



Some Applications of Quasisymmetry

Peter McCullagh

Biometrika, Volume 69, Issue 2 (Aug., 1982), 303-308.

Stable URL:

<http://links.jstor.org/sici?sici=0006-3444%28198208%2969%3A2%3C303%3ASAOQ%3E2.0.CO%3B2-H>

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

Biometrika is published by Biometrika Trust. Please contact the publisher for further permissions regarding the use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/bio.html>.

Biometrika

©1982 Biometrika Trust

JSTOR and the JSTOR logo are trademarks of JSTOR, and are Registered in the U.S. Patent and Trademark Office. For more information on JSTOR contact jstor-info@umich.edu.

©2003 JSTOR

Some applications of quasisymmetry

By PETER McCULLAGH

Department of Mathematics, Imperial College, London

SUMMARY

We give three applications of quasisymmetry. The first two applications involve derived conditional distributions in a study involving matched individuals where the parameters associated with the matching are eliminated by appropriate conditioning. Extensions to matched triplets or to larger blocks lead to generalizations of quasisymmetry. The third application involves reversibility of Markov chains where the connexion between reversibility and quasisymmetry is exploited to construct tests for reversibility and tests for homogeneity of several reversible chains.

Some key words: Matched pair; Quasisymmetry; Retrospective study; Reversible Markov chain.

1. INTRODUCTION

The model of quasisymmetry for a square table of counts $\{R_{ij}\}$ ($1 \leq i, j \leq k$) was introduced by Caussinus (1965) as an extension of the model for symmetry. There are a number of equivalent definitions: probably the simplest of these is a log linear model with row, column and symmetry effects. Putting $\mu_{ij} = E(R_{ij})$, we define quasisymmetry by

$$\log \mu_{ij} = \theta_i + \eta_j + \phi_{ij} \quad (1 \leq i, j \leq k) \quad (1.1)$$

with $\phi_{ij} = \phi_{ji}$ and, without loss of generality, one of the main effects, η say, can be incorporated into ϕ . Since its introduction, the model has been used by a number of authors; see Plackett (1974, p. 50), Bishop, Fienberg & Holland (1975, p. 286) and Haberman (1979, Chapter 8). The original motivation for the model was purely for mathematical and computational simplicity: later authors have also stressed the statistical simplicity associated with the log linear form. Inference for the main effect, θ , is commonly based on the conditional distribution of R_{ij} given $R_{ij} + R_{ji}$ ($i < j$). The purpose of the present paper is to motivate the model from a purely scientific viewpoint, i.e. to indicate the kind of experiments or studies likely to lead to tables satisfying (1.1).

2. PROSPECTIVE STUDIES

2.1. Matched pairs

Suppose that individuals are matched in pairs. Independent observations (Y_{1l}, Y_{2l}) ($l = 1, \dots, n$), with the integers $1, \dots, k$ as their common range, are made on each pair. Here, the first entry refers to the control and the second to the treated individual. Primary interest centres on the effect of treatment on the response probabilities. For the control we write the probability of response j in the form

$$\text{pr}(Y_{1l} = j) = \exp(\alpha_{ij}) / \sum_j \exp(\alpha_{ij}) \quad (2.1a)$$

and we suppose that the effect of treatment is to modify these probabilities additively on the log scale

$$\text{pr}(Y_{12} = j) = \exp(\alpha_{1j} + \Delta_j) / \sum_j \exp(\alpha_{1j} + \Delta_j). \quad (2.1b)$$

For a particular pair (2.1) is merely a reformulation of the response probabilities although there is the implication that $\Delta = (\Delta_1, \dots, \Delta_k)$ is a convenient and useful scale for measuring the effect of treatment. The treatment effect, as measured by Δ , is assumed constant for all pairs. Notice that there is some redundancy in the parameterization (2.1). We could, without loss of generality, set $\alpha_{11} = \Delta_1 = 0$ or impose other similar constraints to ensure identifiability. It follows that there are $n(k-1)$ nuisance parameters $\{\alpha_{ij}\}$ and $k-1$ parameters of interest, Δ .

