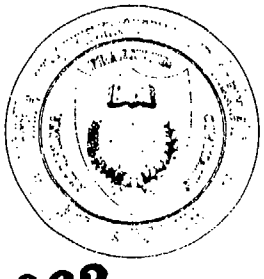
LEVEL II

(12)

CENTER FOR CYBERNETIC STUDIES

The University of Texas
Austin, Texas 78712



Research Report CCS 397

AN MDI MODEL AND AN ALGORITHM FOR COMPOSITE HYPOTHESES TESTING AND ESTIMATION IN MARKETING*

by

A. Charnes
W. W. Cooper
D. B. Learner*
F. Y. Phillips*

Original: July 1981
Revised: September 1981

*Market Research Corporation of America

This Research was partly supported by funds from the Market Research Corporation of America and is partly supported by Project NR 047-021, ONR Contract N00014-81-C-0236 with the Center for Cybernetic Studies, The University of Texas and ONR Contract N00014-81-C-0410 at Carnegie-Mellon University, School of Urban and Public Affairs. Reproduction in whole or part is permitted for any purpose of the United States Government.

CENTER FOR CYBERNETIC STUDIES

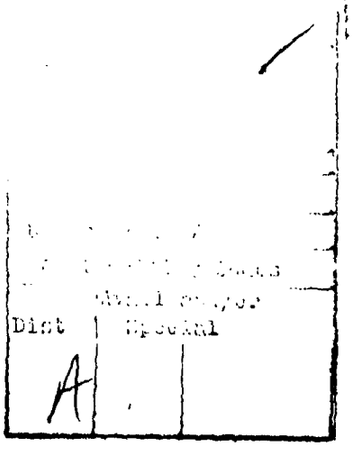
A. Charnes, Director
Business-Economics Building, 203E
The University of Texas at Austin
Austin, TX 78712
(512) 471-1821

ABSTRACT

A strategy is provided for using constrained versions of the MDI (minimum discrimination information) statistic to test and estimate market relations involving composite hypotheses. An algorithm for applying the tests and effecting the estimates is also provided along with numerical illustrations. Other, more general, developments in statistics and mathematical programming (duality) theories and methods are also briefly discussed for their possible bearing on further uses in marketing research and management.

KEY WORDS

Market Segmentation
Market Equilibrium
Switching Equilibrium
MDI Statistic
Composite Hypotheses
Mathematical Programming
External Constraints



1. INTRODUCTION*

This paper centers on a development of constrained information theoretic statistics[1] with accompanying algorithms and illustration for use in marketing. It is pointed toward composite hypothesis testing, as in a market segmentation analysis with explicitly stated constraints. Statistical testing -- in multi-way contingency table analyses, for instance -- is not usually undertaken with explicitly stated arrays of constraints. The recently published book, The Information in Contingency Tables by Gokhale and Kullback[2], however, provides requisite statistical methods and rationales for such treatments. This in turn, opens the way for contact with other parts of the management sciences where complex arrays of constraints are often used to reflect a variety of "policy" and/or "data" conditions.

Our illustrations will be effected from the data published in [20]. Ehrenberg and Goodhardt used these data to show that (a) the Hendry Brand Switching Coefficient approach to market segmentation does not yield very good estimates of brand switching behavior and (b) this market -- like most markets[3] -- is unsegmented (i.e., it consists of a single segment).

---

*The authors are indebted to D.G. Morrison for editorial suggestions that have helped to improve this revision from an earlier draft.

[1] I.e., we are referring to the Kullback-Leibler statistic as found in Kullback [31] rather than the probabilistic or deterministic approaches built around the Shannon-Wiener measure of information as treated in, say, Theil [39] and [40].

[2] See [23].

[3] In a private communication from G. H. Goodhardt dated September 27, 1979, he states that it has been their experience that most markets can be adequately described as being unsegmented.

They do not, however, submit their results to statistical tests.

We shall show how the Ehrenberg-Goodhardt results can be statistically tested relative to other alternatives by reference to their data but we do not propose to enter into yet another discussion of the Hendry approach to market segmentation.[1] Hence, we shall treat their data as simply a numerical illustration for the use of the models and methods we shall supply. Approaching these data in this manner also accords with the fact that we do not know the sample size or conditions of selection. We also do not know the source of the data and, hence, we shall approach these data in a way that can accord with either panel or survey collections of such information from respondents. For similar reasons, we shall steer clear of full-scale segmentation analysis by reference to product benefits, customer (demographic) characteristics, etc., since (a) this accords most closely with the Ehrenberg-Goodhardt results and (b) the requisite data are not available for these purposes anyway. The following development thus provides only a beginning but, as will be seen, it is especially well-suited for statistically testing "nested hypotheses" such as are implicit in market segmentation studies. As we shall also note (and illustrate), however, these constrained information theoretic models and methods possess statistical estimation as well as hypothesis testing properties that may be simultaneously exploited.

---

[1]See, e.g., the critique by Ehrenberg and Goodhardt [21] of the presentation by Kalwani and Morrison [30].

2. BACKGROUND

Marketing professionals are familiar with the use of two-way contingency tables (cross-tabulations). Formally, such two-way analyses can be extended to multi-way contingency tables, but then one confronts a variety of inadequacies in classically available statistical methods. The unsatisfactory nature of our ability to deal with large multi-way contingency tables has begun to give way before many different separate developments in the statistics (and computing sciences) literature. The resulting proliferation of methods has, in turn, given rise to a need for systematization accompanied by a unifying rationale based on methodological as well as conceptual grounds. The recently published book, Discrete Multivariate Analysis, by Bishop, Fienberg and Holland [5] explicitly[1] acknowledges that their effort was undertaken in response to this need. These authors use the "log-linear model" as a basis for "codifying" approaches that might be employed. The approach via log-linear models falls short of what is required, however, to supply the unifying rationale that is needed. This is supplied by Gokhale and Kullback in [23] who use the MDI statistic for this purpose and who show how the log linear model itself can be derived from this statistic[2] along with a variety of other approaches. They also show how extensions beyond any of these other approaches -- e.g., including explicitly adding constraints beyond those of the contingency tables -- can be effected by following along these "information theoretic" lines.

