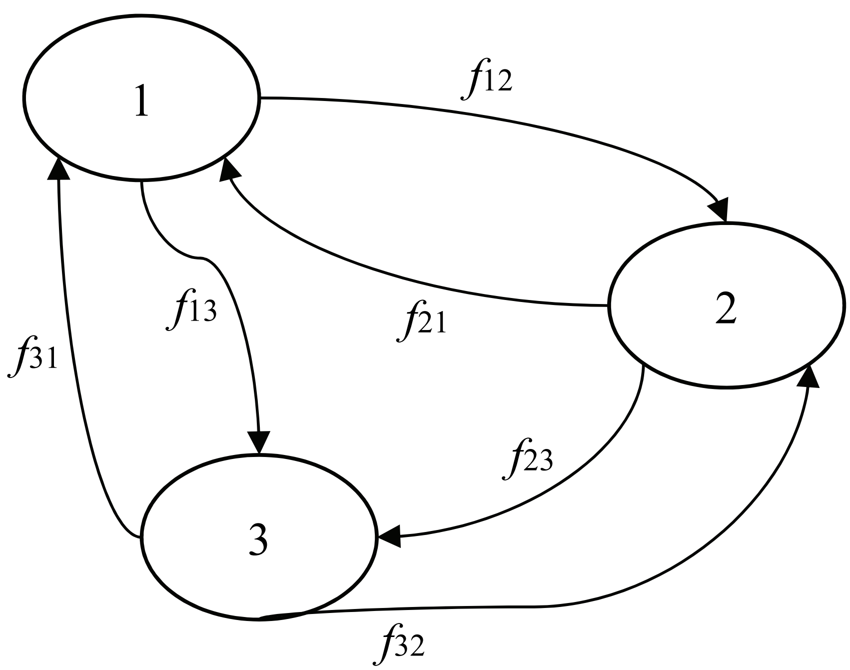
$$\dot{p} = \left(F' - diag(Fe)\right) p, \tag{29}$$

where p is a vector consisting of the probabilities p_i for all the states, $\dot{p}$ denotes the vector of their derivatives (as in the left-hand side (28)). Using property (26) that sum of totals in any k^{th} column and row of the matrix F equals m, (29) can be rewritten as

$$\dot{p} = \left(F' + diag(F'e) - mI\right) p, \tag{30}$$

where I denotes the identity matrix of the m^{th} order.

Figure 1: Transition Intensities for Markov Modeling



Consider the solution of the Chapman-Kolmogorov equations for the steady-state probability when the process is stabilized. If the derivatives in the right-hand side (30) equal zero, then (30) is reduced to

$$(F' + diag\,(F'\,e)\,)\,p = m\,p\,, \qquad (31)$$

which is the same as eigenproblem (25) with the maximum eigenvalue equal m and a unique positive main eigenvector, as discussed in relation to expression (27). Coincidence of these results suggests a useful interpretation: the priority vector (27) makes sense of the eventual probabilities to belong to the discrete states corresponded to the compared items. These probabilities define the preferences among the compared items.

The general dynamic solution for system (30) can be useful for problems in priority modeling. For example, item preference depending on different initial conditions is of interest in maximum differences among the preferences and in their specific behavior (monotonic increase or decrease, oscillating) before the process is stabilized. Priorities behavior after adding new items to the original set may also be considered, taking initial probabilities of the new items equal zero and assuming the pair relations kept due to the given data.

As Bellman (1960) showed, the solution of a homogeneous linear system of differential equations with constant coefficients can be presented as

$$p(t) = P\,diag(\exp(\lambda_j t))\,c\ , \qquad (32)$$

where c is a vector of constants, λ_j are the eigenvalues and P is a corresponded matrix of columns p_j of eigenvectors obtained in solving the problem

$$\left(F' + diag(F'e) - mI\right)p = \lambda\,p\,. \qquad (33)$$

Expression (33) defines the eigenproblem with the matrix at the right-hand side of Chapman-Kolmogorov system (30), and its solution coincides with the solution for problem (25) up to reducing the latter eigenvalues by m.

