The Annals of Statistics
1999, Vol. 27, No. 1, 405–413

EXPLORING INTERACTIONS IN HIGH-DIMENSIONAL TABLES: A BOOTSTRAP ALTERNATIVE TO LOG-LINEAR MODELS[1]

BY BERND STREITBERG

University of Hamburg

Based on a revised Lancaster-type representation of the additive interactions associated with a probability measure, a new approach for the analysis of high-dimensional contingency tables is proposed. The approach is essentially model-free because the additive interaction tensor is merely a convenient reparameterization of the given table. Single interaction terms are investigated using the bootstrap method whose first-order asymptotic validity is immediate. The global structure can be investigated by using the multiple p-values given by Holm's sequentially rejecting multiple testing procedure. The procedure is based on a characterization of the Moebius function as a solution of the simultaneous eigenproblem for all intersection operators in a finite lattice.

1. Introduction: multiplicative and additive interactions. A fundamental notion in the analysis of multiway contingency tables is the concept of *interaction* between more than two variates. There are two basic approaches: the *multiplicative* definition, which originated in Bartlett (1935), and the *additive* approach, which was proposed by Lancaster (1969), although special cases had already been discussed by Lazarsfeld (1961) and Bahadur (1961) in their algebra of dichotomous systems. The multiplicative approach is now firmly embedded within the framework of hierarchical log-linear models: in an m-way table there is *no multiplicative m-variable interaction* iff a proper submodel of the saturated model holds. The additive approach has been investigated further in a few isolated papers [e.g., Zentgraf (1975), Toewe, Bock and Kundt (1985), Darroch and Speed (1983)] but it has not been developed into a consistent method for the analysis of multiway tables. It appears, therefore, largely to have been forgotten.

There are several problems connected with the application of log-linear models.

Received June 1991.

[1]This paper was tentatively accepted on June 25, 1991. Bernd Streitberg died on September 28, 1991 at the age of 44. He was not able to complete a final revision. This version was found by his friends among his files. We publish it here in his memory although it may not include all changes he would have liked to make.

AMS 1991 *subject classifications*. Primary 62H17; secondary 62E30, 62H20, 62G09, 62J15, 05A18, 06A07.

Key words and phrases. Additive interactions, cumulants, contingency tables, bootstrap. Moebius function, multiple comparisons.

1. The presence of a multiplicative interaction between given variables depends not only on the distribution of these variables but also on all other variables considered, unless a rather strict collapsibility condition holds.
2. Condensation of a table without m-way interaction by combining categories can produce an m-way interaction.
3. If confidence intervals and tests are, as it is usually the case, derived via asymptotic expansions, the validity of these approximations for high-dimensional tables in small samples, where empty cells are likely to occur, is often questionable.
4. Essentially log-linear modeling is a model search procedure. It is not clear whether distribution properties of tests and estimators still hold when this fact is taken into account.

I therefore propose to take a fresh look at additive interactions. To begin with, the original definition of additive interactions as given by Lancaster and the other authors can be shown to have a property that is rather unfortunate for an interaction measure: additive interactions do not necessarily vanish even for stochastically independent variables. This becomes apparent once one applies Lancaster's definition to four and higher dimensional tables; the additive interactions in an $ABCD$ table are equal to the products of the AB and the CD interactions if AB is stochastically independent from CD. Fortunately, the definition can be corrected, using a characterization theorem for the Moebius function in finite lattices. This has been done in Streitberg (1990). The resulting revised additive interactions are equal to the Lancaster parameters for $m \leq 3$ and are related to multivariate cumulants.

Second, the theoretical interactions can be estimated from a given empirical table without any model assumptions at all other than the basic multinomial sampling assumption. This is the only sampling model considered here, but the extension to more complicated models like the product-multinomial is straightforward. In practical work, it is highly important to draw a distinction between responses, covariates and explanatory variables. Here merely the pure multivariate situation is covered: all variables are considered as responses and the analysis of their joint distribution is of interest. Extensions to other situations are possible but will not be discussed here.

