



Calhoun: The NPS Institutional Archive
DSpace Repository

Reports and Technical Reports

All Technical Reports Collection

1973-12

Statistical aspects of lumpability hypotheses for Markov chains

Thomas, Marlin Uluess; Barr, Donald Roy

Monterey, California. Naval Postgraduate School

<http://hdl.handle.net/10945/29522>

This publication is a work of the U.S. Government as defined in Title 17, United States Code, Section 101. Copyright protection is not available for this work in the United States.

Downloaded from NPS Archive: Calhoun



Calhoun is the Naval Postgraduate School's public access digital repository for research materials and institutional publications created by the NPS community. Calhoun is named for Professor of Mathematics Guy K. Calhoun, NPS's first appointed -- and published -- scholarly author.

Dudley Knox Library / Naval Postgraduate School
411 Dyer Road / 1 University Circle
Monterey, California USA 93943

<http://www.nps.edu/library>

NPS 55BN55T073121A

NAVAL POSTGRADUATE SCHOOL
Monterey, California



STATISTICAL ASPECTS OF LUMPABILITY
HYPOTHESES FOR MARKOV CHAINS

by

M. U. Thomas

and

D. R. Barr

December 1973

Approved for public release; distribution unlimited.

FEDDOCS
D 208.14/2:NPS-55BN55T073121A

NAVAL POSTGRADUATE SCHOOL
Monterey, California

Rear Admiral M. B. Freeman
Superintendent

J. R. Borsting
Provost

ABSTRACT

Under certain conditions the state space of a discrete parameter Markov Chain may be partitioned to form a smaller "lumped" chain that retains the Markov property. The problem of formulating lumpability hypotheses when the transition probability matrix P is not known and, hence, must be estimated is discussed. An approximate test of these hypotheses is described based on well known non-parametric methods. The procedure is illustrated by an example.

Prepared by:

INTRODUCTION

Under certain conditions the state space of a discrete parameter Markov Chain may be partitioned into subsets of states each of which may be treated as a single state of a smaller chain that retains the Markov property. Such a chain is said to be "lumpable" or "weakly lumpable", depending upon conditions, and elsewhere has been referred to as a "mergable process" [3] and a "chain with collapsed states" [2]. The resulting smaller chain is called a "lumped chain"; conditions allowing lumping are discussed in detail by Kemeny and Snell [4]. A practical problem arises in examining lumpability conditions when the matrix of transition probabilities P is not known, but a number of transitions have been observed. Billingsley [1] discusses related problems of statistical inference for Markov Chains, but statistical considerations of lumpability have apparently not been investigated.

In particular, we consider an aperiodic Markov Chain

$\{X_t: t = 0, 1, 2, \dots\}$ with finite state space $S = \{1, \dots, n\}$ and stationary transition probability matrix $P = [p_{ij}]$. It is convenient to restrict ourselves to the case where $\{X_t\}$ is irreducible. Thus, the chain is described by P , a vector of steady state probabilities $\pi = (\pi_1, \dots, \pi_n)$, and possibly an a priori distribution of initial states, p^0 . Given observations of k transitions of this chain we obtain a matrix of transition counts $[n_{ij}]$ where n_{ij} is the number of transitions into state j from state i , and, of course,

$$\sum_{i=1}^n \sum_{j=1}^n n_{ij} = k.$$

Maximum likelihood estimators for the one-step transition probabilities are given in [1]

