

AD-766 096

STATISTICAL ESTIMATION OF THE COEFFICIENT OF PAIRWISE COMPATIBILITY FOR MENU ITEMS

Joseph L. Balintfy, et al

Massachusetts University

Prepared for:

Office of Naval Research

August 1973

DISTRIBUTED BY:

NTIS

National Technical Information Service
U. S. DEPARTMENT OF COMMERCE
5285 Port Royal Road, Springfield Va. 22151

STATISTICAL ESTIMATION OF THE COEFFICIENT
OF PAIRWISE COMPATIBILITY FOR MENU ITEMS

by

Joseph L. Balintfy

and

Ram C. Dahiya

Reproduced by
NATIONAL TECHNICAL
INFORMATION SERVICE
U.S. Department of Commerce
Springfield, VA 22151

STATISTICAL ESTIMATION OF THE COEFFICIENT
OF PAIRWISE COMPATIBILITY FOR MENU ITEMS

by

Joseph L. Balintfy

and

Ram C. Dahiya

Technical Report No. 10

August, 1973

Prepared under Contract N00014-67-A-0230-0006
(NR 047-100) for the Office of Naval Research

Reproduction in Whole or in Part is Permitted
for any Purpose of the United States Government

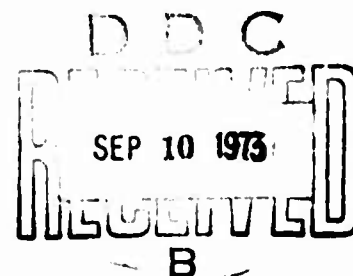
This document has been approved for public release and sale;
its distribution is unlimited

School of Business Administration
Department of General Business and Finance

and

School of Natural Sciences and Mathematics
Department of Mathematics and Statistics

University of Massachusetts
Amherst, Massachusetts



Unclassified

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) School of Business Administration University of Massachusetts at Amherst		2a. REPORT SECURITY CLASSIFICATION Unclassified	
		2b. GROUP Not Applicable	
3. REPORT TITLE Statistical Estimation of the Coefficient of Pairwise Compatibility for Menu Items.			
4. DESCRIPTIVE NOTES (Type of report and inclusive dates) Technical Report No. 10, August, 1973			
5. AUTHOR(S) (First name, middle initial, last name) Joseph L. Balintfy, Ram C. Dahiya			
6. REPORT DATE August, 1973		7a. TOTAL NO OF PAGES 21	7b. NO OF REFS 7
8a. CONTRACT OR GRANT NO N00014-67-A-0230-0006		9a. ORIGINATOR'S REPORT NUMBER(S) Technical Report No. 10	
b. PROJECT NO NR 047-100		9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report) Not Applicable	
c.			
d.			
10. DISTRIBUTION STATEMENT This document has been approved for public release and sale; its distribution is unlimited.			
11. SUPPLEMENTARY NOTES Not Applicable		12. SPONSORING MILITARY ACTIVITY Office of Naval Research, Washington, D.C.	
13. ABSTRACT The coefficient of compatibility between two menu items is defined as the difference between the conditional and marginal probabilities of selecting one after the other. Maximum likelihood estimates from selection statistics are given, and the asymptotic distribution of the estimator is determined so that test of null hypothesis and confidence intervals regarding the coefficients can be obtained. A method is described for building data base for large joint probability matrices from menu selection data.			

Unclassified

Security Classification

14. KEY WORDS	LINK A		LINK B		LINK
	ROLE	WT	ROLE	WT	ROLE
Food Preference					
Menu Planning					
Mathematical Statistics					
Food Service					
Data Banks					

STATISTICAL ESTIMATION OF THE COEFFICIENT
OF PAIRWISE COMPATIBILITY FOR MENU ITEMS

by

Joseph L. Balintfy

and

Ram C. Dahiya

I INTRODUCTION

With recent advancements in the art of representing food and meal preferences by mathematical models [2] along with the possibilities of defining and solving optimum human diet problems by mathematical programming techniques [1], difficult questions emerge concerning the role and measurability of compatibility between menu items.

