Markov Chains Applied to Marketing*

GEORGE P. H. STYAN
and
HARRY SMITH, JR.†

➤ The classical approach to market behavioral analysis rarely uses data provided by the transitional, or switching, habits of the consumer. In this article, the authors have taken types of laundry powders purchased by a housewife to define the state space of a *Markov chain*. Using this model future purchase behavior is predicted, and statistical inferences on the switching habits are made.

In recent years there has been a great deal of interest in the applications of stochastic processes in industry. In particular, *Markov chain* models have been tried in quite a few areas. The number of papers appearing in both advertising and market research journals on this subject has been on the increase in the past few years [3]. This paper is an illustration of certain concepts of discrete time parameter finite Markov chains, a Markov chain being a stochastic process in which the future state of the system is dependent only on the present state and is independent of past history.

Naturally, one of the main concerns of any producer of consumer goods is to get people to use his product. Also important, however, is to get people repurchasing his product once they have used it. That is, the producer wants his customers to be *loyal customers*. However, complete loyalty is seldom found, and thus it is useful for the producer to have information on the switching habits of buyers in the commodity market.

This paper restricts itself to a discussion of laundry soap powder and detergent buying by a small panel of housewives over a six-month period of time. The loyalty and switching probabilities of these women are analyzed, and the limiting distribution of the Markov chain for this particular study is discussed.

THE SURVEY

A survey[1] of a hundred households or families in Leeds, Yorkshire, England, was made between January and June, 1957. The housewives in this survey reported their purchases of laundry cleaning products in a weekly diary. This information was tabulated and condensed for the purpose of this paper into a consideration of four mutually exclusive and wholly exhaustive categories of buying habits. These were

(1) *family buying detergent only,*
(2) *family buying soap powder only,*
(3) *family buying both detergent and soap powder together,*
and (4) *family buying no laundry powder at all.*

These four categories will be referred to as "states" in the Markov analysis. A summary of purchase information is shown in Table 1. One can see that over the 26-week period, the purchases were divided up on the average as follows:

Table 1
DISTRIBUTION OF HOUSEWIVES' PURCHASES*

Week Ending	*Period (k)*	*State* 1	2	3	4
1957					
January 5	1	22	44	9	25
January 12	2	18	44	8	30
January 19	3	18	38	11	33
January 26	4	22	44	10	24
February 2	5	24	39	11	26
February 9	6	25	35	9	31
February 16	7	28	37	9	26
February 23	8	28	39	9	24
March 2	9	20	41	9	30
March 9	10	19	34	5	42
March 16	11	23	40	2	35
March 23	12	21	37	5	37
March 30	13	23	40	5	32
April 6	14	16	45	3	36
April 13	15	24	41	5	30
April 20	16	19	39	4	38
April 27	17	18	44	6	32
May 4	18	21	50	6	23
May 11	19	22	48	4	26
May 18	20	17	49	4	30
May 25	21	22	45	7	26
June 1	22	16	43	5	36
June 8	23	17	41	7	35
June 15	24	20	41	4	35
June 22	25	22	37	5	36
June 29	26	22	37	5	36
Average		21.04	41.23	6.42	31.31

*Total 100 housewives per week.

---

* Originally presented at the 16th Annual ASQC Convention, Cincinnati, Ohio, in May 1962.

† George P. H. Styan is a graduate student in Columbia University's Department of Mathematical Statistics. Harry Smith, Jr., is Head of the Department of Mathematics and Statistics in the Procter and Gamble Company.

[1] Made by Attwood Statistics Ltd., London, England, for Procter and Gamble, Ltd. (then Thos. Hedley and Co. Ltd.), Newcastle-upon-Tyne, England.

(1) Detergents 21.04 percent
(2) Soap powders 41.23 percent
(3) Both powders 6.42 percent
(4) No powder 31.31 percent

The question which immediately arises is this:

"Are these percentages the result of loyal stable buyers, or is there a lot of switching taking place?"