---

[1] pp. 1-3.

[2] pp. 38-39.

The history of the development of the MDI statistic is a rather curious one, at least as far as the American literature of statistics is concerned. It derives from the so-called "Kullback-Leibler statistic" which is of the form

$$(1.1) \qquad I(p:\pi) = \sum_{i=1}^{n} p_i \log \frac{p_i}{\pi_i}$$

where p and π are vectors with components

$$p_i, \ \pi_i \geq 0$$

$$(1.2) \qquad \sum_{i=1}^{n} p_i = \sum_{i=1}^{n} \pi_i = 1 .$$

Here the π_i may represent a set of hypothesized (constant) values which are to be tested relative to the p_i (variable) values estimated from observed data. The π_i may also represent a set of prior probabilities as in Bayesian decision theory and then the p_i become posterior probabilities determined from sample observations. Other interpretations are also possible which accord with different ways of chosing the p_i. When the components of p represent minimizing values, p*, the expression in (1.1) may be replaced by

$$(2) \qquad I(p^*:\pi) = \sum_{i=1}^{n} p_i^* \log \frac{p_i^*}{\pi_i} = \min_{p} \sum_{i=1}^{n} p_i \log \frac{p_i}{\pi_i} ,$$

and this is called the MDI statistic.

The expression in (1.1) is often confused with the Shannon-Wiener measure of information. The latter, however, represents a probabilistic rather than a statistical concept. In fact, as Kullback notes in his book, Information Theory and Statistics[1], L. J. Savage, one of the founders of modern decision theory, thought of it only in these terms when he remarked, "The ideas of Shannon and Wiener, though concerned with probability, seem rather far from statistics." Kullback, however, goes on to show how both (1.1) and (2) can be given a statistical characterization[2] and used to unify an extremely wide variety of statistical concepts and developments.

Kullback's contributions are appropriately acknowledged in the literature which refers to (1.1) as the "Kullback-Leibler statistic."[3] He is also the author of the name Miniumum Discrimination Information (MDI) statistic which name may be explained in the following way: Consider any estimate of p in (1.1) which yields a distribution that significantly differs from the distribution as hypothesized in π. If the p distribution differs significantly from the hypothesized π distribution then a question may remain as to whether some other estimate of p (also consistent with the data) might fail to exhibit significance. The matter is resolved, however, if p is chosen "as close as possible" to π. This, then, is the meaning of minimum discrimination, i.e., p* provides minimum discrimination against the hypothesized π that the data together with any other constraints will allow.

---

[1] [31] p.2.

[2] Also called the "Khinchin-Kullback-Leibler statistic" because of earlier (but simpler) contributions of A. I. Khinchin in his book on the Mathematical Foundations of Statistical Mechanics (New York: Dover Publications, 1949). See [17].

[3] See [25].

The additional constraints allowed in Gokhale-Kullback [23] are of the form

$$(3) \qquad \sum_j a_{ij} p_j = \theta_i, \quad i = 1, 2, \ldots, m .$$

When the θ_i are derived from the data, as in the case of marginal totals for a contingency table, then the constraints are said to be "internal constraints". When they are imposed on the basis of various hypothesized premises, as in, for instance, an assumed segmentation, then the constraints are said to be "external constraints." It is the latter which will be emphasized here.

Notice the flexibility that is allowed by reference to the possibility of testing hypotheses with the π values in the functional or the θ values in the constraints. Choice of the minimizing p* in terms of (1.1), (2) and (3) also suggests the possibility of contact with mathematical programming with its great computational power and the variety of interpretations that are available from the underlying duality relations. These prospects, too, have now been formally comprehended as in [10][1] with the result that an unusually simple (unconstrained) convex programming problem is found to be the dual to minimizing (1.1) subject to (1.2) and (3)[2], which dual is

$$(4) \qquad \operatorname*{Max}_z \sum_{i=0}^{n} \theta_i z_i - \sum_{j=1}^{n} \pi_j \exp\left(\sum_i z_i a_{ij}\right)$$

---

[1] See also [17].

[2] It is perhaps of interest to note that (1.1) provides a "proper goal functional" in the sense of [11] for use in goal programming.

where $\theta_o = 1$, $a_{oj} = 1$, $j = 1, \ldots n$ and "exp" denotes the exponential, $e = 2.718....$ See [10] and [17], and observe that the choices of the components in z are not constrained.

One way to approach the need for the kind of general assessment we are making is to say that the literature of marketing research reflects the fact that the American-English statistics literature has remained relatively quiescent on the subject at least as far as applications are concerned since the appearance of Kullback's book. This has not, however, been true of other statistics literatures. Both Soviet and Japanese statisticians have continued to push forward vigorously along the paths opened by Kullback in his statistical characterizations[1]. The Japanese statistician, H. Akaike, for example, presented a very important paper at a Soviet sponsored conference[2] in which he showed how maximum likelihood estimation (the heart of classical statistics) could be given a formulation asymptotically equivalent to MDI. The MDI method has the unique characteristic of providing both hypothesis testing and statistical estimation regardless of the conclusion of the test. It thus provides a decision theoretic method which unifies hypothesis testing and estimation.