For estimation, either Fisher's k-statistics can be used or, more simply, the same functional as for theoretical distributions. In an explorative approach, all interactions can be tested by referring them to their asymptotic normal distribution. For finite samples, the bootstrap procedure either in its original form or in a suitably defined more efficient form [Efron (1990)] is proposed. The asymptotic validity of the bootstrap here is immediate from a remark by Efron (1979).

2. Map notation for tables and basic definitions. A contingency table is a measure on a product $I = I_1 \times I_2 \times \cdots \times I_m$ of finite sets I_k, $k = 1, 2, \ldots, m$. This is usually given by a density (with respect to counting measure) f: $I \to [0, 1]$ with $f(i) \geq 0$ for all $i \in I$ and $\sum_{i \in I} f(i) = 1$. If n iid

observations are drawn from f, they generate a corresponding empirical measure given by the density $\hat{f}$: $I \to [0,1]$ with $\hat{f}(i)$ equal to the relative frequency of observations in cell i. Obviously $\hat{f}$ is distributed as a scaled multinomial, that is, $n\hat{f} \sim Mult[n, f]$. Let us write $\mathbf{m}$ for the set $\{1, 2, \ldots, m\}$ of variables; then it is convenient to view the set I of cells as the set of all mappings i: $\mathbf{m} \to \bigcup_{k \in \mathbf{m}} I_k$ with $i_k = i(k) \in I_k$. For a table f and a subset $M \subseteq \mathbf{m}$ of the variables the marginal distribution f_M is concentrated on I_M, the set of restricted maps $i|M$: $M \to \bigcup_{k \in M} I_k$ with $(i|M)(k) = i(k)$ for $k \in M$. One has $f_M(j) = \sum_{i:\, i|M = j} f(i)$. The *map notation* also allows a concise definition of tensor products. For $M \in \mathbf{m}$, $N \in \mathbf{m}$ with $M \cap N = \varnothing$ the marginals f_M, f_N can be used to build a product density on $I_{M \cup N}$ by letting, for $i \in I_{M \cup N}$,

$$(f_M \otimes f_N)(i) = f_M(i|M) f_N(i|N).$$

Note that this is simply the product distribution for variables $M \cup N$ which would have been obtained if the variables in M were independent from the variables in N. If, furthermore, $\pi = \{M_1, M_2, \ldots, M_r\}$ is a set partition of $\mathbf{m}$, that is, a set of nonempty, pairwise disjoint subsets of $\mathbf{m}$, called the blocks of π, such that the union of the blocks is equal to $\mathbf{m}$, we let

$$f_\pi = \bigotimes_{M \in \pi} f_M = f_{M_1} \otimes \cdots \otimes f_{M_r},$$

where f_π is a product density on the original set I of cells, obtained from the specified marginals. This is well defined, because $\otimes$ is clearly associative and commutative (sic!) by virtue of the map notation.

Finally, consider the set Π of all set partitions of $\mathbf{m}$. If $m = 4$, for instance, Π has 15 elements. Then Π can be (partially) ordered by refinement, where we write $\pi \leq \pi'$ if each block of π is contained in a single block of π'. For instance $(12, 3, 4) \leq (12, 34)$, which is shorthand for $\{\{1, 2\}, \{3\}, \{4\}\} \leq \{\{1, 2\}, \{3, 4\}\}$. Moreover, for two partitions $\beta, \gamma \in \Pi$ there exists an unique infimum $\beta \wedge \gamma$ and supremum $\beta \vee \gamma$. For example $(12, 34) \wedge (123, 4) = (12, 3, 4)$ and $(1, 2, 34) \vee (13, 2, 4) = (2, 134)$. This means that Π is a lattice. For details see Streitberg (1990); note especially that we talk about *set* partitions and not about *number* partitions, where a canonical lattice structure does not exist. The smallest element of Π is denoted by α (e.g., $\alpha = 1, 2, 3, 4$) and the largest element by ω (e.g., $\omega = 1234$). Here f_α corresponds to the model of complete independence between all variables and f_w is equal to f.