For the moment $t = 0$, solution (32) reduces to $p(0) = Pc$, and solving this linear system with a known vector of initial conditions $p(0)$ results in the vector of constants $c = P^{-1}p(0)$. Thus, the general solution of the differential system is

$$p(t) = P\,diag(\exp(\lambda_j t))\,P^{-1}\,p(0)\,, \qquad (34)$$

The expression $P\,diag(\exp(\lambda_j t))\,P^{-1}$ in (34) is known as matrix exponent. Each component of the vector $p(t)$ is a linear combination of the exponents in (34), and functions $\exp(\lambda_j t)$ behave in accordance with the specific values of λ_j obtained in eigenproblem (33).

As noted, the main eigenvalue in (33) is less by m than the main eigenvalue in (25), so it equals zero, $\lambda_1 = 0$, which corresponds to the constant part of (34) behavior. The other eigenvalues (33) are real numbers or conjugated pairs of complex numbers. As per the Perron-Frobenius theory, all other eigenvalues have less real value than the main eigenvalue, meaning that all real eigenvalues, or real parts in complex eigenvalues, are negative. Thus, the general behavior of solution (34) is defined by a constant part ($\lambda_1 = 0$), by diminishing exponents (real negative eigenvalues), and by oscillating diminishing exponents (complex eigenvalues giving sine and cosine parts of functions). There can also be polynomial items corresponding to equal eigenvalues, although in practical numerical evaluations such cases are rare. The eigenvectors p corresponding to the complex eigenvalues are also complex, however, the total expression (34) yields real values.

The total of the eigenvalues equals the trace of the matrix, which for the matrix in (33) is

$$\begin{aligned} -\sum_{j=1}^{m}\lambda_j &= -Tr\left(F' + diag(F'e) - mI\right) \\ &= -\frac{m}{2} - \frac{m^2}{2} + m^2 \\ &= \frac{m(m-1)}{2} \end{aligned} \qquad (35)$$

The inverted values of the reciprocal intensities in the exponents (34) make sense of the mean time of the transitions among the states, $-1/\lambda_j = \bar{t}_j$, similarly to the interpretation of the parameters in the exponential and Erlang distributions from queueing theory. Thus, (35) corresponds to an ergodic relation saying that the total of the intensities equals the number of connections *m(m-1)/2* among *m* states. The mean intensity can be defined from (35) as

$$\bar{\lambda} = -\frac{1}{m}\sum_{j=1}^{m}\lambda_j = \frac{m-1}{2}, \qquad (36)$$

that corresponds to a mean number of links from initial to final states (Levene & Loizou, 2002). A value λ_j defines the decay of the exponent, therefore, several first eigenvalues from the main λ_1 =0 to about the mean (36) are important in the solution (34). It is interesting to note that the relation (36) can be also interpreted as a harmonic mean for the mean times:

$$\frac{1}{H} = \frac{1}{m}\sum_{j=1}^{m}\frac{1}{\bar{t}_j} = \frac{m-1}{2}. \qquad (37)$$

The mean time for the exponential decay is about 2/(m−1), and after this time the process is stabilizing.

Table 1: Flavors Tested

Code	Flavor
A	Hot Fudge Swirl
B	Caramel Swirl
C	Cheesecake Swirl
D	Walnut
E	Chocolate Chunk
F	Double Chocolate
G	Peanut Butter Chunk
H	Peanut Butter Frosted
I	Peanut Butter Swirl
J	White Chocolate Chunk
K	German Chocolate
L	Raspberry Swirl
M	Snickers
N	Chocolate Frosted
O	Mocha Swirl
P	Chocolate Chip Blonde
Q	S'mores
R	Bailey's Irish Crème Swirl
S	Icing
T	Frosted Mint

Numerical Example

Consider a numerical example using data from a real research project. Twenty flavors (see Table 1) of a snack were ranked by 151 respondents. The first 6 flavors (A to F) were in the current line and the others 14 flavors (G to T) were considered for possible addition to the production line.