Regression (sometimes of a logit or probit variety) and factor analysis are often used techniques in marketing research. Current statistical procedures usually determine the number of terms to be used on a trial and error -- or exhaustive sequential -- basis. Akaike shows in [1],

---

[1]The economist, H. Theil, has also pushed forward vigorously, but largely along lines that are either deterministic or probabilistic in character, to show how an extremely wide variety of social and economic problems might be addressed. See [39] and [40].

[2]See [1].

however, that the MDI approach immediately determines the number of terms to be used from the sample data by the MDI decision theoretic criterion alone.[1]

There are other advantages that can also be secured. For example, as we have elsewhere shown [16], MDI can provide a basis for unifying a great variety of seemingly separate (and different) approaches to marketing research. Here we want to show how to bring it to bear on problems such as the composite hypothesis testing required for market segmentation testing.

3. A SEGMENTATION THEOREM AND ALGORITHM

Turning to the marketing literature, only a few example applications are available and these generally[2] take the form of the simpler "entropy" formula

$$(5) \qquad \sum_{i=1}^{n} p_i \log p_i$$

applied to areas like brand switching or market segmentation.

Examples of uses of the entropy concept in the marketing literature may be found in the articles by Herniter [27], [28] and by Bass [4][3].

[1] Akaike also brought these methods to bear on the so-called "James-Stein paradox" wherein by the use of seemingly irrelevant data one can secure improved estimates which are not only more efficient than the mean (a maximum likelihood estimator) but which also even eliminate the mean from the admissible class of estimators in a decision theoretic sense. See [2] and [3] where Akaike also discusses the inadequacies of Bayesian approaches to this topic. (An amusing and insightful article on the James-Stein paradox may be found in [18] and a more general treatment of the deficiencies of maximum likelihood estimators may found in [41].)

[2] Theil [39] and [40] provides numerous examples in which expressions (1.1) and (2) are used but, as previously noted, these are generally accorded deterministic or probabilistic interpretations and treatments.

[3] See, also Carter [9].

Variants and alternatives to it may be found in the article by Kalwani and Morrison [30] as well as in other articles dealing with the Hendry approach to brand switching and market segmentation analysis. See, e.g., [8] and [26].

We elect to make contact with this part of the marketing literature, but without entering into a full-scale discussion of the Hendry approach to segmentation. An easy way to do this is via the data of Table 1 which Ehrenberg and Goodhardt [19] used to test the Hendry approach and concluded that, on the evidence, this market does not exhibit any segmentation at all.

In arriving at this conclusion, Ehrenberg and Goodhardt use only a first-order analysis of the market shares. Within a Hendry type analysis one must check to insure that the switching constants derived from these ratios "is applicable to the total product class as well as to the individual brands within a product class."[1] More generally one might commence with the (relative) market share shown for each product, and then go on to consider pairs, followed by triplets, and so on, to the $2^n-1=127$ possibilities for the example shown in Table 1. The idea is to check to see whether the resulting ratios are approximately the same in order to avoid (or reduce) the danger of reaching erroneous conclusions from the (perhaps accidental) equality of lower order ratios.

Actually Ehrenberg and Goodhardt use only a first order analysis. That is, they do not undertake any of these higher order calculations, but their approach is nonetheless satisfactory by virtue of the following theorem:

---

[1]Quoted from Kalwani and Morrison [30], p. 420. See also p. 473.

Theorem: If $\frac{A_1}{B_1} = \frac{A_2}{B_2} = \ldots = \frac{A_n}{B_n}$

Then $\frac{A_k}{B_k} = \frac{\sum_{s \in S} A_s}{\sum_{s \in S} B_s}$ for all $k \in \{1, \ldots, n\}$ and all $S \subset \{1, \ldots, n\}$,

where A_j, B_k are real numbers with all $B_k \neq 0$. In other words, when the first order ratios are equal then all the higher order ratios, however they may be formed, will also be equal.

This theorem, which we have proved elsewhere[1], constitutes the first part of our proposed algorithm and enables us to follow Ehrenberg and Goodhart in restricting ourselves to the simple situtation of only first order calculations in the use of formula (6), below. We do not follow Ehrenberg and Goodhardt into an examination of the "entropy function" variants which enter into the Hendry approaches for computing such ratios[2]. We simply observe that the above theorem holds without reference to the way the A_s and B_s values are obtained and then pass on to an initial segmentation via

$$R_i = \frac{p_i - p(i,i)}{p_i\,(1 - p_i)} \qquad (6)$$

where p_i = market share for brand i

$p(i,i)$ = proportion making repeat purchases of brand i.

---

[1] See [14].

[2] The use of these entropy functions in Hendry analyses are discussed in Kalwani and Morrison [30], who apparently believe they can be dispensed with, and also by Ehrenberg and Goodhardt [21] who believe that they are a distinguishing feature of Hendry type analyses.

Before applying this formula to the data of Table 1, below, we reiterate that we are not aware of the way the data in Table 1 were obtained. Neither all rows nor all columns sum to unity -- see Table 2 -- so that evidently the matrix is not Markovian. We could, of course, induce this property by dividing through by the row or column rim values. Ehrenberg and Goodhardt do not use the table in this "Markovian" fashion in [20], however, and the fact that Ehrenberg has been openly critical of the use of first-order Markov processes in describing consumer behavior[1] disinclines us to this approach. We shall therefore interpret the p(i,i) as joint rather than conditional probabilities.