$$\hat{p}_{ij} = n_{ij} / \sum_{j=1}^n n_{ij} = n_{ij} / n_{i\cdot}, \quad (1)$$

where $n_{i\cdot}$ is the observed frequency with which the process visited state i . If a lumpability hypothesis (which we discuss below) was formulated independent of the sample of k transitions, it could be tested in terms of the asymptotic χ^2 -theory involving differences between observed and expected frequencies [1], [6]. More often in practice, however, the hypothesis to be tested is suggested by the sample of observed transitions. Technically, this gives rise to a problem of the simultaneous inference type [5]; for "large" samples, the magnitude of the error (usually lower than the desired size and power) is insignificant. Thus, in approximate tests, based on asymptotic distributions of the test statistics, the effect of formulating the hypotheses to be tested with the aid of the data to be used for the test is usually ignored. In what follows we shall discuss the use of asymptotic χ^2 -theory in testing hypotheses of lumpability. We begin under the assumption that the null hypothesis has been determined, perhaps through reasoning about the physical system being modeled, or perhaps through preliminary examination of data from $\{X_t\}$. Later, in section 4, we make some comments about the problem of hypothesis formulation.

1. LUMPABILITY HYPOTHESES

Consider an n -state Markov Chain $\{X_t: t = 0, 1, 2, \dots\}$.

Formally, we have the following:

Definition: $\{X_t\}$ is lumpable with respect to a partition $\tilde{S} = \{L_1, L_2, \dots, L_m\}$ of S , where $m < n$, if for every initial state probability vector \tilde{p}^0 the resulting chain $\{\tilde{X}_t\}$ is Markovian and the transition probabilities \tilde{p}_{ij} are invariant under choices of \tilde{p}^0 .

A necessary and sufficient condition for $\{X_t\}$ to be lumpable with respect to a partition \tilde{S} of S is that for each pair (L_i, L_j) , the probability of transition from k to some $\ell \in L_j$ is the same for each $k \in L_i$ (Theorem 6.3.2 [4]). We shall use this characterization in stating hypotheses of lumpability. The resulting lumped chain $\{\tilde{X}_t\}$ will be Markovian with transition probabilities \tilde{p}_{ij} , where, for each $k \in L_i$,

$$\tilde{p}_{ij} \equiv \sum_{\ell \in L_j} p_{k\ell}; \quad i, j = 1, \dots, m. \quad (2)$$

The steady state probability vector $\tilde{\pi} = (\tilde{\pi}_1, \dots, \tilde{\pi}_m)$ of $\{\tilde{X}_t\}$ has components

$$\tilde{\pi}_j = \sum_{\ell \in L_j} \pi_\ell; \quad j = 1, \dots, m,$$

and the corresponding prior \tilde{p}^0 is similarly determined from p^0 by pooling over the states in the L_i 's.

To illustrate a lumpable Markov Chain, consider a 5-state chain with transition probabilities $[p_{ij}: (i, j) = 1, \dots, 5]$. Suppose that this chain is lumpable into $\tilde{S} = \{\{1\}, \{2, 4\}, \{3, 5\}\} = \{L_1, L_2, L_3\}$. Then the transition probability matrix for $\{\tilde{X}_t\}$ is given by

$$\tilde{P} = \begin{bmatrix} p_{11} & p_{12} + p_{14} & p_{13} + p_{15} \\ p_{21} & p_{22} + p_{24} & p_{23} + p_{25} \\ p_{31} & p_{32} + p_{34} & p_{33} + p_{35} \end{bmatrix}. \quad (3)$$

It follows from (2) that

$$\begin{aligned} p_{21} &= p_{41}, & p_{22} + p_{24} &= p_{42} + p_{44}, & p_{23} + p_{25} &= p_{43} + p_{45}, \\ p_{31} &= p_{51}, & p_{32} + p_{34} &= p_{52} + p_{54}, & p_{33} + p_{35} &= p_{53} + p_{55}. \end{aligned} \quad (4)$$

Here, $\{\tilde{X}_t\}$ will be Markovian for an arbitrary choice of initial state probability vector. Burke and Rosenblatt [2] give weaker lumpability conditions that apply whenever there exists at least one choice of \tilde{p}^0 such that $\{\tilde{X}_t\}$ is Markovian. In either case, in practice one makes conjectures (in the form of hypotheses) about combining certain states, which result in forming postulated probability transition matrices (of lumped chains), which in turn satisfy conditions such as those in (3) characterizing lumpability into these combined states.