One way of eliminating the nuisance parameters $\{\alpha_{ij}\}$ is to find a statistic whose distribution depends only on Δ . Exact tests and confidence regions can then be constructed, but there is in general no guarantee that the tests so constructed have any strong optimum or near-optimum properties. Nevertheless we proceed by considering all pairs for which the ordered response is either (i, j) or (j, i) ($j \neq i$). The conditional probability that (Y_{11}, Y_{12}) takes the value (i, j) conditional on $(Y_{11}, Y_{12}) = (i, j)$ or (j, i) is readily shown to be $\exp(\Delta_j) / \{\exp(\Delta_i) + \exp(\Delta_j)\}$, the same for each pair. Let n_{ij} be the number of pairs responding (i, j) or (j, i) and let R_{ij} be the number responding (i, j) . It follows that R_{ij} conditional on n_{ij} has the binomial distribution with index n_{ij} and parameter

$$\exp(\Delta_j) / \{\exp(\Delta_i) + \exp(\Delta_j)\}. \quad (2.2)$$

When all types of response are considered simultaneously it is clear that (2.2) implies (1.1) with $\Delta_j = \eta_j - \theta_j$. The likelihood function for Δ conditional on $\{n_{ij}\}$ is the product of $\frac{1}{2}k(k-1)$ independent binomial factors. Likelihood ratio tests and likelihood based confidence regions for Δ are constructed in the usual way.

Essentially the same reference distribution could be obtained by using the full likelihood function and constructing similar regions for Δ by conditioning on the sufficient statistic for the nuisance parameters $\{\alpha_{ij}\}$. Completeness of the conditioning statistic ensures that no exact tests are overlooked, effectively guaranteeing the uniqueness of the reference distribution described above. Details of the argument are given in §2.2 for matched triplets.

The following points concern the applications of the model of quasisymmetry.

(i) Goodness of fit can be tested with $\frac{1}{2}(k-1)(k-2)$ degrees of freedom. If we put $k = 2$ we obtain the model for matched binary pairs (Cox, 1958) for which no goodness-of-fit test exists unless the experiment is replicated.

(ii) Identical pairs are totally ignored and apparently contribute no information concerning the effect of treatment.

(iii) In (1.1) and (2.1) the response categories can be permuted in an arbitrary fashion without materially affecting the model. Thus, if the response categories are ordered no use is made of this information and Δ may not be a convenient measure of the effect of treatment under these conditions.

(iv) The model is not invariant under the groupings of response categories. Alternative models appropriate to ordinal responses and which are invariant under the grouping of adjacent response categories were discussed by McCullagh (1977).

Points (iii) and (iv) above suggest that, in prospective studies, quasisymmetry may be a useful model only if the response categories are on a nominal scale. Most applications of

the model discussed in the statistical literature, however, seem to involve data where the responses are ordinal rather than nominal.

An immediate consequence of the above discussion is that the conditional likelihood score statistic for testing the hypothesis that $\Delta = 0$ is identical to the conditional score statistic for testing symmetry with quasisymmetry as the alternative (Darroch, 1981). Furthermore, since the two likelihood functions are identical and share the same sample space conditionally, it follows that the two score statistics have the same sampling distribution.

2.2. Matched triplets

The natural extension of (2.1) for matched triplets with polytomous response variables (Y_{11}, Y_{12}, Y_{13}) ($l = 1, \dots, n$) is

$$\begin{aligned} \text{pr}(Y_{11} = j) &= \exp(\alpha_{1j} + \Delta_{1j}) / \sum_j \exp(\alpha_{1j'} + \Delta_{1j'}), \\ \text{pr}(Y_{12} = j) &= \exp(\alpha_{1j} + \Delta_{2j}) / \sum_j \exp(\alpha_{1j'} + \Delta_{2j'}), \\ \text{pr}(Y_{13} = j) &= \exp(\alpha_{1j} + \Delta_{3j}) / \sum_j \exp(\alpha_{1j'} + \Delta_{3j'}). \end{aligned} \quad (2.3)$$

All observations are assumed independent. Here, the vectors $\Delta_1, \Delta_2, \Delta_3$ measure the effect of treatments: we could without loss of generality set $\Delta_1 = 0$ and $\alpha_{11} = 0$ for each l . To construct similar regions for the treatment parameters it is convenient to write each response as an indicator vector of length k . Thus the triplet (Y_{11}, Y_{12}, Y_{13}) becomes (Z_{11}, Z_{12}, Z_{13}) , where each Z is of length k . The sufficient statistic for $\{\alpha_{1j}\}$ is $\{Z_l\}$, the vector sum of the responses for each triplet, and the required conditional distribution is that of the data given $\{Z_l\}$. In other words, we require the distribution of (Y_{11}, Y_{12}, Y_{13}) given that (Y_{11}, Y_{12}, Y_{13}) is some permutation of the integers (i, j, s) . The required conditional distribution, where it is not degenerate, is

$$\text{pr}\{(Y_{11}, Y_{12}, Y_{13}) = (i, j, s) | Z_l\} = \exp(\Delta_{1i} + \Delta_{2j} + \Delta_{3s}) / \sum \exp(\Delta_{1i} + \Delta_{2j} + \Delta_{3s}), \quad (2.4)$$

where the sum in the denominator is taken over all distinct permutations of the integers (i, j, s) .