Applying (6) to the data of Table 1 we obtain the following results

$$
\begin{array}{lll}
 & R_1 = 0.660 & R_4 = 0.719 \\
(7) & R_2 = 0.656 & R_5 = 0.750 \\
 & R_3 = 0.723 & R_6 = 0.704
\end{array}
$$

Using the above theorem we then assert that i and j are in the same classification[2] when $R_i \doteq R_j$. Following Ehrenberg and Goodhardt [20], we have used the average of the row and column marginal values in arriving at these results. This assumes "market share equilibrium"[3] between the

---

[1] See [19] including his rejoinder to the Massey and Morrison critique noted there.

[2] The approximation indicated by "$\doteq$" is made clear in (7) and (8) and subsequently subjected to statistical tests. A general (more precise) statement of this approach is formalized as the SSI theorem in [14].

[3] See [20].

two periods, a hypothesis that is subject to test along lines that we shall indicate. This, however, is not the only possible equilibrium that might be of interest for marketing purposes. For instance, we shall introduce the idea of "switching equilibrium" and show how this, too, may be tested.

Ehrenberg and Goodhardt conclude from their analyses in [20] that this market does not segment into different classes. I.e., they conclude that there is only one "homogeneous" market in which all products compete.[1] They do not test this hypothesis by statistical methods, however, as we shall do. This will be done as a byproduct of our use of a series of composite hypothesis testing procedures with an accompanying algorithm starting with the following three different classes (or segments) which our rule of approximate equality ($R_i \doteq R_j$) provisionally suggests.

$$(8) \qquad \begin{aligned} I_1 &= \{1,2\} \\ I_2 &= \{3,4,6\} \\ I_3 &= \{5\} \end{aligned}$$

TABLE 1

The Observed Brand-Switching Percentages

for Six Brands of Breakfast Cereals

(Two successive purchases per consumer)

Product	First Purchase	Second Purchase 1	2	3	4	5	6	All Buyers
1	Corn Flakes	(23.8)	7.7	3.3	2.0	1.3	1.4	39.5
2	Weetabix	6.4	(14.3)	3.3	1.1	.8	1.1	27.0
3	Shredded Wheat	3.6	3.2	(5.4)	.8	.8	.7	14.5
4	Sugar Puffs	3.2	1.5	1.1	(3.1)	.8	.3	10.0
5	Puffed Wheat	1.7	.6	.4	.6	(1.5)	.1	4.9
6	Brand X	1.0	.4	.6	.3	.3	(1.5)	4.1
	All Buyers	39.7	27.7	14.1	7.9	5.5	5.1	100.0

[1] See Ehrenberg and Goodhart [20].

4. VULNERABILITY RATIOS

These R_i values are sometimes referred to as "switching constants" -- e.g., in a Hendry type analysis. We prefer the term "vulnerability ratios", however, so that the above development produces classes of equally vulnerable products.

Justification for this change in terminology my be presented in terms of (6) by rewriting it in the following "verbalized" form:

$$R_i = \frac{1 - \frac{\text{repeat purchase rate for } i}{\text{brand share for } i}}{1 - \text{brand share for } i} \tag{9}$$

It is, as can be seen, a consumer outflow measure for brand i.[1] Without allowing for inflows it becomes a measure of brand i's vulnerability to such outflows. Hence, our characterization of it as a "vulnerability ratio."

The smaller the value of this ratio, the more stable (or invulnerable) is the brand in the sense that smaller R_i reflect an increasing proportion of repeat buyers in the brand's present market share. For instance, if $R_i = 0$, then the brand share for i = the repeat purchase rate for i, and all of this brand's customers are repeat buyers.

---

[1] This would be true even if the p(i,i) were interpreted as conditional probabilities, as in a Markoff type approach, rather than the joint probability interpretation which we are using.

We are keeping these interpretations general and restricting our terminology accordingly. For instance, we do not assess the reasons for the observed behavior and we allow for a variety of assumptions (e.g., of a probabilistic character) which may accord with more specialized statistical distributions used by others. As a case in point we may cite Sabavala and Morrison [38] who use a Beta-Binomial development to obtain what they refer to as a "Loyalty Index" for viewing TV programs.[1] In this case more special (and precise) characterizations are possible since the condition $R_i = 0$ then also implies a very bipolar population in which consumers all have p values near 0 or 1 for any such brand.

Results like the above yield additional hypotheses for testing -- i.e., situations particularized further below the general situation $p_i = p_{ii}$ which applies in our case -- which, of course, is consistent not only with high or low p_i values but with intermediate values as well.

The use of specializing assumptions may also restrict the range of possible hypotheses on a priori grounds. For instance, an assumed Beta-Binomial will also restrict the range of R_i values to $0 \leq R_i \leq 1$ whereas our more general situation admits of values $R_i > 1$. Within the range of such restrictions, then, we may regard the Sabavala-Morrison "Loyalty Index" as the complement of our "Vulnerability Ratio".

---

[1]We are indebted to D.G. Morrison for pointing out the additional possibilities and relations derivable from this article.

In general, we may think of brands with $0 \leq R_i < 1$ as being relatively invulnerable. At $R_i = 1$ we have

$$p(i,i) = p_i^2 \quad . \tag{10}$$

If we interpret these ratios in terms of probabilities, then we can say that at $R_i = 1$ the repeat purchases are statistically independent identically distributed events. Finally, at values $R_i > 1$ we would have $p_i^2 > p(i,i)$ in which case repeat buying would be even worse than random and the brand highly vulnerable.[1]

This kind of information can be put to use in developing market strategies. For instance, it directs attention to the possibilities of attracting customers from brands with R_i values exceeding one's own. Conversely, it warns of possible inroads from brands with smaller R_i values, which inroads may then have some staying power that increases, e.g., because of repeat purchasing propensities.

Differences in observed R_i values may not be statistically significant, in which case the above interpretations will need to be modified to accord with this fact. Testing for statistical significance of these and other hypotheses is a task to which we shall now turn, but our focus on this aspect of testing does not mean that we believe that other considerations such as, e.g., application of marketing insight, are unimportant.