To answer this question, one should consider information on successive purchases and formulate estimates of transition probability matrices. The first transition probability matrix estimate is shown in Table 2. For

Table 2

FIRST TRANSITION PROBABILITY MATRIX

		Second Week			
		Detergent	Soap powder	Both powders	No powder
First Week	Detergent	0.5454	0.0909	0.0455	0.3182
	Soap powder	0.0227	0.7955	0.0227	0.1591
	Both powders	0.1111	0.2222	0.6667	0
	No powder	0.1600	0.2000	0	0.6400

example, of the 22 households which purchased only detergents in the first week, 54.54 percent or 12 households purchased only detergents in the second week. This is an estimate of detergent loyalty. 9.09 percent of the 22 households switched to soap powder only, and so on. Similar two-period matrices were formed to cover the 26 weeks completely. Some 25 transition matrices were therefore constructed. These matrices are shown in frequency form in Table 3. They demonstrate that the buying habits of women are dependent upon their previous purchases. However, it is necessary to test such an implication statistically, and the test to be employed examines the "order of the Markov chain."

ORDER OF THE CHAIN

If a purchase of a particular type of laundry powder is dependent not only on the last purchase, but on the previous r (r > 1) purchases made, then the model being applied here is said to be an r^{th} *order* Markov chain.

Consider the null hypothesis:

H_0: A zero order Markov chain (or independent trials sequence); against the alternative hypothesis:

H_1: A first order Markov chain (or dependence being *only* on the state occupied immediately previously).

Following Anderson and Goodman [1], a χ^2 test of homogeneity is formulated in the appendix at the end of this article to test these hypotheses.

This test was carried out on the 25 matrices shown in Table 3, with the results as shown in Table 4. Each transition matrix yields a highly significant value of χ^2_9.

Table 3

TRANSITION FREQUENCY MATRICES*

$$F_1 = \begin{bmatrix} 12, 2, 1, 7 \\ 1, 35, 1, 7 \\ 1, 2, 6, 0 \\ 4, 5, 0, 16 \end{bmatrix} \quad F_2 = \begin{bmatrix} 11, 3, 1, 3 \\ 2, 28, 3, 11 \\ 0, 0, 7, 1 \\ 5, 7, 0, 18 \end{bmatrix} \quad F_3 = \begin{bmatrix} 13, 2, 0, 3 \\ 3, 27, 4, 5 \\ 3, 2, 6, 0 \\ 44, 13, 0, 16 \end{bmatrix}$$

$$F_4 = \begin{bmatrix} 15, 2, 2, 3 \\ 2, 28, 3, 11 \\ 2, 2, 6, 0 \\ 5, 7, 0, 12 \end{bmatrix} \quad F_5 = \begin{bmatrix} 14, 2, 1, 7 \\ 3, 25, 3, 8 \\ 1, 3, 5, 2 \\ 7, 5, 0, 14 \end{bmatrix} \quad F_6 = \begin{bmatrix} 16, 2, 1, 6 \\ 1, 23, 4, 7 \\ 2, 3, 3, 1 \\ 9, 9, 1, 12 \end{bmatrix}$$

$$F_7 = \begin{bmatrix} 18, 3, 1, 6 \\ 2, 26, 5, 4 \\ 3, 3, 2, 1 \\ 5, 7, 1, 13 \end{bmatrix} \quad F_8 = \begin{bmatrix} 16, 2, 2, 8 \\ 2, 28, 2, 7 \\ 0, 4, 4, 1 \\ 2, 7, 1, 14 \end{bmatrix} \quad F_9 = \begin{bmatrix} 13, 1, 0, 6 \\ 2, 29, 1, 9 \\ 0, 3, 3, 3 \\ 4, 1, 1, 24 \end{bmatrix}$$

$$F_{10} = \begin{bmatrix} 11, 3, 0, 5 \\ 1, 27, 0, 6 \\ 1, 1, 2, 1 \\ 10, 9, 0, 23 \end{bmatrix} \quad F_{11} = \begin{bmatrix} 14, 0, 1, 8 \\ 1, 28, 2, 9 \\ 0, 0, 2, 0 \\ 6, 9, 0, 20 \end{bmatrix} \quad F_{12} = \begin{bmatrix} 16, 3, 0, 2 \\ 1, 29, 0, 7 \\ 1, 2, 2, 0 \\ 5, 6, 3, 23 \end{bmatrix}$$