The following lemma is simple—the proof can be safely left to the reader—but fundamental. Assume that f is itself a product density, say $f = f_\pi$, where π is a partition of $\mathbf{m}$. This means that f can be decomposed into a tensor product as in the formula above. What do the product densities of such an f look like?

LEMMA 1. *Product densities of a decomposable density*: *if* $f = f_\pi$, *the product density* f_β *for* $\beta \in \Pi$ *is given by* $(f_\pi)_\beta = f_{\beta \wedge \pi}$.

The proof is left as an exercise. Note the utility of the partition notation.

3. The definition of additive interactions. Additive interactions are an alternative parameterization of f meant to show clearly the dependencies between the given variates. The fundamental idea, which is due to Lancaster, can already be seen in the case $m = 2$. Here we compare the original table f with the table $f_{1,2}$ which would have been obtained under independence of the two variates. It is very natural to consider the "residuals" $\kappa = f - f_{1,2}$. Note that κ is again a table of the same size as f, which has in cell (a, b) the difference $f(a, b) - f_1(a)f_2(b)$. We shall not discuss χ^2-statistics based on these residuals, because the χ^2-approach cannot be generalized to the m-dimensional case, unless a restrictive "complete independence" assumption is made. Instead we focus on κ itself and observe three simple properties, which can be used for a general definition of additive interactions for the m-way case:

DEFINITION 1. Additive interactions: the table $\kappa = T(f)$ of additive interactions corresponding to a given distribution f on I is a table of the same dimension as f, that is, $T(f)$: $I \to \mathscr{R}$, where $\mathscr{R}$ is the set of real numbers, with the following three properties.

A1. Additivity axiom: κ is a linear combination of the product densities f_π derived from f.
A2. Normalization axiom: the coefficient of the original table $f_\omega = f$ in this linear combination is unity.
A3. Interaction axiom: if f is equal to a proper product density, that is, $f = f_\pi$ for $\pi < \omega$, κ is identically equal to zero.

The central axiom is (A3), stating a minimal property of any "decent" interaction measure: it should vanish whenever the variables are completely or groupwise independent. Axiom (A1) captures the idea of additivity, that is, linearity in the product densities and (A2) is a condition put forward in order to prevent a trivial form of nonuniqueness. It is astonishing that these three simple properties already constitute a valid definition: κ exists and is unique. Before we consider the general case, let us work out the definition of κ for the case $m = 3$. From (A1), one finds that

$$\kappa = c_1 f + c_2 f_{1,23} + c_3 f_{2,13} + c_4 f_{3,12} + c_5 f_{1,2,3},$$

where the five coefficients are to be determined using Lemma 1. If one puts $f = f_{1,2,3}$, one has by (A3) $0 = c_1 + c_2 + c_3 + c_4 + c_5$ and three more equations are obtained by letting f equal one of the other proper product tables. Together with (A2) these equations determine κ uniquely and one finds for the entry of κ in cell (a, b, c),

$$\begin{aligned}\kappa(a, b, c) = {} & f(a, b, c) - f_1(a)f_{23}(b, c) - f_2(b)f_{13}(a, c) \\ & - f_3(c)f_{12}(a, b) + 2f_1(a)f_2(b)f_3(c).\end{aligned}$$

This equation has been written in conventional "cellwise" notation, and not in the "vector" notation used in the definition and the following general propositions.

THEOREM 1. *Existence and uniqueness of additive interactions*: κ, *as defined by* (A1), (A2) *and* (A3), *exists and is uniquely given by*

$$\kappa = \sum_{\pi \in \Pi} (-1)^{|\pi|-1}(|\pi| - 1)! f_\pi,$$

where $|\pi|$ *denotes the number of blocks in a partition* π.