5. TYPICAL FORMS

We will use specializations of (1.1) and (2) and (3) to illustrate our procedures for testing the structure of a market in the form of a series of "nested hypotheses" starting with (8). We now introduce a double subscript for the discrete probability distributions with which we are concerned. The model is[2]

---

[1] One would generally expect such situations to be rare and fairly short-lived -- although one could, clearly, arrange to have a product and accompanying market strategies which would produce such behavior almost at will.

[2] The well-known SUMT program of Fiacco and McCormick [22] is available for solving this kind of problem via its dual.

$$\text{Minimize} \quad I(p_{ij}:\pi_{ij}) \equiv \sum_{i,j} p_{ij} \ln \frac{p_{ij}}{\pi_{ij}}$$

subject to

(11)
$$1 = \sum_{i,j} p_{ij}, \quad p_{ij} \geq 0, \ \forall\, i,j,$$

$$0 = \hat{R}_i - \hat{R}_j, \ \text{etc.},$$

where

$$\hat{R}_i = \frac{1 - p(i,i)/\bar{p}_i}{1 - \bar{p}_i}$$

(12)
$$\hat{R}_j = \frac{1 - p(j,j)/\bar{p}_j}{1 - \bar{p}_j}$$

and the $\bar{p}_1$ and $\bar{p}_j$ indicate posited market share values, i.e., these values are <u>not</u> to be estimated by this minimization procedure. Here $p(i,i) \equiv p_{ii}$ and the p_{ij} represent proportions switching between products i and j over the two purchase occasions.

The estimated $\hat{R}_i = \hat{R}_j$ represent vulnerability ratios hypothesized to accord with "nestings" indicated by arrangements like

(13)
$$0 = \hat{R}_i - \hat{R}_j$$
$$0 = \hat{R}_j - \hat{R}_k$$

when products i, j and k are hypothesized to be in the same vulnerability class.

To simplify notation we omit the circumflexes on the R_i, R_j, R_k and designate <u>optimal</u> estimates by

(14)
$$R_i^* = \frac{1 - p_{ii}^*/\bar{p}_i}{1 - \bar{p}_i}$$

$$R_j^* = \frac{1 - p_{jj}^*/\bar{p}_j}{1 - \bar{p}_j}$$

The resulting

$$I^* = I\,(p_{ij}^*\,,\ \pi_{ij}) \tag{15}$$

may then be used to test the hypothesis that the minimizing p_{ij}^* choices do not deviate from the distribution associated with the π_{ij}. The π_{ij} represent "switching proportions" hypothesized in our case to accord with the data represented in Table 2,

TABLE 2

i \ j	1	2	3	4	5	6
1	.238	.077	.033	.020	.013	.014
2	.064	.143	.033	.011	.008	.011
3	.036	.032	.054	.008	.008	.007
4	.003	.015	.011	.031	.008	.003
5	.017	.006	.004	.006	.015	.001
6	.010	.004	.006	.003	.003	.015

where the π_{ij} data of Table 2 are the Table 1 data expressed as fractions. The p_{ij}^* values are to be secured by solving (11) using <u>all</u> of the data (off-diagonal as well as diagonal) in Table 2.

A characterization which provides access to the statistical properties as developed in Kullback [31] for the MDI statistic may be obtained from the often made remark that it has been widely observed (i.e., in many markets) that brand switching is proportional to brand share[1] and,

[1]See Ehrenberg and Goodhardt [20]

with R_i^* regarded as a switching constant, this is reflected in (6). However, this "empirical law" may itself be regarded as being implied by the assumption that the underlying distribution which yields these switches is multinomial in character with probabilities p_i, $i = 1,2,\ldots,n$. It then follows from the multivariate form of the central limit theorem that the observations tend to the multi-normal distribution with increasing sample size.[1] Thus, from these considerations alone, one may anticipate such an empirical law directly from the fact that the dispersion (covariance) matrix for large samples will be symmetric with off-diagonal terms proportional to the product of the corresponding brand shares. Conversely, a finding of statistically significant deviations will result in rejection of the "empirical law" that switching is proportional to brand share. Rejection of this hypothesis can serve, in marketing terms, as an alert to management that the market structure is changing.

6. ALGORITHM AND TESTING PROCEDURE

The theorem that we provided in Section 3 allowed us to avoid one part of the combinatorial number of possibilities that might be encountered in a market segmentation analysis. Another part of these combinatorial difficulties is encountered when we come to effecting our stastical tests since one might (in principle) have to consider a great number of possible $\hat{R}_i$, $\hat{R}_j$ pairs to test in (11).

---

[1] For a more detailed discussion with accompanying further references, see [14].

We can, of course, restrict our tests to the already suggested possibilities in (8) when we are confident that this will not omit other possibilities that might also be valid. More generally, however, and even in the present case, a greater variety of possibilities may need to be considered. We therefore suggest the following as a second part to the algorithm which we have already provided via the theorem in Section 3.

This second part of the algorithm, which is to be used when statistical testing is to occur, may be stated as follows:

> Order the R_i values from smallest to greatest as $R_{(1)}, R_{(2)}, \ldots, R_{(n)}$, as in Table 3 below. See (7). By MDI, test the hypothesis that $R_{(1)}$ and $R_{(2)}$ are equal at a specified level of significance. If the resulting value of the MDI statistic rejects the hypothesized equality then segment $R_{(1)}$ from $R_{(2)}$. Next begin with $R_{(2)}$ as a smallest R_i to test for a further segment apart from $R_{(3)}$. If, instead, the test accepts the hypothesized equality of $R_{(1)}$ and $R_{(2)}$, add the additional condition to $R_{(1)} = R_{(2)} = R_{(3)}$. Test with the MDI statistic. If the test accepts the hypothesized equality of $R_{(2)}$and $R_{(3)}$ go on to new higher R_i by adding the new pertinent additional equality conditions.