$$F_{13} = \begin{bmatrix} 14, 1, 0, 8 \\ 1, 31, 2, 6 \\ 1, 2, 0, 2 \\ 0, 11, 1, 20 \end{bmatrix} \quad F_{14} = \begin{bmatrix} 15, 0, 0, 1 \\ 2, 30, 2, 11 \\ 1, 0, 2, 0 \\ 6, 11, 1, 18 \end{bmatrix} \quad F_{15} = \begin{bmatrix} 16, 3, 1, 4 \\ 1, 29, 1, 10 \\ 0, 0, 2, 3 \\ 2, 7, 0, 21 \end{bmatrix}$$

$$F_{16} = \begin{bmatrix} 14, 1, 0, 4 \\ 1, 32, 0, 6 \\ 0, 0, 4, 0 \\ 3, 11, 2, 22 \end{bmatrix} \quad F_{17} = \begin{bmatrix} 14, 2, 0, 2 \\ 1, 35, 1, 7 \\ 0, 2, 4, 0 \\ 6, 11, 1, 14 \end{bmatrix} \quad F_{18} = \begin{bmatrix} 17, 2, 0, 2 \\ 1, 36, 1, 12 \\ 0, 1, 3, 2 \\ 4, 9, 0, 10 \end{bmatrix}$$

$$F_{19} = \begin{bmatrix} 15, 2, 0, 5 \\ 0, 37, 1, 10 \\ 1, 2, 1, 0 \\ 1, 8, 2, 15 \end{bmatrix} \quad F_{20} = \begin{bmatrix} 15, 0, 1, 1 \\ 1, 38, 2, 8 \\ 1, 0, 3, 0 \\ 5, 7, 1, 17 \end{bmatrix} \quad F_{21} = \begin{bmatrix} 13, 3, 0, 6 \\ 1, 31, 3, 10 \\ 0, 3, 2, 2 \\ 2, 6, 0, 18 \end{bmatrix}$$

$$F_{22} = \begin{bmatrix} 12, 1, 0, 3 \\ 1, 29, 4, 9 \\ 1, 1, 3, 0 \\ 3, 10, 0, 23 \end{bmatrix} \quad F_{23} = \begin{bmatrix} 12, 0, 0, 5 \\ 1, 28, 0, 12 \\ 1, 3, 3, 0 \\ 6, 10, 1, 18 \end{bmatrix} \quad F_{24} = \begin{bmatrix} 13, 2, 0, 5 \\ 4, 26, 0, 11 \\ 0, 0, 4, 0 \\ 5, 9, 1, 20 \end{bmatrix}$$

$$F_{25} = \begin{bmatrix} 14, 3, 0, 5 \\ 3, 28, 1, 5 \\ 0, 0, 4, 1 \\ 5, 6, 0, 25 \end{bmatrix}$$

* F_k denotes transition from period k to period k + 1, k = 1,2,...,25

Table 4

χ^2 TEST RESULTS FOR ZERO VERSUS FIRST ORDER MARKOV CHAIN*

k	χ^2_9	k	χ^2_9	k	χ^2_9
1	91.08	9	89.55	18	96.98
2	96.08	10	82.49	19	77.70
3	79.74	11	91.57	20	111.23
4	68.99	12	89.53	21	68.29
5	58.49	13	69.88	22	90.85
6	44.83	14	89.68	23	78.61
7	51.76	15	87.96	24	118.77
8	68.09	16	139.53	25	124.22
		17	100.11		

* k denotes the transition from week k to week k + 1 [k = 1,2, ...,25], and χ^2_9 the calculated value of χ^2 with nine degrees of freedom as given by (1) in the appendix.

The null hypothesis of a zero order Markov chain is *rejected* at the 0.1 percent level in all 25 cases. That is, the hypothesis that the observations at successive periods are statistically independent is rejected.

With diary information, sequential tests for higher order Markov chains can be made. These tests require a very large number of observed transitions, and such data were unavailable in this study. Only a first order Markov chain is to be considered in this article.