PROOF. The proof follows from a characterization of the Moebius function in arbitrary finite lattices. A standard lattice-type proof, proceeding via induction, has been given in Streitberg (1990). In the Appendix of this paper an alternative proof is given with the hope of provoking more insight by showing that the theorem is essentially the solution of an eigenvalue problem for a class of matrices in the incidence algebra on Π.

Theorem 1 relates additive interactions to multivariate cumulants. While the general formula for multivariate cumulants is well known [Speed (1983)] and obviously is analogous to the formula for additive interactions given above, a deeper reason for the relationship has been given in Streitberg (1990). There, an axiomatic definition for cumulants has been proposed and it has been shown that for each m-dimensional random vector in the Banach space L^m, the multivariate cumulant is equal to the Lebesgue integral of the product of these random variables with respect to a certain signed measure. For the discrete case this signed measure is given by the additive interaction measure of Definition 1. Here we will express this relationship in a simpler language. If f is a distribution on the set I of cells, we can, with a certain abuse of notation, regard I as a random vector which takes the value i with probability $f(i)$. Now consider a fixed, but arbitrary cell c of the table. Given c, we associate to I a binary random vector I^c with components $I_1^{c_1}, \ldots, I_m^{c_m}$, where $I_k^{c_k} = 1$ iff $I_k = c_k$ for $k \in \mathbf{m}$. For example, let $m = 4$ in a 3^4-table and assume that I takes the value $(2, 3, 1, 2)$. Then for $c = (3, 3, 2, 2)$ one finds that I^c takes the value $(0, 1, 0, 1)$. The distribution of I^c is obtained from the distribution of I by dichotomizing the categories of every variable k into two classes "c_k" and "not c_k."

COROLLARY 1. *Cumulants and additive interactions*: *the entry of* κ *in any fixed cell* c *is equal to the* m*-variate cumulant* $\operatorname{cum}(I^c)$.

PROOF. Fix a cell c and specialize the formula in Theorem 1 to this cell by considering $\kappa(c)$. The multivariate cumulant of an m-dimensional random vector X is given by the formula in Theorem 1, if one could have $f_\pi(c)$ equal to the moment product $\prod_{M \in \pi} E[\prod_{k \in M} X_k]$. This is achieved by defining $X = I^c$.

As an example, let $c = (a, b)$. Then $\kappa(a, b) = \operatorname{cov}[I_1^a, I_2^b]$. Note that, for $m > 1$, the sum of κ over any coordinate is equal to 0, because a degenerate random variable is independent from all other variates. For a subset $M \subseteq \mathbf{m}$, we let the marginal additive interactions κ_M equal the additive interactions of the marginal table f_M and for a partition π we define the interaction

product κ_π as the (tensor) product of the κ_M, $M \in \pi$. For example, $\kappa_{1,2}(a, b) = \kappa_1(a)\kappa_2(b) = E[I_1^a]E[I_2^b] = f_1(a)f_2(b)$. The sum of the interaction products is equal to f; for example, $f = \kappa_{12} + \kappa_{1,2}$. This is true for general m-way tables, giving for the revised interactions the analogue of an expansion derived by Bahadur (1961) for the Lazarsfeld–Lancaster interactions. In the original expansion, product terms like $\kappa_{12}\kappa_{34}$ were absent.

COROLLARY 2. *Revised Bahadur expansion*: *the original f can be decomposed into a sum over all interaction products*

$$f = \sum_{\pi \in \Pi} \kappa_\pi .$$

Furthermore, for any $\beta \in \Pi$, $f_\beta = \sum_{\pi \le \beta} \kappa_\pi$.

The proof is immediate by Moebius inversion.