2. TEST OF LUMPABILITY

Let us denote the hypothesis that $\{X_t\}$ is lumpable into $\tilde{S} = \{L_1, \dots, L_m\}$ by the partition \tilde{S} itself, and suppose we take as the alternate the composite hypothesis that $\{X_t\}$ is not lumpable into \tilde{S} :

$$H_0: \{L_1, \dots, L_m\} \text{ v.s. } H_a: \text{not } - \{L_1, \dots, L_m\}.$$

With the characterization of lumpability discussed above, H_0 is equivalent to stating that, in addition to satisfying the conditions of a stochastic matrix, $[p_{ij}]$ satisfies conditions (2).

The random variable

$$\sum_{i=1}^n \sum_{j=1}^n \frac{(n_{ij} - n_{i \cdot} p_{ij})^2}{n_{i \cdot} p_{ij}} \quad (5)$$

is asymptotically (as $k \rightarrow \infty$) χ^2 -distributed with $n(n-1)$ degrees of freedom ([1], Theorem 5.3). However, the p_{ij} are unknown, so we must apply the well known procedure of reducing degrees of freedom to account for estimation of parameters in (5). For example, Roy [6] (Theorem 5, page 126) states the rule in a form appropriate for the current context. We take the random variable (5), with the p_{ij} 's replaced by the corresponding maximum likelihood estimators \hat{p}_{ij} , as the test statistic. H_0 is rejected if the calculated value of this statistic falls above the tabulated χ^2 quantile corresponding to the desired level of significance, α . Note that in calculating the \hat{p}_{ij} we must use the constraints, such as those given in (4) for our example, corresponding to H_0 . In addition, we must use the constraints to determine the appropriate reduction in the degrees of freedom.

For a given null hypothesis $\{L_1, \dots, L_m\}$, let λ_i denote the number of the original n states present in L_i . By proper initial choice of labels for the states in S , it is possible to state the null hypothesis in the form

$$H_0: \tilde{S} = \{ \{1, 2, \dots, \lambda_1\}, \{\lambda_1 + 1, \dots, \lambda_1 + \lambda_2\}, \dots, \\ \{ \sum_{j=1}^{m-1} \lambda_j + 1, \dots, n \} \}.$$

We wish to estimate the n^2 parameters p_{ij} subject to

$$\sum_{j=1}^n p_{ij} = 1; \quad i = 1, 2, \dots, n$$

(n constraints), and

$$\sum_{j \in L_s} p_{i'j} = \sum_{j \in L_s} p_{i''j}; \quad i' \neq i'' \text{ both in } L_i; \quad i = 1, 2, \dots, m, \quad (6)$$

adding $\sum_{i=1}^m (\lambda_i - 1) \cdot m = m(n - m)$ constraints. Thus, using Roy's rule-of-thumb, the number of "independent" parameters we need to estimate is $n^2 - n - m(n - m)$, so the degrees of freedom of the test statistic is simply $n^2 - [n^2 - n - m(n - m)] = n + m(n - m)$.

The maximum likelihood estimators \hat{p}_{ij} of the p_{ij} under the above constraints can be derived using Lagrangian multipliers with the log likelihood function $\sum_{i,j} n_{ij} \log p_{ij}$. The form of these estimators have the following intuitively appealing interpretation: suppose $k \in L_i$ and $q \in L_s$, in order to estimate p_{kq} , first form a maximum likelihood estimate of $\sum_{j \in L_s} p_{kj}$, where L_s contains q . By equation (6), it is not surprising that this estimate turns out to be a "pooled" estimate,

$$\widehat{\sum_{j \in L_s} p_{kj}} = \frac{\sum_{k \in L_i} \sum_{j \in L_s} n_{kj}}{\sum_{k \in L_i} n_{k\cdot}}, \quad (7)$$

where, as before, $n_{k\cdot} = \sum_j n_{kj}$. The proper allocation of the combined estimate (7) over the individual cells p_{kj} , for each $j \in L_s$, is obtained by weighting (7) by the relative frequencies $n_{kj} / \sum_{j \in L_s} n_{kj}$.