A meal selected by, or planned for an individual usually consists of a set of menu items which are mutually complementary in the sense that each item represents one of the courses of the meal. It is tempting to consider the utility of a meal in terms of the utilities of its components. Recent results in multiattribute utility theory open the way for a variety of representations, and this is where the issue of compatibility comes in.

The simplest additive utility model would imply that the utility of a meal is equal to some weighted sum of the univariate utilities of the menu items in the meal. Additivity, therefore, means that the utility of the meal is completely explained by the utilities of the items irrespective

of their relative combination. Empirical evidence [5] is, however, overwhelming that this is not so. People find some combinations of items more or less compatible than others, meaning that compatibility is also a factor in the utility of a meal.

If the set of menu items under consideration is preference and therefore utility independent, a measure of compatibility can be derived from the multiattribute utility model of Keeney [6]. This model expresses the utility of a meal as a weighted sum of additive utilities plus a weighted sum of all possible crossproducts of univariate utilities. Depending upon the sign of the coefficients, this second sum may increase or decrease the value of additive utilities, and hence would represent the effect of compatibility in a given combination of items. Unfortunately, again, utility independence can not be assumed to be true for any set of menu items, and not even for any set of pairwise combinations of items. Consequently, research is still in progress to find the appropriate expression of compatibility in utility terms.

In a recent report by Balintfy and Sinha [3] a meal planning model is presented which does not require the representation of compatibility of menu items in the utility measure of the meals. The idea is advanced that a mathematical programming model which maximizes an additive utility function could determine which items would be most liked by a given population in a given time period, and from this fixed set of items a separate scheduling activity could combine the items into a sequence of meals on the criterion of compatibility alone. This way the concepts of the utility and compatibility of a meal are conveniently separated; the first being included only

in the objective function additively, while the second appears among the scheduling constraints, and hence is no longer linked directly to utility measures. Consequently, it can be represented by any measure for which data collection is feasible.

This paper suggests the introduction of a statistical measure for compatibility of menu items in the form of a coefficient of pairwise compatibility of menu items. This coefficient is defined on a cartesian product set of two sets of menu items where the sets under consideration are identified with two nonidentical courses of the meal, such as entrees and vegetables. Limiting the notion of compatibility to pairwise compatibility alone is arbitrary, but it is the first step of the investigation which may lead to further extensions later.

The first part of the paper provides the definition, some properties and examples of the coefficient of pairwise compatibility. The second part describes the method of estimation and the probability distribution of the estimator. The last part is devoted to the techniques of estimating coefficients for the whole cartesian product set from subsets of data as it is usually available through a sequence of selective menu schedules. The application of this coefficient to menu scheduling algorithms and experience with data collection will be the subject of later reports.

II PROBLEM DEFINITION

Let X represent a finite set of menu items such as entrees, and let Y represent another finite set of items such as vegetables. Then the Cartesian product set $U = [X \times Y]$ represents all pairwise combinations of menu items under consideration. The elements of U can be considered as elementary meals consisting of only two courses. If x_i is one of the entrees, i.e., $x_i \in X$, and y_j is the j -th vegetable, i.e., $y_j \in Y$, then the pairwise combination of items (x_i, y_j) denoted by u_{ij} is the element of U , i.e., $(x_i, y_j) = u_{ij} \in U$.

If one considers sets X and Y as choices of menu items, such as entrees and vegetables respectively, then it is conceivable that an individual will have a linear preference ordering at any point of time for all combinations (x_i, y_j) which is based on his utility for x_i , his utility for y_j and his utility for the compatibility between x_i and y_j in a meal. For a rational person, this preference ordering is revealed by his choice-behavior in selecting a particular combination of items in a sequence of trials. Consequently, one can assign probability measures for the sets X , Y and U as follows:

$$\begin{aligned} p(x_i) &\geq 0 \quad ; \quad \sum_{x_i \in X} p(x_i) = 1 \\ p(y_j) &\geq 0 \quad ; \quad \sum_{y_j \in Y} p(y_j) = 1 \\ p(x_i, y_j) &\geq 0 \quad ; \quad \sum_{u_{ij} \in U} p(x_i, y_j) = 1 \end{aligned}$$