This proposition gives a way of understanding high-dimensional interactions. The three-way interaction of a three-way table, for instance, is the residual that is left over after all lower dimensional interactions are accounted for. Expansions like the above are often used for an approximation of f by omitting higher order terms. As in the related Edgeworth expansion, this is not always a good idea, because the truncated expansion is not necessarily a valid density (i.e., nonnegative). The problem of obtaining, say, *ML*-estimators for the low-order terms under a given "truncation" model is unsolved both for the original and the revised interactions. In the next section I propose instead a "model-free" application of the Bahadur expansion.

The correspondence *additive interactions* ⇔ *cumulants* is very fruitful because concepts and theorems from one of the fields can be freely transferred to the other. A few hints might be sufficient.

First, a theory of conditional interactions can be derived from known results about conditional cumulants, for example, MacCullagh (1987).

Second, the theory of generalized cumulants by Speed (1986a) shows how the interactions of the m-way table f are related to the interactions of m'-way tables obtained by regarding groups of variables as single variables. The further generalizations to polypolykays by Carney (1968), Speed (1986b) and others point the way toward an interaction approach for more complicated cell structures than full Cartesian products.

Finally, the other direction of the correspondence, viewing the conventional cumulants as functionals of interaction measures, gives a possibility for robustifying these creatures, after unveiling some of their mystery.

> *Even today, the statistical theory of cumulants wears a halo of mystery that we still are a long way from dispelling. We do not hesitate to predict that cumulants and perpetuants will soon be inserted in the mainstream of mathematics.* [Rota (1986)]

There are, however, some open questions. One problem is the expression of κ for 2^m-tables as a "nice" homogeneous polynomial of order m in the cell

entries $f(i)$, $i \in I$. For instance, if $I_k = \{1, 2\}$ for $k = 1, 2$, one finds $\kappa(1, 1) = f(1, 1)f(2, 2) - f(1, 2)f(2, 1)$. For a 2^4-table, κ is a polynomial of degree 4 in the 16 cell entries with more than 1000 terms.

Interactions correspond to multidimensional covariances. In order to obtain parameters corresponding to multidimensional correlations, one would be interested in inequalities for interactions. This is essentially unsolved. Arguing from the cases $m = 2$ or $m = 3$, one might conjecture the following: if $i \in I$ is an arbitrary cell of an m-way table I with $m > 1$, one has $|\kappa(i)| \leq \frac{1}{2}^m$. This conjecture is, however, already wrong for $m = 4$—consider a 2^m-table with only two nonzero entries of $\frac{1}{2}$ at $(0, 0, \ldots 0)$ and $(1, 1, \ldots 1)$. Then for $m > 1$: $|\kappa| = (2^m - 1)|B_m|/m$, where B_m is the mth Bernoulli number.

4. A bootstrap approach. If $\kappa = T(f)$ is the interaction table of a distribution f, we use the same function $\hat{\kappa} = T(\hat{f})$ of $\hat{f}$ as an estimator of κ. Note that this estimator is not unbiased and that one could use k-statistics if an unbiased estimator is desired. For instance, in a two-way table, an unbiased estimator of $\kappa(a, b)$ is given by $((n - 1)/n)\hat{f}(a, b) - \hat{f}(a)\hat{f}(b)$. Here T is a "civilized" statistic in the sense of the δ-method, that is, for $n\hat{f} \sim \text{Mult}[n, f]$, $\hat{\kappa}$ is asymptotically multivariate normal. However, T is an algebraically complicated function of $\hat{f}$ and the derivation of the asymptotic covariance matrix requires efficient computer algebra systems. Such a system for formal tensor algebra has been written by the author (*TENSOR*, written in APL2), but the resulting formulas are by no means simple. Therefore it is proposed to bootstrap both the additive interactions and the terms in the Bahadur expansion, that is, the decompositions of the individual cell entries into interaction product components. The bootstrap samples are tables f^*, generated from $\hat{f}$ by multinomial sampling, that is, conditionally on $\hat{f}$ one has $nf^* \sim Mult[n, \hat{f}]$. The first-order asymptotic validity of using the empirical distribution of $T(f^*)$ is immediate because in the strong sense $\hat{f} \to f$ and also, conditional on $\hat{f}$, $f^* \to \hat{f}$. Therefore, $T(f^*)$, conditionally on $\hat{f}$, and $T(\hat{f})$ have the same asymptotic distribution.