The results of the modeling are presented in Table 2. The table shows Thurstone Scaling – coefficients of regression (19), their t-statistics, raw Thurstone scale (12), standard scale in 0-1 range (13), and ranks of the flavors. In the regression approach (12) the upper triangle (190 values) of the z-matrix with elements (11) is used, m−1 dummies (18) without the first flavor A, and no intercept. Centering the coefficients of regression by subtracting their mean from each, the coefficients are rescaled until their total equals zero (6). The rescaled coefficients of regression coincide with the raw Thurstone scale (12). In regression, t-statistics are also obtained for the coefficients that correspond to checking the significance of the difference between the level of each flavor in comparison with the first flavor A (thus, except for flavor E, all the others are different from flavor A). Flavor A was best one on the Thurstone scale, this explains the negative signs of the regression coefficients and t-values. If any other flavor is excluded from the regression, its coefficient would be of positive and negative signs for the flavors preferred and non-preferred to this one, respectively, and the t-values would estimate the significance of the difference of all the levels from the fixed one. Coefficient of multiple determination in the model equals $R^2 = 0.966$, so approximation of the data by the Thurstone scale is very high. If the total z-matrix is used to construct a regression (19) with an intercept, again by $m-1$ dummies (18), then the coefficients of regression will be the same as described above. Also in the rescaled set of coefficients the value of the intercept equaled the coefficient for the excluded flavor. The results of the weighted regression (23)-(24) practically coincide with the regular regression by this data.

Table 2 also presents a Thurstonian estimation by the logistic BTL model (16) and its standard scaling by (13) to 0-1 range. Note that the raw logistic scale slightly differs from the Thurstone raw scale, however, the standardized scales in 0-1 range are practically undistinguishable and both the normal and logistic estimations also yield the same ranks.

Finally, Table 2 shows the results of Markov modeling for the eventually reached steady-state probabilities. These probabilities correspond to the elements of the main eigenvector in problems (25), (31) or (33) and define shares of choosing each of the flavors under consideration. Ranks of flavors by the Markov probabilities coincide with the ranks by the Thurstone scale for this data.

Reaching the steady-states in Markov processing can be considered by solution (34) of the Chapman-Kolmogorov differential equations (28). At first the behavior of the current six flavors in the line is constructed, using initial conditions with equal probability for all 6 states (see Figure 2). The reached steady-state probabilities in this set are 0.219, 0.123, 0.135, 0.123, 0.205, and 0.195, for the flavors from A to F, respectively. Using these probabilities as initial values for the current flavors and zero initial values for other possibly introduced flavors, another Markov model by all 20 flavors is constructed (see Figure 3). Note that 14 possibly introduced flavors would push down the current flavors' shares. The flavors A, E and F remain best, however, two new flavors – M and N – could become more attractive than the other current flavors. Thus, the mutual behavior of all current flavors are considered with these two best candidates for the line extension (see Figure 4). Figure 4 shows that newly introduced flavors M and N can overcome three of the current flavors, thus, it makes sense to substitute the current B and D flavors for these new ones if the size of the line will continue to consist of only six flavors. It is interesting to note that the eigenvalues (34) in all these eigenproblems are real numbers so the flavor curves behavior consists in just exponential change, without oscillations corresponding to the complex numbers. This indicates a consistent relation among the pair comparison data and the robust results of both Thurstone and Markov evaluations.

Conclusion

This article considered preference evaluation by pair comparison data. Thurstonian scaling via multiple regression and Markov chain modeling by Chapman-Kolmogorov equations was explored. A Thurstone scale as a regression model a special design of dummy variables was used for estimation. Coefficients of regression represent the levels of the items by the Thurstone scale. Simultaneously the standard errors and t-statistics for the coefficients of regression were obtained along with the coefficient of multiple determination so that precision and statistical significance of the differences among the items could be estimated.

The Thurstone model defines a scale of differences, and its standardized zero-one range corresponds to the interval Thurstone scale. With regression, non-linear scaling can be considered, a hierarchy Bayesian model using other variables (for example, demographics) or any other technique known in regression modeling. Also considered was the Bradley-Terry-Luce logistic model of pair comparison that produces a scaling of the Thurstonian type with the results very close to the Thurstone-Mosteller-Daniels model.