For the sake of notational simplification $p(x_i) = p_i$, $p(y_j) = p_j$ and $p(x_i, y_j) = p_{ij}$ notations also will be used in subsequent parts of the paper. The first two of these notations will be referred to as the marginal probability of selection, while the last term is called the joint probability of selection. These terms are related by the well known identities

$$p(x_i) = \sum_{y_j \in Y} p(x_i, y_j)$$

$$p(y_j) = \sum_{x_i \in X} p(x_i, y_j)$$

Decisions in choosing any item combination from set U are usually made sequentially in the sense that a choice is made, say from set X first (entrees), followed by a choice from set Y , implying that the choice of y_j may be conditioned by the choice of a particular x_i . Consequently, one can talk about the conditional probability of selecting y_j , given that x_i is already selected according to the formula

$$p(y_j | x_i) = \frac{p(x_i, y_j)}{p(x_i)}$$

which means that the conditional probability of selections is completely defined by the corresponding marginal and joint probability measures. Furthermore, it is also known that if $p(x_i, y_j) = p(x_i)p(y_j)$, then $p(y_j | x_i) = p(y_j)$, which means that the selection of y_j is unconditional of the choice of x_i .

It is proposed that a coefficient of pairwise compatibility of menu items be created on the basis of the above discussed conditional dependence

criterion. Accordingly a coefficient θ_{ij} is defined as

$$(1) \quad \theta_{ij} = p(y_j | x_i) - p(y_j)$$

to express the compatibility between menu items x_i and y_j according to the following rules of interpretation:

$\theta_{ij} > 0$ item x_i is compatible with item y_j .

$\theta_{ij} < 0$ item x_i is not compatible with item y_j .

$\theta_{ij} = 0$ the items are neither compatible nor incompatible, hence they are utility additive.

It is noteworthy to add that the above definition of the θ_{ij} coefficients can express degrees of compatibilities or the lack of it as well as statistically significant differences from zero.

Example: Consider a set of three entrees, $X = \{x_1, x_2, x_3\}$ and three vegetables, $Y = \{y_1, y_2, y_3\}$. Let $P = ||p_{ij}||$ be the matrix of joint probabilities and $p(x)$ and $p(y)$ the vectors of marginal probabilities over X and Y respectively. Let some arbitrary values of these probabilities be as follows:

$$p(x) = \begin{bmatrix} 1/3 \\ 1/3 \\ 1/3 \end{bmatrix} ; \quad p(y) = \begin{bmatrix} 11/36 \\ 14/36 \\ 11/36 \end{bmatrix} ; \quad P = \begin{bmatrix} 1/18 & 1/6 & 1/9 \\ 5/36 & 1/9 & 1/12 \\ 1/9 & 1/9 & 1/9 \end{bmatrix}$$

where the elements p_{ij} indicate the joint probability in the i -th row and the j -th column.

From these data by formula (1) the matrix of conditional probabilities

$p(y_j|x_i)$ denoted by $P_{y|x}$ can be computed, and yields

$$P_{y|x} = \begin{bmatrix} 1/6 & 1/2 & 1/3 \\ 5/16 & 1/3 & 1/4 \\ 1/3 & 1/3 & 1/3 \end{bmatrix}$$

Then subtracting $p(y)$ from each column of $P_{y|x}$ yields the matrix of the coefficients of compatibility, $||\theta_{ij}||$

$$||\theta_{ij}|| = \begin{bmatrix} -5/36 & 4/36 & 1/36 \\ 4/36 & -2/36 & -2/36 \\ 1/36 & -2/36 & 1/36 \end{bmatrix}$$

It is interesting to note that the coefficient θ_{ij} will take negative and positive values even in the last row of the table where the input data would imply indifference. One reason for this is that $p(x_3y_j) \neq p(x_3)p(y_j)$. Furthermore, just because (x_3y_1) , (x_3y_2) , (x_3y_3) are equiprobable does not mean that (x_3y_j) are indifferent. In fact, if one looks at the choices of x_1 and x_2 combined, then one finds that y_1, y_3 are equally preferred, but y_2 is preferred over both y_1 and y_3 . Now just the fact that y_3 gets equal weight to y_1 and y_2 in combinations with x_3 means that y_3 does not go well with x_3 and this is indicated by the number $-2/36$. The above paradox can be generalized as follows. Suppose $p(y_1), p(y_2), \dots, p(y_J)$ are the marginal probabilities of $y_1, y_2, \dots, y_J$ found from (I-1) entrees. Now suppose the I -th entree is selected, so that $p(y_j|x_I) = 1/J$ for all j uniformly. Also, let $\max p(y_j) = p(y_m) \neq 1/J$. Then this implies