STATIONARITY

As is indicated by the matrices in Table 3, the loyalty probability for each state is much higher than the switching probability between two different states. To consider the possibility of *regularities* throughout the 26-week period, one may graph the loyalty probabilities[2] for each state, as in Figure 1. One can see relatively stable loyalties, with the exception of Figure 3 where the wild fluctuations are clearly due to the small frequencies observed for this state.

This leads to the possibility that the transition probabilities are "stationary," that is, independent of time or purchase period. A system with stationary transition probabilities is called a *homogeneous Markov chain.*

Denote the transition probability[2] from state i to state j at periods k and k + 1 respectively as $p_{ij}(k)$, $i,j = 1,2,3,4$; $k = 1,2,\ldots,25$. Then to test the hypothesis of stationary transition probabilities, consider the null hypothesis:

$$H_0\colon p_{ij}(k) = p_{ij} \qquad \text{for } all\ k = 1,2,\ldots,25$$

against the composite alternative hypothesis:

$$H_1\colon p_{ij}(k) \qquad \text{dependent on the period } k.$$

A likelihood ratio test can be used to test these hypotheses and is formulated in the appendix. Asymptotically one finds an equivalent standard normal variate as given by (3) in the appendix, which for the data under analysis proves to be 0.69. This is clearly not significant, since the corresponding significance level is over 24 percent.

Hence one has insufficient evidence to reject the null hypothesis and hence may consider a homogeneous Markov chain model. Thus one represents the system by a single "stationary" transition probability matrix P, the maximum likelihood estimate of which is shown in Table 5. The switching pattern of the system is therefore taken to be independent of time.

LIMITING DISTRIBUTION

Inspection of **P** shows that it is possible to move from every state to every other state. No particular switch is therefore impossible. Such a matrix defining a Markov chain has the property that

Table 5

MAXIMUM LIKELIHOOD ESTIMATE OF THE *STATIONARY* TRANSITION PROBABILITY MATRIX **P**

		Second Week			
		Detergent	*Soap powder*	*Both powders*	*No powder*
	Detergent	0.6724	0.0857	0.0229	0.2190
First	Soap powder	0.0367	0.7179	0.0444	0.2010
Week	Both powders	0.1235	0.2407	0.5123	0.1235
	No powder	0.1465	0.2584	0.0219	0.5732

$$\lim_{n\to\infty} \mathbf{P}^n = \mathbf{E}$$

where **E** is an *idempotent* matrix with all its rows the same, and each row adding to unity. A matrix is said to be *idempotent* when its square equals itself, in this case $\mathbf{E} = \mathbf{E}^2$. The row defining **E** will be a vector of probabilities, and also the left-hand characteristic vector of **P** corresponding to its characteristic root of unity. If **e** denotes a column vector of unities, then one may write $\mathbf{E} = \mathbf{e}\,\mathbf{l}'$, and $\mathbf{l}'\,\mathbf{P} = \mathbf{l}'$, where $\mathbf{l}'$ is the row in question. Thus if the Markov chain starts with probabilities given by the elements in $\mathbf{l}'$, it will always have these probabilities. One says that $\mathbf{l}'$ defines the limiting or stationary distribution of the Markov chain. Thus in this case the elements of $\mathbf{l}'$ will contain the shares of the market attained by the various states if the switching pattern defined by **P** were to persist for a long period of time. It thus gives an indication of where the market is heading, which is a useful piece of information for the market strategist. Comparison of $\mathbf{l}'$ with the current market shares will indicate how far from stationarity the current market *distribution* is.

In this case, one finds that

$$\mathbf{l}' = (21.14,\ 40.85,\ 6.14,\ 31.87) \text{ percent}$$

while observed share vectors (from Table 1) are found to be

Week 1:	(22.00,	44.00,	9.00,	25.00)	percent
Week 26:	(22.00,	37.00,	5.00,	36.00)	percent
Average:	(21.04,	41.23,	6.42,	31.31)	percent

The closeness of $\mathbf{l}'$ to the average market share indicates that the market *distribution* is approximately stationary throughout the 26 weeks considered.