Another possibility for multiple pair comparison evaluation was suggested based on stochastic Markov chain modeling for discrete states and continuous time of transitions probabilities. This approach uses pair comparisons data for intensity of transitions among the states (items) for constructing the Chapman-Kolmogorov system of differential equations and solving for the dynamic as well as for the eventually reached steady-state probabilities. While in the Thurstone model only differences are meaningful, the Markov states approach elaborates a ratio scale of probabilities to choose each of the items in comparison. Thus, the Thurstone and Markov models correspond to the relative and absolute preference estimates. The considered methods of priority evaluation are convenient and simple and could enrich both theoretical modeling and practical applications for various multiple criteria decision making problems.

References

Bar-Niv, M., & Lipovetsky, S. (1995). Deciding circular priorities in insolvency situations, *International Journal of Operations and Quantitative Management*, *1*, 183-195.

Bellman, R. (1960). *Introduction to matrix analysis*. New York, NY: McGraw-Hill.

Table 2: Thurstone Scale, Logistic Scale and Markov States

Flavor	Thurstone Scale					Logistic Scale		Markov State
	Regression Coefficients	t-statistics	Raw Scale	Range 0-1	Rank	Raw Scale	Range 0-1	Probability %
A. Hot Fudge Swirl	0	0	0.49	1.00	1	0.44	1.00	9.93
B. Caramel Swirl	-0.36	-14.1	0.12	0.64	7	0.11	0.63	5.51
C. Cheesecake Swirl	-0.35	-13.4	0.14	0.65	6	0.13	0.65	5.75
D. Walnut	-0.38	-14.8	0.11	0.62	8	0.09	0.61	5.42
E. Chocolate Chunk	-0.01	-0.3	0.48	0.99	2	0.43	0.99	9.75
F. Double Chocolate	-0.05	-2.1	0.44	0.95	3	0.39	0.95	9.13
G. Peanut Butter Chunk	-0.47	-18.2	0.02	0.53	10	0.02	0.53	4.59
H. Peanut Butter Frosted	-0.80	-30.3	-0.31	0.20	18	-0.28	0.20	2.72
I. Peanut Butter Swirl	-0.42	-16.3	0.07	0.58	9	0.06	0.58	4.99
J. White Chocolate Chunk	-0.55	-21.4	-0.06	0.45	11	-0.06	0.45	4.04
K. German Chocolate	-0.60	-23.2	-0.11	0.40	13	-0.10	0.40	3.86
L. Raspberry Swirl	-1.00	-38.6	-0.51	0.00	20	-0.46	0.00	1.97
M. Snickers	-0.33	-12.6	0.16	0.67	5	0.15	0.67	5.90
N. Chocolate Frosted	-0.17	-6.5	0.32	0.83	4	0.29	0.83	7.39
O. Mocha Swirl	-0.71	-27.5	-0.22	0.29	17	-0.20	0.29	3.12
P. Chocolate Chip Blonde	-0.66	-25.3	-0.17	0.34	14	-0.15	0.35	3.43
Q. S'mores	-0.57	-21.9	-0.08	0.43	12	-0.07	0.43	3.96
R. Bailey's Irish Crème Swirl	-0.68	-26.2	-0.19	0.32	15	-0.17	0.33	3.37
S. Icing	-0.71	-27.4	-0.22	0.29	16	-0.20	0.29	3.11
T. Frosted Mint	-0.96	-37.2	-0.47	0.04	19	-0.43	0.04	2.07