Accordingly, since the conditional probability (2.4) is the same for all triplets, we arrange the data in a three-way table of counts $\{R_{ijs}\}$, where R_{ijs} is the number of triplets responding (i, j, s) . Let n_{ijs} be the sum of R_{ijs} over all distinct permutations of the indices. Then R_{ijs} given n_{ijs} has the binomial distribution with parameter given by (2.4) and

$$\log\{E(R_{ijs} | n_{ijs})\} = \Delta_{1i} + \Delta_{2j} + \Delta_{3s} + \phi_{ijs}, \quad (2.5)$$

where ϕ_{ijs} is symmetric in its indices and depends on n_{ijs} . Thus (2.5) is the natural extension of quasisymmetry appropriate to matched triplets. Estimates of the treatment parameters and of their precision can be found by maximizing either the product multinomial log likelihood function not involving ϕ or by maximizing the Poisson log likelihood function for the log linear model (2.5). These two methods yield identical answers (Palmgren, 1981).

The extension to higher dimensions is immediately clear.

3. MATCHED PAIRS IN RETROSPECTIVE STUDIES

We suppose that for each diseased individual or case, D , there is one matching control, \bar{D} , and that the risk factor J has k levels. For example, J might refer to blood group

(O, A, B, AB). For the moment we assume that J has no structure such as order or factorial arrangement. With the risk factor at level j the probability of disease is written

$$\text{pr}(D|J = j) = \exp(\mu + \Delta_j) / \{1 + \exp(\mu + \Delta_j)\}. \quad (3.1)$$

In other words, with $J = j$ the odds for the disease is increased by the factor $\exp(\Delta_j - \Delta_{j'})$ relative to $J = j'$. This factor is assumed to apply throughout the population under investigation and is called the relative risk or odds ratio of j to j' .

Let π_j be the proportion of individuals in risk category j . Then, for a diseased individual chosen at random the probability of observing risk category j is

$$\left\{ \frac{\pi_j \exp(\mu + \Delta_j)}{1 + \exp(\mu + \Delta_j)} \right\} / \left\{ \sum_{j'} \frac{\pi_{j'} \exp(\mu + \Delta_{j'})}{1 + \exp(\mu + \Delta_{j'})} \right\}. \quad (3.2)$$

For a randomly chosen control the corresponding expression is

$$\left\{ \frac{\pi_j}{1 + \exp(\mu + \Delta_j)} \right\} / \left\{ \sum_{j'} \frac{\pi_{j'}}{1 + \exp(\mu + \Delta_{j'})} \right\}. \quad (3.3)$$

By analogy with the discussion of §3.1 we consider any pair for which the ordered risk categories are either (i, j) or (j, i) , where the first entry refers to the case. The conditional probability that the case has risk category j is readily shown to be $\exp(\Delta_j) / \{\exp(\Delta_i) + \exp(\Delta_j)\}$, independent of $\{\pi_j\}$ and μ and this expression applies to any such pair. Consequently if the data are arranged in a square table of counts $\{R_{ij}\}$, where R_{ij} is the number of matched pairs for which the case has risk category j and the control has risk category i and if (3.1) with the usual assumption of independence holds then the table must be quasisymmetric.

Pike, Casagrande & Smith (1975), without noting the connexion with quasisymmetry, suggest estimating the $k-1$ relative risks subject to mutual consistency. Their maximum likelihood method is exactly equivalent to fitting quasisymmetry using the product-binomial log likelihood function.

Finally we note that the relationship between the probabilities (3.2) and (3.3) is exactly the same as the relationship between (2.1a) and (2.1b): treatment affects the probabilities additively on the log scale. Consequently any optimality or uniqueness properties that apply to the conditional method of §2 apply also to retrospective studies. Again, as in §2, the extension to matched triplets or to higher dimensions is immediately clear.

When the levels of the risk variable have some structure such as order or if they correspond to a factorial arrangement of several factors it is usually necessary to examine the corresponding contrasts of the Δ 's. If, for example, the levels of the risk factor are formed by grouping some quantitative variable, linear and quadratic contrasts of the Δ 's would often be appropriate.