that the m -th vegetable gives a worse combination with x_I . Also, if $\min p(y_j) = p(y_n)$, then this implies that y_n goes best with the I -th entree. Hence we will still get negative and positive values for indicating compatibility of the I -th entree with different vegetables.

The above example also shows that the coefficients of compatibility depend intimately upon the cardinality of the sets of items involved in defining the measure. Therefore, for the purpose of finding coefficients which are invariant under various conditions of application, the cardinality of U should be as large as possible. Since choices satisfying this condition are rarely, if ever, available and hence observable on selective menus, a procedure is outlined in the last section of the paper to allow the construction of the joint probability matrix P from estimates of its submatrices.

III THE ESTIMATION OF THE COEFFICIENTS OF COMPATIBILITY

In order to estimate θ_{ij} , defined in section II, we start with n independent meal selections. If we want to estimate θ_{ij} for a population, then n is a random sample of people belonging to this population. On the other hand, n could be different meal selections by the same individual if one is interested in estimating θ_{ij} for an individual. Let

r_{ij} = number of meal selections (out of n) containing (x_i, y_j) .

r_i = number of meal selections (out of n) containing menu item x_i .

r_j = number of meal selections (out of n) containing menu item y_j .

$i = 1, 2, \dots, I$; $j = 1, 2, \dots, J$.

Then the maximum likelihood estimator of θ_{ij} is given by

$$(2) \quad \hat{\theta}_{ij} = r_{ij}/r_i - r_j/n$$

The estimator $\hat{\theta}_{ij}$ is an unbiased estimator of θ_{ij} , since

$$E(r_j/n) = p_j$$

and

$$\begin{aligned} (3) \quad E(r_{ij}/r_i) &= E \left[\frac{1}{r_i} E(r_{ij}|r_i) \right] \\ &= E \left[\frac{1}{r_i} r_i \frac{p_{ij}}{p_i} \right] \\ &= \frac{p_{ij}}{p_i} \end{aligned}$$

where E denotes expectation with respect to r_i . Derivation of (3) uses the fact that $r_{ij}|r_i$ is a binomial random variable with parameters r_i and p_{ij}/p_i .

Now we develop the asymptotic distribution of $\hat{\theta}_{ij}$ which is required for testing the hypothesis that $\theta_{ij} = 0$. This distribution can also be used in order to find the confidence interval of θ_{ij} .

Theorem

The asymptotic distribution of $\sqrt{n} (\hat{\theta}_{ij} - \theta_{ij})$ is $N(0, \sigma_{ij}^2)$, where

$$\sigma_{ij}^2 = p_{ij}(1 - p_{ij}/p_i + 2p_{ij} - 2p_i)/p_i^2 + p_j(1 - p_j)$$

Proof:

$$\begin{aligned}\sqrt{n}(\hat{\theta}_{ij} - \theta_{ij}) &= \sqrt{n} \left[\frac{r_{ij}}{r_i} - \frac{r_j}{n} - \frac{p_{ij}}{p_i} + p_j \right] \\ &= \sqrt{n} \frac{n}{r_i} \left(\frac{r_{ij}}{n} - \frac{r_i}{n} \frac{p_{ij}}{p_i} \right) - \sqrt{n} \left(\frac{r_j}{n} - p_j \right) \\ &= \alpha_n U_n - V_n\end{aligned}$$

where

$$\begin{aligned}\alpha_n &= \frac{n}{r_i} \\ U_n &= \sqrt{n} \left(\frac{r_{ij}}{n} - \frac{r_i}{n} \frac{p_{ij}}{p_i} \right) \\ V_n &= \sqrt{n} \left(\frac{r_j}{n} - p_j \right)\end{aligned}$$