To illustrate our suggested procedures for using these MDI values, we now refer to the following R_i values from the results portrayed in (7), but arranged as indicated by this part of our algorithm.

TABLE 3

i:	2	1	6	4	3	5
R_i:	.650	.660	.704	.719	.723	.750

Applying this algorithm we test the hypotheses indicated by the nestings (from top to bottom) under the column labelled "Null Hypothesis" in Table 4. With this formulation, Ehrenberg and Goodhardt apparently are correct, or at least their "one homogeneous market" conclusion cannot be rejected on the basis of these data for all sample sizes, $N \leq 6{,}500$ and significance level $\alpha = 0.95$.[1]

TABLE 4

Problem	Null Hypothesis	I*	d.f.#	$N \leq 6500$ $\alpha = 0.95$
1	$R_2 = R_1$	Negligible	1	accept
2	$R_2 = R_1 = R_6$	1.24×10^{-4}	2	accept
3	$R_2 = R_1 = R_6 = R_4$	3.60×10^{-4}	3	accept
4	$R_2 = R_1 = R_6 = R_4 = R_3$	5.94×10^{-4}	4	accept
5	$R_2 = R_1 = R_6 = R_4 = R_3 = R_5$	8.54×10^{-4}	5	accept

See Gokhale and Kullback [23] or [24].

There is, however, another more stringent test of market homogeneity based on requiring the estimated p_{ij} to conform to the conditions for market share equilibrium which requires the adjunction of additional constraints. There are also other kinds of market equilibria that might need to be considered.[2] To create a better perspective on some of the

[1] Ehrenberg and Goodhardt did not supply information on sample sizes or other statistical characterizations.

[2] We leave aside the more expedient computational and interpretational possibilities afforded from (4) in order to focus on these more immediate issues.

the latter possibilities we shall therefore first consider how a "switching equilibrium" model might be formulated for joint hypothesis testing and estimation via

$$\min I(p:\pi)$$

subject to

$$\text{(16)} \qquad \begin{aligned} & p_{ij} = p_{ji}, \quad \forall\, i \neq j \\ & \sum_{i=1}^{6} \sum_{j=1}^{6} p_{ij} = 1 \\ & p_{ij} \geq 0 \quad \forall\, i,j \end{aligned} \bullet$$

The resulting estimates are portrayed in Table 5 with $I^* = 0.007$ and 15 degrees of freedom. The hypothesis of "switching equilibrium" is <u>not</u> rejected for all $N \leq 1430$ at $\alpha = 0.95$.

TABLE 5

p_{ij}^* Switching Equilibrium Values, $i \neq j$

i \ j	1	2	3	4	5	6
1	.240	.071	.035	.025	.015	.012
2	.071	.144	.033	.013	.007	.007
3	.035	.033	.054	.009	.006	.007
4	.025	.013	.009	.031	.007	.003
5	.015	.007	.006	.007	.015	.002
6	.012	.007	.007	.003	.002	.015

Apparently the market portrayed in Table 1 is consistent with the hypothesized switching equilibrium and the p_{ij}^{*} values portrayed in Table 5 represent "best" estimates of the switching probabilities in the sense that they are "as close as it is possible to get" to the π_{ij} values in Table 2 while retaining the symmetry conditions for switching equilibrium as reflected in (16).

Turning next to the question of market-share equilibrium we now formulate our model for purposes of testing this model by incorporating these conditions explicitly as follows:[1]

$$\min \; I(p:\pi)$$

subject to

$$\begin{aligned} &\sum_{j=1}^{6} p_{ij} = \bar{p}_i \\ &\sum_{i=1}^{6} p_{ij} = \bar{p}_j \\ &\hat{R}_i = \hat{R}_j \\ &\vdots \\ &\hat{R}_k = \hat{R}_\ell \\ &p_{ij} \geq 0, \qquad \forall\, i,j \end{aligned} \tag{17}$$

where $\bar{p}_i$ and $\bar{p}_j$ represent the average of the rim values from Table 2 with $\bar{p}_i = \bar{p}_j$ whenever $i=j$ and $\sum_{i=1}^{6} \bar{p}_i = \sum_{j-1}^{6} \bar{p}_j = 1$. In other words, we formalize the assumption of an hypothesized market share equilibrium and then repeat the

---

[1]Compare with (11) and (12). Although Ehrenberg and Goodhardt are clear in their discussion, no such tests were conducted by them -- possibly because classical statistical mechanisms are not designed to deal with such explicitly constrained models involving "external constraints" in rather complex arrays.

same tests as in (11) to ascertain whether the conclusion of a single homogeneous market is correct and obtain the results shown in Table 6.

TABLE 6

MARKET SEGMENTATION WITH
MARKET SHARE EQUILIBRIUM

Problem	Null Hypothesis	I*	d.f.	N=3750 α=0.95
1	$R_1=R_2$	2.78×10^{-3}	12	accept
2	$R_1=R_2=R_6$	3.02×10^{-3}	13	reject
3	$R_1=R_2$; $R_6=R_4$	2.88×10^{-3}	13	accept
4	$R_1=R_2$; $R_6=R_4=R_3$	2.89×10^{-3}	14	accept
5	$R_1=R_2$; $R_6=R_4=R_3=R_5$	2.99×10^{-3}	15	accept

For N=3,750 the hypothesis of market homogeneity would be rejected since at this sample size (and above) the market segments into two. The hypothesized singleton classification in (8) is also rejected in favor of joining this classification together with one in I_2. However, for N < 3,750 this classification is rejected and the Ehrenberg-Goodhardt conclusion [13] of a single homogeneous market continues to be maintained as before.