Tests and confidence intervals for all interaction statistics can be obtained in several ways:

1. use "t-tests" with conventional bootstrap estimators for the variances;
2. use the conventional bootstrap quantiles;
3. use more efficient variance or quantile estimators [Efron (1990)].

A systematic study of the different approaches has not been done. For the more efficient estimators a (n, n)-matrix has to be inverted, if Efron's formulas are used. By assuming cellwise constant Hajek bootstrap projections, this can be reduced to the inversion of a (p, p)-matrix where $p = |I|$ is the number of cells in the table.

For a 2^m-table, each marginal distribution is characterized by a single interaction statistic. By using any of the bootstrap approaches, an ANOVA-type interaction table is obtained, where p-values for the interaction statis-

tics are determined by their bootstrap distribution. Because many tests are conducted, one would consider a multiple testing procedure, say by using Holm–Bonferroni multiple p-values. For the theory of multiple testing, compare the fundamental paper by Sonnemann (1982) and the papers in Bauer, Hommel and Sonnemann (1987). The multiple p-values for r given p-values $p_1, p_2, \ldots, p_r$ are defined by $q_1, q_2, \ldots, q_r$ where

$$q_i = \min\Big(1, \max_{j:\, p_j \le p_i} (r - r_j + 1)p_j\Big).$$

Here $r_j \in \{1, 2, \ldots, n\}$ is the rank of p_j among $p_1, p_2, \ldots, p_r$. These are the p-values corresponding to the sequentially rejecting Holm procedure [Holm (1979)]. The bootstrap procedure can also be applied to the cell decompositions, that is, the Bahadur expansion of $\hat{f}$. The method allows a detailed exploration of the interactions which are present in a multidimensional frequency table.

APPENDIX

PROOF OF THEOREM 1. We use Iverson's (1962) convention throughout: if $\mathscr{P}$ is a proposition, $(\mathscr{P})$ is equal to 1 if $\mathscr{P}$ is true and equal to 0 if it is false. See also Graham, Knuth and Patashnik (1989). From Lemma 1 and Definition 1, one has $\kappa = \sum c_\pi f_\pi$ with $c_\omega = 1$ and $\sum_{\pi:\, \beta = \pi \wedge \tau} c_\pi = 0$ for all $\beta, \tau < \omega$. Define the zeta matrix $Z = (z_{\beta\gamma})$ by $z_{\beta\gamma} = (\beta \le \gamma)$ and, for $\tau \in \Pi$, intersection matrices $K_\tau = (k^\tau_{\beta\gamma})$ by $k^\tau_{\beta\gamma} = (\beta = \gamma \wedge \tau)$. Because the elements of Π can be arranged such that Z is a triangular matrix with unity along the main diagonal, the inverse $M = (m_{\beta\gamma}) = Z^{-1}$ exists and is called the Moebius function of Π. By Rota (1964), $m_{\pi\omega} = (-1)^{|\pi|-1}(|\pi| - 1)!$. The Moebius function solves the eigenproblem of all intersection matrices simultaneously: $ZK_\tau M = D_\tau$ where $D_\tau = (d^\tau_{\beta\gamma})$ is diagonal with $d^\tau_{\beta\beta} = (\beta \le \tau)$, because $\sum_\gamma (\beta \le \gamma)(\gamma = \pi \wedge \tau) = (\beta \le \pi \wedge \tau) = (\beta \le \tau)(\beta \le \pi)$. Now let $c = (c_\pi)$ and $e = (e_\pi)$ with $e_\pi = (\pi = \omega)$, then from $K_\tau c = 0$ for all $\tau < \omega$ it follows together with $c_\omega = 1$ that $c = Z^{-1}e$, that is, $c_\pi = m_{\pi\omega}$. □

For the concept of Moebius inversion, see Moebius (1832), Rota (1964), Aigner (1975) or Stanley (1989).