Since (U_n, V_n) have a limiting distribution and $\alpha_n \longrightarrow \frac{1}{p_i}$, making use of the result in Rao[7,p.319], the asymptotic distribution of $\sqrt{n} (\hat{\theta}_{ij} - \theta_{ij})$ is the same as the asymptotic distribution of T_n where

$$(4) \quad T_n = \frac{1}{p_i} U_n - V_n$$

Now in order to obtain the asymptotic distribution of T_n , we define

$$X_{ijk} = \begin{cases} 1 & \text{if } k\text{-th meal selection contains } (x_i, y_j) \\ 0 & \text{otherwise} \end{cases}$$

$$k = 1, 2, \dots, n$$

Then

$$T_n = \frac{\sqrt{n}}{p_i} \sum_{k=1}^n (X_{ijk} - \frac{p_{ij}}{p_i} X_{i \cdot k} - p_i X_{\cdot jk} + p_i p_j) / n$$

$$= \sqrt{n} \sum_{k=1}^n Y_{ijk} / n ,$$

where

$$X_{i \cdot k} = \sum_{j=1}^J X_{ijk} ,$$

$$X_{\cdot jk} = \sum_{i=1}^I X_{ijk} ,$$

$$Y_{ijk} = (X_{ijk} - \frac{p_{ij}}{p_i} X_{i \cdot k} - p_i X_{\cdot jk} + p_i p_j) / p_i .$$

We need the first two moments of Y_{ijk} in order to obtain the asymptotic distribution of T_n . It is straightforward to show that

$$E(T_n) = 0$$

In order to obtain $V(Y_{ijk})$, we need the following results.

$$E(X_{ijk}) = p_{ij} ,$$

$$V(X_{ijk}) = p_{ij} q_{ij} , \quad \text{where } q_{ij} = 1 - p_{ij} ,$$

$$\text{Cov}(X_{ijk}, X_{\ell jk}) = -p_{ij} p_{\ell j} \quad i \neq \ell ,$$

$$V(X_{\cdot jk}) = p_j q_j , \quad q_j = 1 - p_j ,$$

$$V(X_{i \cdot k}) = p_i q_i ,$$

$$\begin{aligned} \text{Cov}(X_{ijk} X_{\cdot jk}) &= \sum_{\ell=1}^I \text{Cov}(X_{ijk} X_{\ell jk}) \\ &= p_{ij} (1 - p_{ij}) - \sum_{\substack{\ell=1 \\ \ell \neq i}}^I p_{ij} p_{\ell j} \\ &= p_{ij} q_j , \end{aligned}$$

$$\text{Cov}(X_{ijk}, X_{i \cdot k}) = p_{ij} q_i ,$$

$$\begin{aligned} \text{Cov}(X_{i \cdot k}, X_{\cdot jk}) &= E(X_{i \cdot k} X_{\cdot jk}) - E(X_{i \cdot k}) E(X_{\cdot jk}) \\ &= P(X_{ijk}=1) - p_i p_j , \\ &= p_{ij} - p_i p_j . \end{aligned}$$

Making use of the above results in the following equation,

$$\begin{aligned} p_i^2 V(Y_{ijk}) &= V(X_{ijk}) + \frac{p_{ij}^2}{p_i^2} V(X_{i \cdot k}) + p_i^2 V(X_{\cdot jk}) - 2 \frac{p_{ij}}{p_i} \text{Cov}(X_{ijk} X_{i \cdot k}) - \\ &\quad - 2 p_i \text{Cov}(X_{ijk}, X_{\cdot jk}) + 2 p_{ij} \text{Cov}(X_{i \cdot k}, X_{\cdot jk}) , \end{aligned}$$

and simplifying, we have

$$(5) \quad \sigma_{ij}^2 = V(Y_{ijk}) = p_{ij} (1 - p_{ij}/p_i + 2p_{ij} - 2p_i)/p_i^2 + p_j q_j$$

Since Y_{ijk} , $k = 1, 2, \dots, n$, are independent and identically distributed random variables with mean zero and variance σ_{ij}^2 given by (5), by Central Limit Theorem and equation (4) we have

$$T_n \xrightarrow{d} N(0, \sigma_{ij}^2) ,$$

where $\xrightarrow{d}$ denotes convergence in distribution. Consequently, we have proved the result in Theorem that

$$\sqrt{n}(\hat{\theta}_{ij} - \theta_{ij}) \longrightarrow N(0, \sigma_{ij}^2) .$$

Now in order to obtain confidence interval of θ_{ij} , we have the following Lemma which can easily be proved by making use of result in Cramer [4, p254].