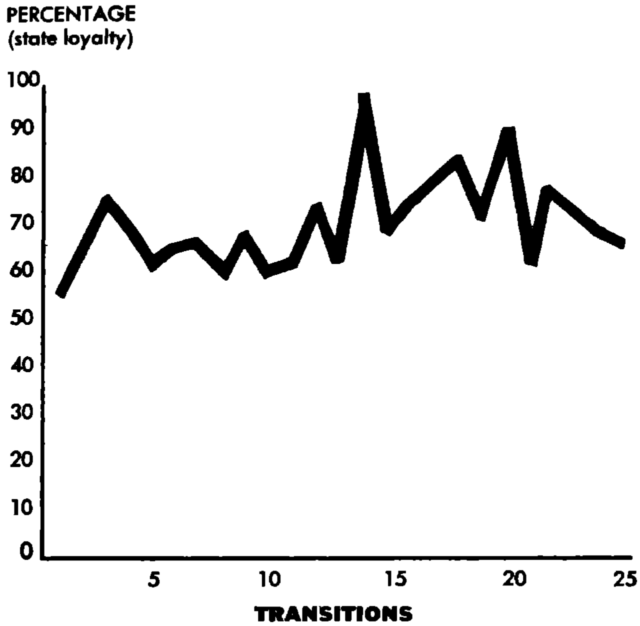
It has thus been found that the data from this survey fit fairly well to the model of a first order Markov chain with stationary transition probabilities and with, on the average, a stationary market share distribution.

PREDICTION

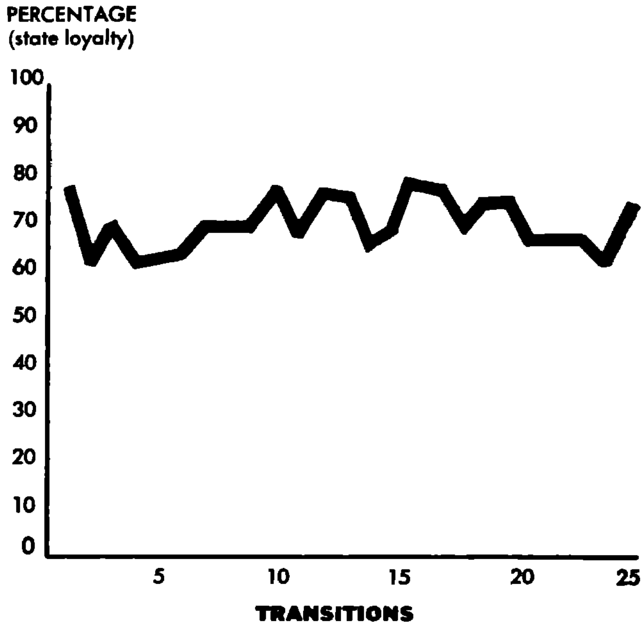
One of the possible further applications of Markov chains is prediction of future market positions. If one assumes the next position primarily dependent on the

[2] Maximum likelihood estimates of these are found by dividing the transition frequency (as shown in Table 3) by the initial period frequency (as shown in Table 1), and it is these that are used.