The maximum likelihood estimates of the p_{kq} are thus given by

$$\hat{p}_{kq} = \frac{\sum_{k \in L_i} \sum_{j \in L_s} n_{kj}}{\sum_{k \in L_i} n_{k\cdot}} \left[\frac{n_{kq}}{\sum_{j \in L_s} n_{kj}} \right]. \quad (8)$$

Replacing the unknown p_{ij} in (5) by their estimates \hat{p}_{ij} given above results in a test statistic which is distributed approximately χ^2 with $n + m(n - m)$ degrees of freedom.

In summary, the procedure for conducting a test of the hypothesis \tilde{S} of lumpability, at approximately the α -level of significance, is as follows:

1. Use the observed record $\{x_1, x_2, \dots, x_{n+1}\}$ to form the transition frequency matrix (n_{ij}) .
2. Compute the estimates \hat{p}_{ij} given in (8).
3. Calculate the value of the test statistic (5), with \hat{p}_{ij} in place of the unknown p_{ij} .
4. Reject the hypothesis of lumpability if the calculated value of the test statistic exceeds the tabulated $(1 - \alpha)$ th quantile of the χ^2 -distribution with $n + m(n - m)$ degrees of freedom.

3. A NUMERICAL EXAMPLE

Consider a special case of our earlier example, where

$$P = \begin{bmatrix} .3 & .1 & .2 & .1 & .3 \\ .1 & .3 & .1 & .3 & .2 \\ .5 & .1 & 0 & .1 & .3 \\ .1 & .5 & .2 & .1 & .1 \\ .5 & 0 & .1 & .2 & .2 \end{bmatrix},$$

with $\tilde{S} = \{\{1\}, \{2, 4\}, \{3, 5\}\}$, so

$$\tilde{P} = \begin{bmatrix} .3 & .2 & .5 \\ .1 & .6 & .3 \\ .5 & .2 & .3 \end{bmatrix}.$$

We generated 1000 transitions with P , using a table of random numbers, resulting in the frequency matrix

$$(n_{ij}) = \begin{bmatrix} 84 & 31 & 52 & 31 & 112 \\ 22 & 46 & 13 & 54 & 33 \\ 69 & 9 & 0 & 13 & 31 \\ 14 & 83 & 33 & 23 & 16 \\ 118 & 0 & 23 & 48 & 42 \end{bmatrix} .$$

Imagine P is unknown, and we wish to use the data in (n_{ij}) to test $H_0: \tilde{S}$. The usual (without lumpability constraints) maximum likelihood estimate of P is given by

$$\hat{P}^* = \begin{bmatrix} .27 & .10 & .17 & .10 & .36 \\ .13 & .27 & .08 & .32 & .20 \\ .57 & .07 & 0. & .11 & .25 \\ .08 & .49 & .20 & .14 & .09 \\ .51 & 0. & .10 & .21 & .18 \end{bmatrix} .$$

Under the hypothesized lumpability conditions, the matrix of estimates

(\hat{p}_{ij}) is

$$\hat{p} = \begin{bmatrix} .270 & .100 & .170 & .100 & .360 \\ .107 & .281 & .080 & .330 & .202 \\ .530 & .081 & 0. & .117 & .272 \\ .107 & .479 & .190 & .133 & .091 \\ .530 & 0. & .096 & .198 & .176 \end{bmatrix} .$$

The value of the test statistic is 3.03, which falls well below the $\alpha = .05$ χ^2 critical value 19.68 with $5 + 3(5 - 3) = 11$ degrees of freedom. We would thus conclude the observed data is consistent with the hypothesis of lumpability, in the sense that the test value is not significant at the .05 level. Of course, in this case with P known, H_0 is known to be true; the "test" is simply an illustration of how we would have proceeded if P had not been known.