Lemma 1:

The asymptotic distribution of $\frac{\sqrt{n}(\hat{\theta}_{ij} - \theta_{ij})}{\hat{\sigma}_{ij}}$ is $N(0,1)$, where $\hat{\sigma}_{ij}$ is an estimator of σ_{ij} obtained on replacing p_{ij} by r_{ij}/n , p_i by r_i/n and p_j by r_j/n in σ_{ij} .

The test statistic

$$z_{ij} = \sqrt{n} \hat{\theta}_{ij} / \hat{\sigma}_{ij}$$

can now be used for testing the null hypothesis

$$H_0 : \theta_{ij} = 0 ,$$

where z_{ij} has a standard normal distribution for large n .

IV OBTAINING JOINT PROBABILITY ESTIMATES FROM SUBSETS OF THE FULL PRODUCT SET

It is obvious that the estimation of the coefficient of compatibility, θ_{ij} , for any pair of menu items is related to the estimation problem of the corresponding joint probability term p_{ij} which in turn is the function of the cardinality of the product set $[X \times Y]$. Ideally, therefore, sets X (entrees) and Y (vegetables) should be large enough to include all the items an individual may routinely encounter in his menu selection activity. The problem here is, of course, that the number of choices which can be offered at any time in a selective menu are much less than the elements in X or Y . Consequently, observations are feasible only on subsets of $[X \times Y]$ and a technique is necessary to project the joint probabilities of the subsets into those of the full set under consideration. We will show first that the reverse process is always feasible, and from that the procedure of correct projection easily follows.

Let matrix $P = ||p_{ij}||$ denote the joint probabilities of selection when sets X (entrees) and Y (vegetables) are of cardinality I and J respectively, and when the cartesian product set $U = (X \times Y)$ represents the full set of pairwise combinations of all the items under consideration. Let S_k be a subset of U , and consider K such subsets where $S_k \subset U$ for $k = 1, 2, \dots, K$ holds. The subset S_k itself is the cartesian product set of item sets X_k and Y_k with cardinality $I_k < I$ and $J_k < J$ respectively. For each subset S_k there exists a corresponding submatrix of the joint probabilities P denoted by P_k .

Now consider the joint probability matrix $P(S_k)$ defined only on the limited choices X_k and Y_k . Clearly, $P_k \neq P(S_k)$, but they are related

by a simple rule of proportionality if one makes the legitimate assumption that the preference ordering of the individual remains unchanged. In this case

$$(6) \quad P(S_k) = P_k/a_k$$

where

$$a_k = \sum_{i=1}^{I_k} \sum_{j=1}^{J_k} p_{ij}$$

and p_{ij} is the element of the known matrix P_k . The constant a_k is nothing else but the sum of the probabilities in matrix P_k which is also equal to the sum of the marginal probabilities for the given submatrix with respect to subsets X_k or Y_k and hence a_k is less than one.

Now it is obvious that we can project P_k into $P(S_k)$, but our problem is to go backward, because in reality only estimates of $P(S_k)$ will be available, and the P_k submatrices, and eventually the matrix P will have to be reconstructed by some process.

First we make the assumption that $U = \bigcup_{k=1}^K S_k$ in such a way that the K subsets are not disjoint, but at least $K-1$ nonempty intersections of the type $S_k \cap S_\ell$ exist. Let $\Omega_{k\ell}$ denote such an intersection set.

Then we have from (6)

$$\sum_{\Omega_{k\ell}} p_{ij}(S_k) = \sum_{\Omega_{k\ell}} p_{ij}/a_k$$

and

$$\sum_{\Omega_{k\ell}} p_{ij}(S_\ell) = \sum_{\Omega_{k\ell}} p_{ij}/a_\ell$$

Consequently

$$a_k/a_\ell = \sum_{\Omega_{k\ell}} p_{ij}(S_\ell) / \sum_{\Omega_{k\ell}} p_{ij}(S_k) = b_{k\ell}$$

or

$$a_k - b_{k\ell} a_\ell = 0$$

where $b_{k\ell}$ is considered known, and a_k and a_ℓ are to be determined.