REFERENCES

AIGNER, M. (1975). *Kombinatorik I. Grundlagen und Zähltheorie*. Berlin.

BAHADUR, R. R. (1961). A representation of the joint distribution of responses to n dichotomous items. In *Studies in Item Analysis and Prediction* (H. Solomon, ed.) 158–176. Stanford.

BARTLETT, M. S. (1935). Contingency table interactions. *J. Roy. Statist. Soc. Suppl.* **2** 248–252.

BAUER, P., HOMMEL, G. and SONNEMANN, E., eds. (1987). *Multiple Hypotheses Testing*. Springer, Berlin.

CARNEY, E. J. (1968). Relationship of generalized polykays to unrestricted sums for balanced complete finite populations. *Ann. Math. Statist.* **39** 643–656.

DARROCH, J. N. and SPEED, T. P. (1983). Additive and multiplicative models and interactions. *Ann. Statist.* **11** 724–738.

EFRON, B. (1979). Bootstrap methods. *Ann. Statist.* **7** 1–26.
EFRON, B. (1990). More efficient bootstrap computations. *J. Amer. Statist. Assoc.* **85** 79–89.
GRAHAM, R. L., KNUTH, D. E. and PATASHNIK, O. (1989). *Concrete Mathematics. A Foundation for Computer Science*. Addison-Wesley, Reading, MA.
HOLM, S. (1979). A simple sequentially rejective multiple test procedure. *Scand. J. Statist.* **6** 65–70.
IVERSON, K. E. (1962). *A Programming Language*. New York.
LANCASTER, H. O. (1969). *The Chi-Squared Distribution*. London.
LAZARSFELD, P. F. (1961). The algebra of dichotomous systems. In *Studies in Item Analysis and Prediction* (H. Solomon, ed.) 111–157. Stanford.
MACCULLAGH, P. (1988). *Tensor Methods in Statistics*. Chapman and Hall, London.
MOEBIUS, A. F. (1832). Über eine besondere Art der Umkehrung der Reihen. *J. Reine Angew. Math.* **9** 105–132.
ROTA, G. C. (1964). On the foundations of combinatorial theory I. Theory of Moebius functions. *Z. Wahrsch. Verw. Gebiete* **2** 340–368.
ROTA, G. C. (1986). Introduction. In *Percy Alexander MacMahon*: *Collected Papers II*: *Number Theory*, *Invariants and Applications*. Cambridge.
SONNEMANN, E. (1982). Allgemeine Lösungen multipler Testprobleme. *EDV in Medizin und Biologie* **13** 120–128.
SPEED, T. P. (1983). Cumulants and partition lattices. *Austral. J. Statist.* **25** 378–388.
SPEED, T. P. (1986a). Cumulants and partition lattices II: generalized k-statistics. *J. Austral. Math. Soc. Ser A* **40** 34–53.
SPEED, T. P. (1986b). Cumulants and partition lattices III: multiply indexed arrays. *J. Austral. Math. Soc. Ser. A* **40** 161–182.
STANLEY, P. (1989). *Enumerative Combinatorics* **1**. Monterey, CA.
STREITBERG, B. (1990). Lancaster interactions revisited. *Ann. Statist.* **18** 1878–1885.
TOEWE, J., BOCK, J. and KUNDT, G. (1985). Interactions in contingency table analysis. *Biometrikal J.* **27** 1–7.
ZENTGRAF, R. (1975). A note on Lancaster's definition of higher order interactions. *Biometrika* **62** 375–378.

CORRESPONDENCE TO:
GÜNTHER SAWITZKI
INSTITUT FÜR ANGEWANDTE MATHEMATIK
IM NEUENHEIMER FELD 294
D-69120 HEIDELBERG
GERMANY
E-MAIL: gs@statlab.uni-heidelberg.de