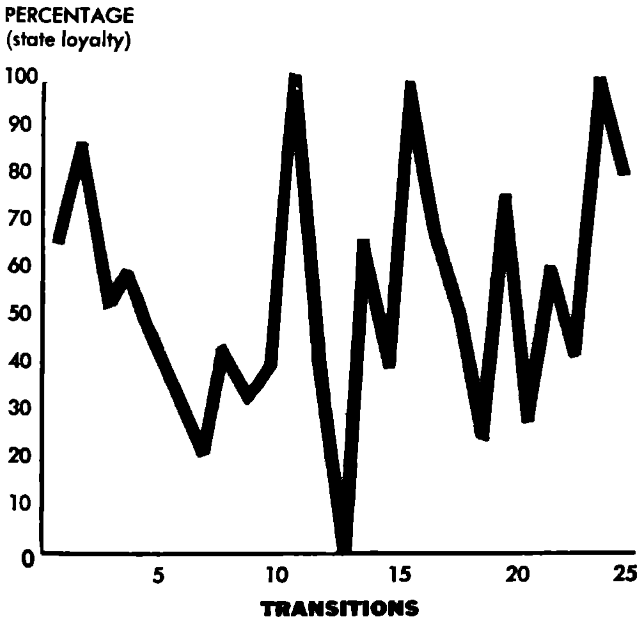
FIGURE 1
LOYALTY PROBABILITIES BY STATE



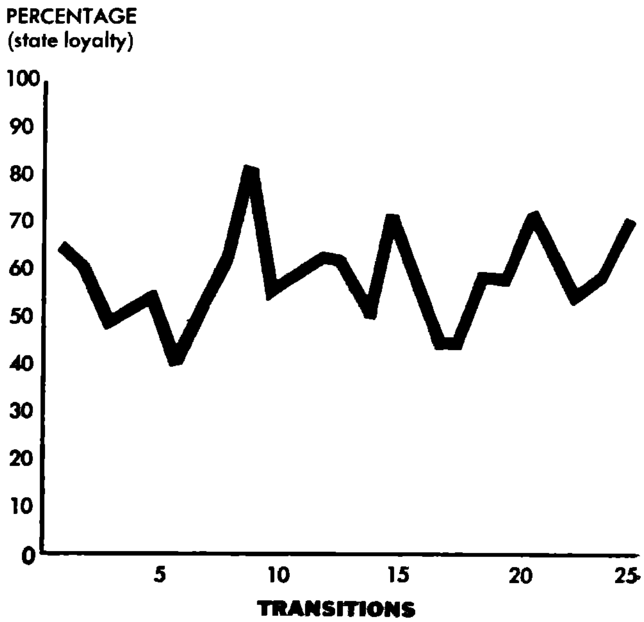
A. DETERGENT LOYALTY



B. SOAP POWDER LOYALTY



C. BOTH POWDER LOYALTY



D. NO POWDER LOYALTY

present position, then the following method of prediction can be used.

Let p'_k denote the row vector of market positions for period k. Then an estimate of p'_{k+1} is found by forming the product $p'_k P$, where P is the transition probability matrix, estimated over the largest number of previous transitions for which the χ^2-test confirmed stationarity of the transition probability elements of P. For example, the first prediction can be made for period 3, using p'_2, with P estimated by F_1 in Table 3. That is $p'_2 P$ estimates p'_3. Further estimates follow similarly.

However, because of the stationary nature of the data at hand, it is unnecessary to use this procedure for prediction in this case. Prediction here can be made by the stationary distribution (I') alone, which is in fact the same as the naïve persistence model with these data.

SUMMARY

Consecutive purchases of laundry cleaning powders by a small panel of housewives over a six-month period have been considered. First, a classification was made of the various powders, into four mutually exclusive and wholly exhaustive states. Frequency data for these four states were examined and found to be fairly stable. Transition probability matrices were estimated, giving the probability of occurrence of state j in week k + 1, immediately after state i in week k. The duration of the stage of the system is approximately equal to the period [2] of use of a laundry powder. Each such matrix was considered to define a separate Markov chain, and with the aid of a statistical test, evidence was found that it was of first-order rather than of zero-order, the independent trials sequence.

All the transition matrices were then tested to see if they were simply random deviates from some underlying matrix. Again using a statistical test, evidence was found to support this. The unique matrix thus found was considered to define a homogeneous Markov chain. Inspection showed that the average distribution of the market during the period concerned was very nearly equal to the stationary distribution of the Markov chain.

Finally a method of using the transition probabilities as a predictive device to estimate future market positions was described. Markov chains are thus seen to be useful in marketing as a predictive device for short and long range purposes. They can also be used to indicate which brand in the market place gains (or loses) its customers from (or to) which other brand.

MATHEMATICAL APPENDIX

Homogeneity Test for Order

Let $F_k = [f_{ij}(k)]$, denote a transition frequency matrix, where $f_{ij}(k)$ denotes the frequency of occurrence of state j in period k + 1 and state i in period k (i,j = 1,2,...,m; k = 1,2,...,t).