Figure 2: State Probability of Current Six Flavors in Markov Model

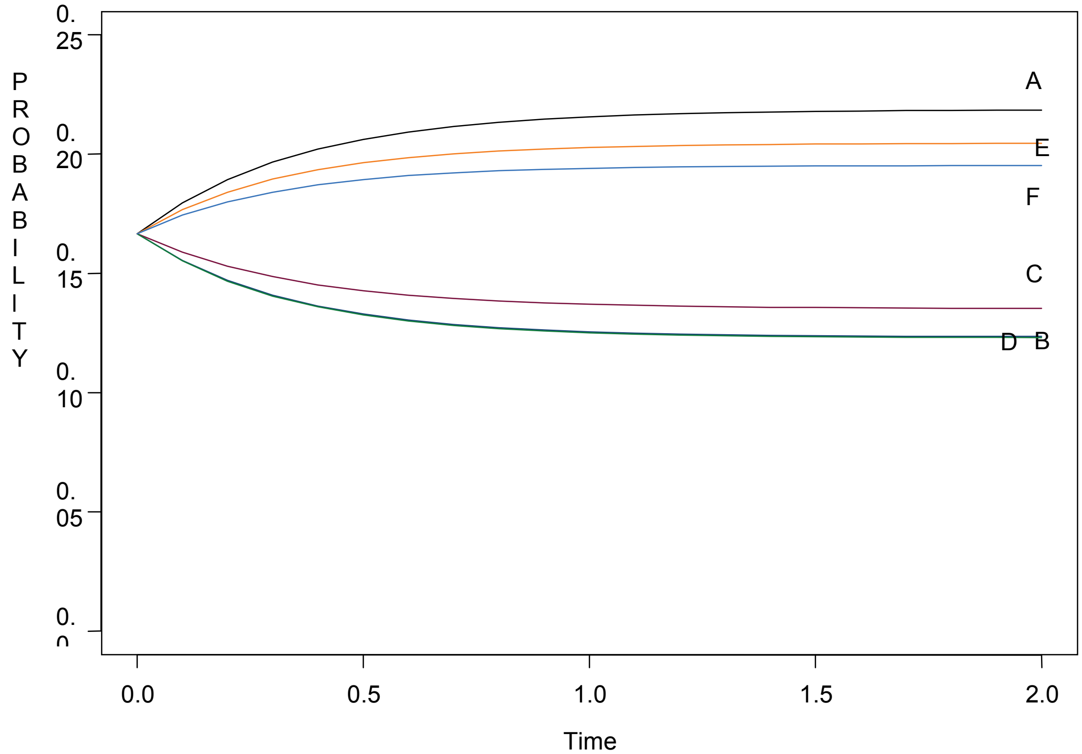


Figure 3: State Probability for 20 Flavors in Markov Model

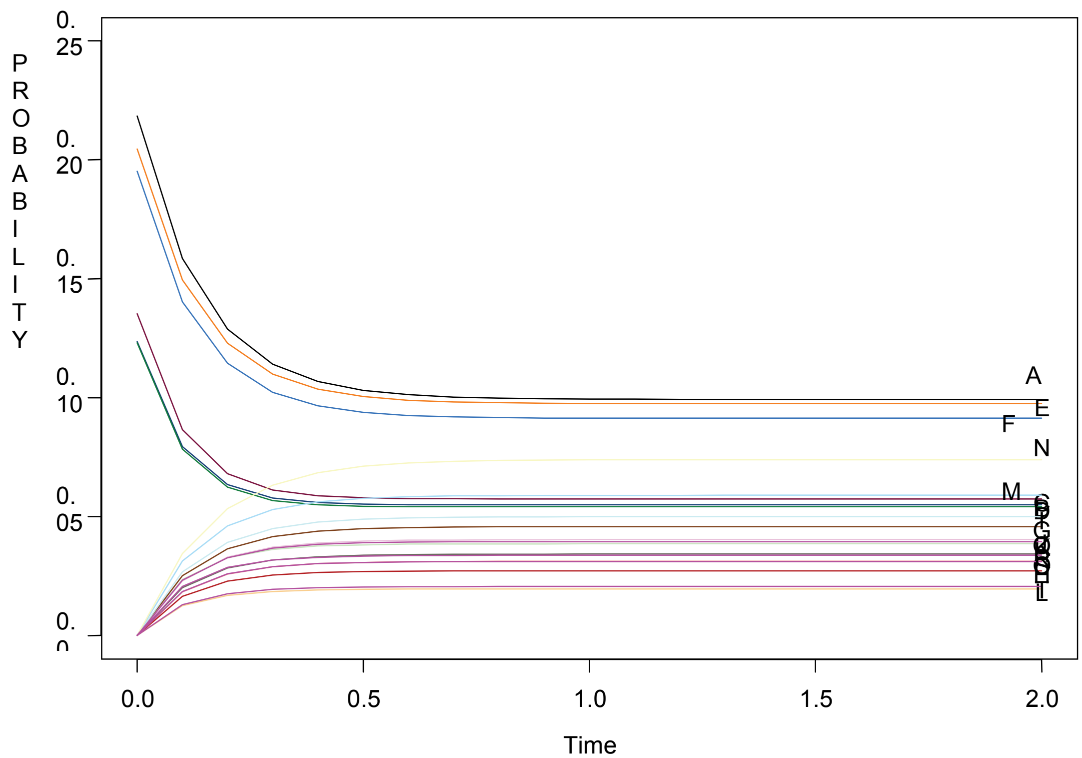
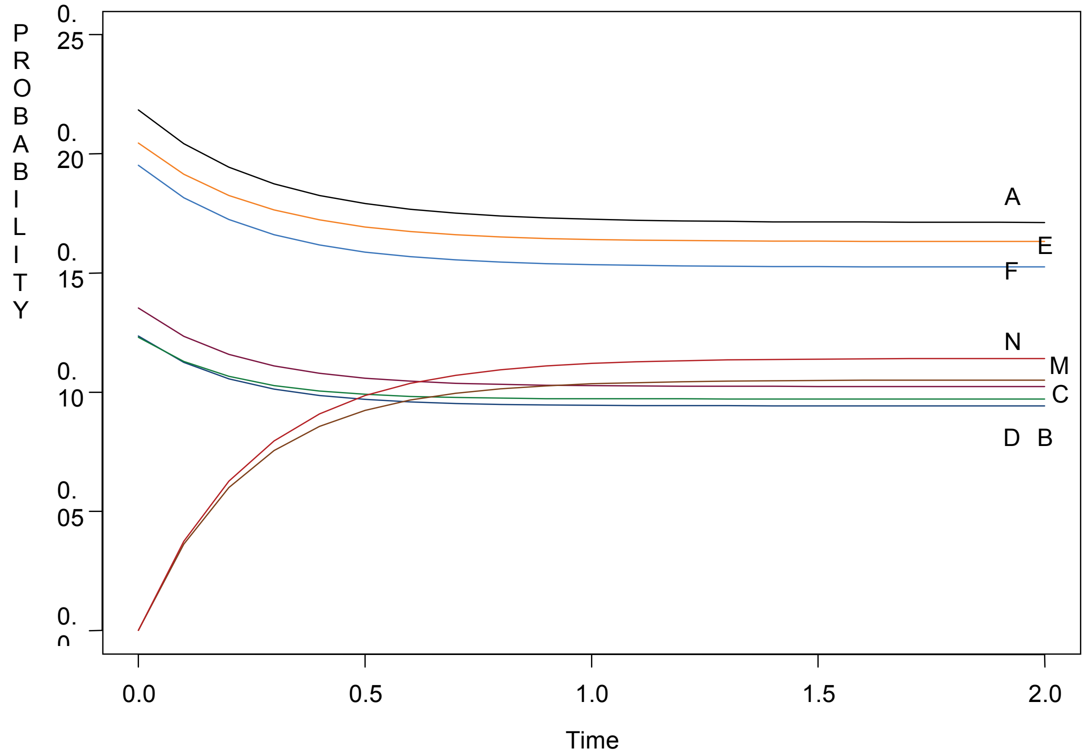


Figure 4: State Probability for Chosen Eight Flavors in Markov Model



Bock, R. D., & Jones, L. V. (1968). *The measurement and prediction of judgment and choice*. San Francisco, CA: Holden-Day.

Bradley, R. A., & Terry, M. E. (1952). Rank analysis of incomplete block designs: The method of paired comparisons. *Biometrika*, *39*, 324-345.

Conklin, M., & Lipovetsky, S. (1999). Efficient assessment of self-explicated importance using latent class Thurstone scaling. *The 10th Annual Advanced Research Techniques Forum*, Santa Fe, NM: American Marketing Association.