In conclusion, we also provide the p_{ij}^* estimates corresponding to the hypothesis of problem 5 in Table 6. These are presented in Table 7. Note that these values are consistent with the indicated segmentation in N = 3,750. At this sample size, the results in Tables 4 and 6 are consistent and at $N \leq 1,430$ they are also consistent with the concept of switching equilibrium. At sample sizes larger than N = 1,430, however, the hypothesis of switching equilibrium is rejected while the hypothesis of market-share equilibrium

continues to be maintained for $1,430 \leq N \leq 3,750$. Evidently, the two concepts are not the same. Thus, even with an hypothesized and validated market-share equilibrium the vulnerability ratio continues to provide valuable information by the way it summarizes the net switches in and out and by the way it designates the more vulnerable products by reference to repeat buying relative to market share.[1]

TABLE 7

ESTIMATES OF SWITCHING/REPEAT PROBABILITIES IN A TWO-SEGMENT MARKET UNDER MARKET-SHARE EQUILIBRIUM

i \ j	1	2	3	4	5	6
1	.239	.076	.033	.024	.012	.012
2	.065	.143	.034	.014	.007	.010
3	.035	.031	.054	.010	.007	.006
4	.028	.013	.010	.031	.006	.002
5	.017	.006	.004	.007	.016	.001
6	.012	.005	.007	.004	.003	.014

[1]See the discussion in Section 4.

CONCLUDING REMARKS

A great variety of other possibilities are also present, of course, but the above examples, with accompanying algorithms and theorems, at least provide a start for the use of MDI methods in marketing where external as well as internal constraints (e.g., those of the ordinary "rim condition" varieties) are to be considered. In the opening sections of this paper, we noted how statistically formulated versions of information theory have been able to supply a basis for unifying a great body of statistical methodology (and concepts). As should be clear from our examples, the use of classical statistical methods on testing and estimation problems such as we have been considering would be onerous and difficult at best. The MDI approaches handle them easily and also open a variety of additional possibilities. This includes possibilities for further testing of sharpened hypotheses when one is willing to make more specialized assumptions about the statistical distributions and/or (as in Sabavala and Morrison [38]) the underlying probability models governing the behavior of individual consumers. It also opens the possibility of contacts with mathematical programming (and its duality relations) such as are provided in [17]. In addition, these models with their explicity formulated constraints can provide a basis for sharpening managerial insight into marketing planning and control. The ultimate managerial benefit is that approaches along these lines can serve to direct management to activities that ought to be undertaken and this, after all, is what market research is (or should be) about if it is to be anything more than a series of academic exercises.

REFERENCES

[1] Akaike, H., "Information Theory and an Extension of the Maximum Likelikhood Principle", 2nd International Symposium on Information Theory, (B. N. Petrov and F. Csaki, eds.) pp. 267-281, Akademiai Kiaido, Budapest (1973)

[2] _____, "An Extension of the Method of Maximum Likelihood and the Stein's Problem", Ann. Inst. Statist. Math. 29, Part A, 153-164 (1977).

[3] _____, "A New Look at the Bayes Procedure", Biometrika 65, No. 1, pp. 53-59, (1978).

[4] Bass, F. M., "The Theory of Stochastic Brand Preference and Brand Switching", Journal of Marketing Research 10, pp. 361-375.

[5] Bishop, Y. M. M., S. Fienberg, and P. W. Holland, Discrete Multivariate Analysis, Cambridge, MA: MIT Press, (1975).

[6] Blattberg, R. C. and S. K. Sen, "Market Segmentation and Stochastic Brand Choice Models", Journal of Marketing Research 13, Feb. 1976, pp. 34-45.

[7] Brockett, P. L., A. Charnes and W. W. Cooper, "M.D.I. Estimatation via Unconstrained Convex Programming", Research Report CCS 326 (Austin, TX: University of Texas, Center for Cybernetic Studies, 1978).

[8] Butler, D. and B. Butler, Jr. Hendro-Dynamics: Fundamental Laws of Consumer Dynamics, (Cambridge, MA: The Hendry Corporation, Chapter 1, 1970, and Chapter 2, 1971).

[9] Carter, J. C., An Entropic-Based Partitioning Approach to the Analysis of Competition, Ph.D. Thesis, Columbia University, 1975 (Ann Arbor, MI: University Microfilms International, 1979).

[10] Charnes, A. and W. W. Cooper, "Constrained Kullback-Leibler Estimation, Generalized Cobb-Douglas Banlance and Unconstrained Convex Programming", Rendiconti di Academia Nazionale dei Lincei, Series VIII, Vol. 56, April, 1974.

[11] _____ and _____, "Goal Programming and Multiple Objective Optimizations; Part I", European Journal of Operational Research 1, No. 1, Jan. 1977, pp. 39-54.

[12] _____ and _____, Management Models and Industrial Applications of Linear Programming, (New York: John Wiley & Sons, Inc., 1961).

[13] _____ and _____, and D. B. Learner, "Constrained Information Theoretic Characterizations in Consumer Purchase Behavior", J. Opl. Res. Soc., No. 9, 1978, pp. 832-842.

[14] _____, _____, _____, and F. Y. Phillips, "A Theorem and Approach to Market Segmentation", in R. P. Leone, ed., Proceedings -- Market Measurement and Analysis, published by TIMS College of Marketing and TIMS, 1980.