Let $\sum_{i=1}^{m} f_{ij}(k) = f_{.j}(k)$; $\sum_{j=1}^{m} f_{ij}(k) = f_{i.}(k)$

$$\sum_{i=1}^{m}\sum_{j=1}^{m} f_{ij}(k) = \sum_{i=1}^{m} f_{i.}(k) = \sum_{j=1}^{m} f_{.j}(k) = f_{..}(k)$$

Anderson and Goodman [1] show that

$$\sum_{i=1}^{m}\sum_{j=1}^{m} \frac{f_{i.}(k)\left[\frac{f_{ij}(k)}{f_{i.}(k)} - \frac{f_{.j}(k)}{f_{..}(k)}\right]^2}{f_{.j}(k)/f_{..}(k)} \sim \chi^2_{(m-1)^2} \quad \ldots (1)$$

where ~ indicates that the ratio of the distribution functions of the left-hand side to that of the right-hand side converges to 1 as $f_{..}(k)$ tends to infinity.

In this case m = 4, $f_{..}(k) = 100$. Hence after some simplification, it can be seen that

$$100\left[\sum_{i=1}^{4}\sum_{j=1}^{4}\frac{f^2_{ij}(k)}{f_{i.}(k)f_{.j}(k)} - 1\right] \sim \chi^2_9$$

Likelihood Ratio Test for Sationarity

Let $F = \sum_{k=1}^{t} F_k = [f_{ij}]$, where $\sum_{k=1}^{t} f_{ij}(k) = f_{ij}$.

Then the maximum likelihood estimate of the stationary transition probability[2] between states i and j is

$$\hat{p}_{ij} = f_{ij}/f_{i.}, \text{ where } f_{i.} = \sum_{j=1}^{m} f_{ij},$$

and the maximum likelihood estimate of the transition probability between states i and j at periods k and k + 1 is

$$\hat{p}_{ij}(k) = f_{ij}(k)/f_{i.}(k) \qquad (k = 1,2,\ldots,t)$$

Following Anderson and Goodman [1], one defines the likelihood ratio criterion λ as

$$\lambda = \prod_{i,j=1}^{m}\prod_{k=1}^{t}\left(\frac{\hat{p}_{ij}}{\hat{p}_{ij}(k)}\right)^{f_{ij}(k)} \quad \ldots (2)$$

For large $f_{..}(k)$ one has

$$-2 \log_e \lambda \sim \chi^2_{m(m-1)(t-1)}.$$

In this case m = 4, t = 25, so with some reduction of (2) one finds that approximately

$$4.6052\left[\sum_{k=1}^{25}\sum_{i=1}^{4}\sum_{j=1}^{4} f_{ij}(k)\log_{10} f_{ij}(k) + \right.$$

$$\sum_{i=1}^{4} f_{i.} \log_{10} f_{i.} - \sum_{i=1}^{4} \sum_{j=1}^{4} f_{ij} \log_{10} f_{ij} -$$

$$\left. \sum_{k=1}^{25} \sum_{i=1}^{4} f_{i.}(k) \log_{10} f_{i.}(k) \right] \sim \chi^2_{288}.$$

When $m(m-1)(t-1) > 100$, as in this case, tables of χ^2 are not readily available. In such cases Fisher's approximation to an equivalent normal deviate can be used, i.e.,

$$\sqrt{2\chi^2_\nu} \sim N(\sqrt{2\nu - 1}, 1) \qquad \ldots (3)$$

where ν is the number of degrees of freedom. In this case one finds

$$\sqrt{2\chi^2_{288}} \sim N(23.98, 1).$$

REFERENCES

1. T. W. Anderson and L. A. Goodman, "Statistical Inference about Markov Chains," *Annals of Mathematical Statistics*, 28 (1957), 89-110.
2. R. A. Howard, "Stochastic Process Models of Consumer Behavior," *Journal of Advertising Research*, 3 (September 1963).
3. B. Lipstein, "The Dynamics of Brand Loyalty and Brand Switching," *Annual Conference of the Advertising Research Foundation*, September 1959.