4. COMMENTS

We have discussed a test of a given lumpability hypothesis; the problem of using the observed data both to formulate the hypothesis as well as test it has been mentioned. Even if one disregards this problem, there is a very significant problem in how to use the data to formulate appropriate hypotheses. A solution of this problem would be of great interest, for example, in large computer based information systems, where man's intuition is not sufficient to cope with the range of possible alternatives.

REFERENCES

- [1] Billingsley, P., Statistical Inference for Markov Processes, University of Chicago Press, Chicago, 1961.
- [2] Burke, C. J. and Rosenblatt, M., "A Markovian Function of a Markov Chain," *Ann. Math. Stat.*, Vol. 29, No. 4, 1958, pp. 1,112-22.
- [3] Howard, R. A., Dynamic Probabilistic Systems, Volume I: Markov Models, Wiley, New York, 1971.
- [4] Kemeny, J. G. and Snell, J. L., Finite Markov Chains, Van Nostrand, Princeton, New Jersey, 1967.
- [5] Miller, R. B., Simultaneous Statistical Inference, McGraw-Hill, New York, 1966.
- [6] Roy, S. N., Some Aspects of Multivariate Analysis, John Wiley and Sons, Inc., New York, 1957.

INITIAL DISTRIBUTION LIST

	No. of Copies
Defense Documentation Center (DDC) Cameron Station Alexandria, Virginia 22314	12
Dean of Research Code 023 Naval Postgraduate School Monterey, California 93940	1
Library (Code 0212) Naval Postgraduate School Monterey, California 93940	2
Library (Code 55) Department of Operations Research and Administrative Sciences Naval Postgraduate School Monterey, California 93940	3
Dr. Thomas Varley Office of Naval Research Arlington, Virginia 2217	1
Professor Ralph Disney Department of Industrial & Operations Engineering University of Michigan Ann Arbor, Michigan 48104	1
Professor P. Billingsley Department of Statistics University of Chicago Chicago, Illinois 60637	1
Professor C. J. Burke Department of Mathematics California State College Hayward, California 94542	1
Professor G. J. Lieberman Department of Operations Research Stanford University Stanford, California 94305	1
Dr. Richard D. Smallwood Xerox Palo Alto Research Center Palo Alto, California 94304	1

No. of Copies

Professor E. A. Brill	1
Professor R. W. Butterworth	1
Professor J. D. Esary	1
Professor D. P. Gaver	1
Professor P. A. W. Lewis	1
Professor K. T. Marshall	1
Professor P. R. Milch	1
Professor F. R. Richards	1
Professor R. A. Stephan	1
Professor A. R. Washburn	1
Department of Operations Research and Administrative Sciences Naval Postgraduate School Monterey, California 93940	
Professor D. R. Barr	12
Department of Operations Research and Administrative Sciences Naval Postgraduate School Monterey, California 93940	
Professor M. U. Thomas	12
Department of Operations Research and Administrative Sciences Naval Postgraduate School Monterey, California 93940	

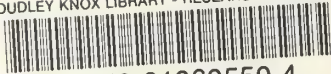
UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER NPS55BN55T073121A	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) STATISTICAL ASPECTS OF LUMPABILITY HYPOTHESES FOR MARKOV CHAINS		5. TYPE OF REPORT & PERIOD COVERED Technical Report
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) M. U. Thomas and D. R. Barr		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Postgraduate School Monterey, California 93940		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS		12. REPORT DATE December 1973
		13. NUMBER OF PAGES 16
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report)
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution 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) Lumpable Markov Chains, Markov Process with Collapsible States Mergable Process, Markov Chains, Statistical Test		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) Under certain conditions the state space of a discrete parameter Markov Chain may be partitioned to form a smaller "lumped" chain that retains the Markov property. The problem of formulating lumpability hypotheses when the transition probability matrix P is not known and, hence, must be estimated is discussed. An approximate test of these hypotheses is described based on well known non-parametric methods. The procedure is illustrated by an example.		

U167729

DUDLEY KNOX LIBRARY - RESEARCH REPORTS



5 6853 01060550 4