[15] _____, _____, _____, _____, "An MDI Procedure for Vulnerability Segmentation Tests", (Austin, TX: The University of Texas, Center for Cybernectic Studies, Research Report CCS 376, Sept. 1980).

[16] _____, _____, _____, _____, "The MDI Method as a Generalization of Logit, Probit and Hendry Models in Marketing", Market Research Corporation of America PDA-RP#70, 1980.

[17] _____, _____, and L. Seiford, "Extremal Principles and Optimization Dualities for Khinchin-Kullback-Leibler Estimation", Math. Operationsforsch. Statist., Ser. Optimization, Vol. 9, No. 1 (1978).

[18] Efrow, B., and C. Morris, "Stein's Paradox in Statistics", Scientific American, 236, No. 5, May 1977, pp. 119- 127.

[19] Ehrenberg, A.S.C., "An Appraisal of Markov Brand-Switching Models", Journal of Marketing Research 2, 1965, pp. 347-373. See also A.S.C. Ehrenberg, "On Clarifying M and M", Journal of Marketing Research 5, 1968, pp. 228-229. See also [34], below.

[20] _____, and G. J. Goodhardt, "The Hendry Brand Switching Coefficient", ADMAP, August 1974, pp. 232-238. Also available from London, ASKE Research Ltd.

[21] _____, _____, "The Switching Constant", Management Science 25, No. 7, July 1979, pp. 703-705.

[22] Fiacco, A. J. and G. P. McCormick, Nonlinear Programming: Sequential Unconstrained Minimization Techniques (New York: John Wiley & Sons, Inc. 1968).

[23] Gokhale, D. V., and S. Kullback, The Information in Contingency Tables, (New York: Marcel Dekker, New York, 1978).

[24] _____, _____, "The Minimum Discrimination Information Approach in Analyzing Categorical Data", Commun. Statist. - Theor. Meth., A7(10), 987-1005, (1978).

[25] Haynes, K. E., F. Y. Phillips and J. W. Mohrfeld, "The Entropies: Some Roots of Ambiguity", Socio-Economic Planning Sciences, May (1980).

[26] Hendry Corporation, Hendrodynamics: Fundamental Laws of Consumer Dynamics, Chapter 1 (1970), Chapter 2 (1971).

[27] Herniter, J., "An Entropy Model of Brand Purchase Behavior", JMR, Vol. X, Nov. 1973.

[28] _____, "A Comparison of the Entropy Model and the Hendry Model", JMR, Vol. XI, Feb. (1974).

[29] Kalwani, M. U., "The Entropy Concept and the Hendry Partitioning Approach", (Cambridge, MA: Sloan School of Management, MIT, WP#1072-79, June 1979).

[30] _____, and D. G. Morrison, "A Parsimonious Description of the Hendry System", Management Science 23, No. 5, Jan. (1977), pp. 467-477.

[31] Kullback, S., Information Theory and Statistics, John Wiley, New York (1959).

[32] _____, and R. A. Leibler, "On Information and Sufficiency", Ann. Math. Statist. 22, pp. 79-86 (1951).

[33] Learner, D. B., and F. Y. Phillips, "Controllable Forecasting in Marketing", Presented at ORSA/TIMs Conference on Market Measurement and Analysis, Stanford CA, March (1979).

[34] Massy, W.F. and D.G. Morrison, "Comments on Ehrenberg's Appraisal of Brand-Switching Models", Journal of Marketing Research 5, 1968, pp. 225-228.

[35] Phillips, F. Y., Some Information Theoretic Methods for Management Analysis in Marketing and Resources, Ph.D. Dissertation, University of Texas at Austin, (1978).

[36] _____, "A Guide to MDI Statistics for Planning and Manage-Model Building", Institute for Constructive Capitalism Techical Series #2, The University of Texas at Austin (1980).

[37] Rubinson, J. R. and F. M. Bass, "A Note on 'A Parsimonious Description of the Hendry System'", Paper No. 68, (Lafayette, Ind.: Krannert School of Management, Purdue University, March 1978).

[38] Sabavala, D. J. and D. G. Morrison, "A Model of TV Show Loyalty", Journal of Advertising Research 17, No. 6, Dec. 1977, pp. 35-43.

[39] Theil, H., Economics and Information Theory, (Amsterdam: North Holland Publishing Co., 1967).

[40] _____, Principles of Econometrics, (New York: John Wiley & Sons, Inc. 1971).

[41] Weiss, L. and J. Wolfowitz, Maximum Probability and Related Topics, (Berlin: Springer-Verlag, 1974).

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER CCS 397	2. GOVT ACCESSION NO. AD-A109	3. RECIPIENT'S CATALOG NUMBER 147
4. TITLE (and Subtitle) An MDI Model and an Algorithm for Composite Hypotheses Testing and Estimation in Marketing.		5. TYPE OF REPORT & PERIOD COVERED 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) A. Charnes, W. W. Cooper, D. B. Learner, and F. Y. Phillips		8. CONTRACT OR GRANT NUMBER(s) N00014-81-C-0236
9. PERFORMING ORGANIZATION NAME AND ADDRESS Center for Cybernetic Studies, UT Austin Austin, Texas 78712		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research (Code 434) Washington, D.C.		12. REPORT DATE September 1981 13. NUMBER OF PAGES 31
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

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

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Market segmentation, market equilibrium, switching equilibrium, MDI statistic, composite hypotheses, mathematical programming, external constraints.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

A strategy is provided for using constrained versions of the MDI (minimum discrimination information) statistic to test and estimate market relations involving composite hypotheses. An algorithm for applying the tests and effecting the estimates is also provided along with numerical illustrations. Other, more general, developments in statistics and mathematical programming (duality) theories and methods are also briefly discussed for their possible bearing on further uses in marketing research and management.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)