Daniels, H. E. (1950). Rank correlation and population models, *Journal of the Royal Statistical Society*, *Series B*(*12*), 171-181.

David, H. A. (1988). *The method of pair comparisons*, *2nd Ed*. London: Griffin.

Edwards, A. L. (1957). *Techniques of attitude scale construction*. New York, NY: Appleton-Century-Crofts.

Ennis, D. M., & Johnson, N. L. (1993). Thurstone-Shepard similarity models as special cases of moment generating functions. *Journal of Mathematical Psychology*, *37*,104-110.

Green, P. E., & Tull, D. S. (1978). *Research for marketing decisions*. New Jersey: Prentice-Hall.

Hillier, F. S., & Lieberman G. J. (1974). *Introduction to operations research*. San Francisco, CA: Holden-Day.

Levene, M., & Loizou, G. (2002). Kemeny's constant and the random surfer. *The American Mathematical Monthly*, *109*, 741-745.

Lipovetsky, S., & Conklin, M. (2001). Dual priority-antipriority Thurstone scales as AHP eigenvectors. *Engineering Simulation*, *18*, 631-648.

Lipovetsky, S., & Conklin, M. (2002). Robust estimation of priorities in the AHP. *European Journal of Operational Research*, *137*, 110-122.

Lipovetsky, S., & Conklin, M. (2004a). Thurstone scaling via binary response regression. *Statistical Methodology*, *1*, 93-104.

Lipovetsky, S., & Conklin, M. (2004b). Nonlinear Thurstone scaling via SVD and Gower plots. *International Journal of Operations and Quantitative Management*, *10*, 259-273.

Lipovetsky, S. (2005). Analytic hierarchy processing in Chapman-Kolmogorov equations. *International Journal of Operations and Quantitative Management*, *11*, 219-228.

Lipovetsky, S. (2006). Van Westendrop price sensitivity in statistical modeling, *International Journal of Operations and Quantitative Management*, *12*, 141-156.

Lipovetsky, S. (2007a). Thurstone Scaling in Order Statistics, *Mathematical and Computer Modelling*, *45*, 917-926.

Lipovetsky, S. (2007b). Priority and choice probability estimation by ranking, rating, and combined data. *Journal of Statistical Theory and Practice*, *1*, 265-278.

Lipovetsky, S. (2008). Bradley-Terry choice probability in maximum likelihood and eigenproblem Solutions. *International Journal of Information Technology & Decision Making*, *7*, 395-405.

Long, J. S. (1997). *Regression models for categorical and limited dependent variables*. London: SAGE Publications.

Luce, R. D. (1959). *Individual choice behavior: A theoretical analysis*. New York, NY: Wiley.

Luce, R. D., & Suppes, P. (1965). Utility, preference and subjective probability. In *Handbook of Mathematical Psychology*, R. D. Luce, R. R. Bush & E. Galanter (Eds.), 249-410. New York, NY: Wiley.

MakKay, D. B., Bowen, W. M., & Zinnes, J. L. (1996). A Thurstonian view of the analytic hierarchy process. *European Journal of Operational Research*, *89*, 427-444.

Mosteller, F. (1951). Remarks on the method of paired comparisons. *Psychometrika*, *16*, 203-218.

Stern, H. (1990). Models for distributions on permutations. *Journal of the American Statistical Association*, *85*, 558-564.

Thurstone, L. L. (1927). A law of comparative judgment. *Psychological Reviews*, *34*, 273-286.

Thurstone, L. L. (1959). *The measurement of values*. Chicago, IL: University of Chicago Press.

Thurstone, L. L., & Jones L. V. (1957). The rational origin for measuring subjective values. *Journal of the American Statistical Association*, *52*, 458-471.

Torgerson, W. S. (1958). *Theory and Methods of Scaling.* New York, NY: Wiley.

Zinnes, J. L., & MacKay, D. B. (1989). Probabilistic multidimensional analysis of preference ratio judgments. In *New developments in psychological choice modeling*, G. De Soete, H. Feger & K. C. Klauer (eds.), 177-205. Amsterdam: North-Holland.