4. REVERSIBLE MARKOV CHAINS

Let Q be a matrix of probabilities satisfying (1.1) with $\sum_j q_{ij} = 1$ for each $i = 1, \dots, k$. We have that $\log q_{ij} = \theta_i + \phi_{ij}$ or

$$q_{ij} = \exp(\phi_{ij}) / \sum_j \exp(\phi_{ij}), \quad (4.1)$$

where $\phi_{ij} = \phi_{ji}$. In matrix notation Q is the product of a diagonal matrix and a symmetric matrix. It follows (Kelly, 1979, p. 10) that Q is the transition matrix of a

reversible Markov chain. Furthermore, any reversible Markov chain has a quasisymmetric transition matrix.

The above result has a number of immediate applications. First, using the Markov property and taking the initial state to be fixed, we can concentrate on the transition counts, which are sufficient. A test for quasisymmetry is equivalent to a test for reversibility assuming the Markov property. The same test for reversibility can be used for a single unbroken chain or for fragments of the same or homogeneous chains, provided that the initial state of each fragment is taken as fixed. The above result follows because the log likelihood function for a Markov chain with fixed initial state is identical to the log likelihood formed by considering each row of the transition matrix as an independent multinomial observation with probabilities satisfying (4.1). The corresponding likelihood ratio or score tests are numerically identical but, because of the difference in the sample spaces, the distributions of these statistics are only equivalent asymptotically. The global test for reversibility has $\frac{1}{2}(k-1)(k-2)$ degrees of freedom. Two-state Markov chains in equilibrium are always reversible.

A second application involves the comparison of two or more reversible Markov chains. For simplicity we take two unbroken chains with transition counts $R_{ij}^{(1)}$ and $R_{ij}^{(2)}$ respectively. Assuming reversibility in each case, suppose we wish to test for homogeneity. In an obvious notation we have $H_0: q_{ij}^{(1)} = q_{ij}^{(2)}$ ($i \leq j$) which is equivalent to $H_0: \phi_{ij}^{(1)} = c\phi_{ij}^{(2)}$ ($i \leq j$), where c is an arbitrary constant. Conceptually the simplest method for testing H_0 is to regard each row of the transition matrices as a multinomial response with fixed total and with probabilities satisfying (4.1). The hypothesis is then tested in the usual way on $\frac{1}{2}k(k+1)-1$ degrees of freedom by fitting two models, the first under H_0 and the second under $H_A: \phi_{ij}^{(1)} \neq \phi_{ij}^{(2)}$.

A second and numerically identical method, but one that is computationally more convenient, is to treat the transition counts for each chain as Poisson observations satisfying the log linear model of quasisymmetry. In contrast to previous applications we are here interested in equality of the symmetry effects, the two main effects being taken as nuisance parameters. Under H_0 , estimates of ϕ can be obtained from the combined table: fitted values are found by using these estimates of ϕ in conjunction with the row totals for each table. The likelihood ratio statistic is found by differencing the goodness of fit statistics in the usual way.

I wish to thank the referees for helpful comments on an earlier version of this paper and in particular the referee whose suggestion led to the work of §4.

REFERENCES

- BISHOP, Y. M. M., FIENBERG, S. E. & HOLLAND P. W. (1975). *Discrete Multivariate Analysis: Theory and Practice*. Cambridge: Massachusetts Institute of Technology.
- CAUSSINUS, H. (1965). Contribution à l'analyse statistique des tableaux de corrélation. *Ann. Fac. Sci. Univ. Toulouse* **29**, 77–182.
- COX, D. R. (1958). Two further applications of a model for binary regression. *Biometrika* **45**, 562–5.
- DARROCH, J. N. (1981). The Mantel–Haenszel test of marginal symmetry; fixed-effects and mixed models for a categorical response. *Int. Statist. Rev.* **49**, 285–307.
- HABERMAN, S. J. (1979). *Analysis of Qualitative Data, 2: New Developments*. New York: Academic Press.
- KELLY, F. P. (1979). *Reversibility and Stochastic Networks*. Chichester: Wiley.
- MCCULLAGH, P. (1977). A logistic model for paired comparisons with ordered categorical data. *Biometrika* **64**, 449–53.
- PALMGREN, J. (1981). The Fisher information matrix for log linear models arguing conditionally on observed explanatory variables. *Biometrika* **68**, 563–6.

- PIKE, M. C., CASAGRANDE, J. & SMITH, P. G. (1975). Statistical analysis of individually matched case-control studies in epidemiology: Factor under study a discrete variable taking multiple values. *Br. J. Prev. Soc. Med.* **29**, 196–201.
- PLACKETT, R. L. (1974). *The Analysis of Categorical Data*. London: Griffin.

[*Received March 1981. Revised October 1981*]