If we consider the $K-1$ intersections between the K subsets, and make the $k+1 = \ell$ substitution, we obtain a system of $K-1$ equations

$$a_k - c_k a_{k+1} = 0 ; \quad k = 1, 2, \dots, K-1$$

where

$$c_k = b_{k,k+1}$$

Also, we have

$$a_1 + a_2 + \dots + a_K = C$$

This system has a unique solution for any given value of C . In matrix vector form we have

$$\begin{bmatrix} 1 & -c_1 & 0 & \dots & \dots & \dots & 0 \\ 0 & 1 & -c_2 & & & & \vdots \\ \vdots & & \ddots & \ddots & & & \vdots \\ \vdots & & & \ddots & \ddots & & \vdots \\ \vdots & & & & \ddots & \ddots & \vdots \\ 0 & & & & & 1 & -c_{K-1} \\ 1 & 1 & 1 & \dots & \dots & 1 & 1 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ \vdots \\ \vdots \\ a_K \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ \vdots \\ \vdots \\ 0 \\ C \end{bmatrix}$$

which can be stated equivalently by the matrix algebraic symbols

(7) $Ma = d$
 where

a and d are column vector symbols for the variable a_k ; $k = 1, 2, \dots, K$ and for the constant vector at the right hand side. The solution of (7) can be written as

$$a = M^{-1}d$$

which is equivalent to

$$a = CM_K$$

where M_K denotes the K -th (last) column of M^{-1} . Because of the special structure of M the components of M_K , and hence that of a , can be explicitly computed as follows:

$$\begin{aligned} a_1 &= C \cdot c_1 \cdot c_2 \cdot \dots \cdot c_{K-1} / \Delta \\ a_2 &= C \cdot c_2 \cdot c_3 \cdot \dots \cdot c_{K-1} / \Delta \\ &\vdots \\ a_{K-1} &= C \cdot c_{K-1} / \Delta \\ a_K &= C / \Delta \end{aligned}$$

where

$$\Delta = |M| = 1 + c_{K-1} + c_{K-2} \cdot c_{K-1} + \dots + c_1 \cdot c_2 \cdot \dots \cdot c_{K-1}$$

in this case.

Since

$$a_1 + a_2 + \dots + a_K = C$$

a unique solution can be found for any value of C . The correct value of C obtains by normalizing the resulting P matrix elements to satisfy the

$$\sum_{i=1}^I \sum_{j=1}^J p_{ij} = 1$$

conditions for probability measures.

The above derivations open the way for experiments and experimental designs with the objective of obtaining estimates of P from observations available on a sequence of selection statistics from selective menus. It is believed that the procedures outlined are necessary and sufficient for the statistical basis of establishing food compatibility data banks for computerized food service organizations.

REFERENCES

- [1] Armstrong, R.D., J.L. Balintfy, P. Sinha, "A Multistage Scheduling Algorithm for Preference Maximized Selective and Nonselective Menus," ONR Technical Report No. 7, School of Business Administration, University of Massachusetts, July 1973.
- [2] Balintfy, J.L., W.J. Duffy, P. Sinha, "Modeling Food Preferences Over Time," to appear in the Journal of Operations Research.
- [3] Balintfy, J.L. and P. Sinha, "A Nonlinear Programming Approach to Utility Maximized Menu Plans," ONR Technical Report No. 9, School of Business Administration, University of Massachusetts, August 1973.
- [4] Cramer, H., Mathematical Methods of Statistics, Princeton University Press, Princeton, 1945.
- [5] Eindhoven, J. and D.R. Peryam, "Measure of Preferences for Food Combinations," Food Technology, Vol. 13 (1959) pp.379-384.
- [6] Keeney, R.L., "Utility Functions for Multiattributed Consequences," Management Science, Vol. 18 (1972) pp.276-287.
- [7] Rao, C.R., Linear Statistical Inference and its Applications, John Wiley & Sons, New York 1965.