

CUMULANTS AND PARTITION LATTICES IV: A.S. CONVERGENCE OF GENERALISED k -STATISTICS

T. P. SPEED

(Received 24 August 1984)

Communicated by T. C. Brown

Abstract

Generalised k -statistics associated with multi-indexed arrays of random variables satisfying a generalised form of exchangeability are studied. By showing that they form multi-indexed reversed martingales and that the associated family of σ -fields possesses certain conditional independence properties, conditions for the a.s. convergence of generalised k -statistics are obtained. When the arrays of random variables are sums of independent arrays of independent effects, as is the case with the standard random effects anova models, the limits are identified as the associated generalised cumulants.

1980 *Mathematics subject classification* (*Amer. Math. Soc.*): 60 G 42, 62 F 12, 62 J 10.

Keywords and phrases: cumulant, k -statistic, partial exchangeability, anova models, variance component estimation, asymptotics.

1. Introduction

In the previous paper in this series we gave generalisations of the classical cumulants of Thiele and k -statistics of Fisher to certain multi-indexed arrays of random variables. Amongst other things we showed that when the arrays were sums of independent arrays of independent effects, as is the case with the standard random effects anova models, our generalised cumulants reduced to cumulants and to the products of cumulants of those effects, and our generalised k -statistics were their best unbiased estimators within a natural class of estimators. When this discussion was specialised to order two, we found that our generalised cumulants were just the usual components of variance in anova

models, whilst the generalised k -statistics were their natural unbiased estimators arising from the anova decomposition of the total sum of squares.

The main aim of the later papers in this series is the exploration of complex analyses of variance under assumptions wider than the usual ones of normality and additivity. As explained in the previous paper in the series, this work was begun independently by Kempthorne (1952) and Cornfield and Tukey (1956), and continued over a number of years by Kempthorne and Tukey and their students. However, neither of these lines of development included any discussion of the asymptotics of variance component estimation, and in a generalised sense this is the main topic of the current paper. We show in Section 4 below that under a natural moment condition, all of the usual unbiased estimators of components of variance converge a.s. under the most general possible assumptions, namely those which are clearly necessary for the expressions involved to make sense. Our broad framework is similar to that adopted by Miller (1973, 1977) with three vital differences: we only make symmetry assumptions on the distributions of our arrays, whereas he assumes normality; we have no fixed effects present, that is, the means of our array elements are all constant whereas his discussion is really for the general mixed model; and our estimates of components of variance are the standard unbiased ones, whereas he uses the maximum likelihood estimates, which are biased (although asymptotically unbiased) even in the absence of fixed effects. In addition our results for variance components are but a special case of similar ones for generalised k -statistics. It would naturally be of interest to extend the techniques used below to allow the inclusion of structured mean values, and work towards this end is in progress.

We turn now to a brief description of the approach of this paper. In essence the main argument is simple and well known; it is the task of setting up the framework and filling in the details which seems rather formidable. Averages of random variables whose joint distribution satisfies certain symmetry constraints frequently give rise to reversed martingales, see for example Meyer (1966, page 151) or paper II in this series, and these converge a.s. without further moment conditions. Dubins and Pitman (1979) show that similar conclusions can be derived with multi-indexed averages only when extra moment conditions and a permutability or conditional independence assumption hold. In our case the main work needed to show this is not probabilistic but algebraic: defining the groups which embody our generalised exchangeability, identifying the relevant subgroups and describing their orbits for the averaging, defining the associated sub- σ -fields, and deriving the necessary conditional independences. Much of this has been done elsewhere, and once it is all organised the main result follows quickly.

The paper is arranged as follows. We begin with a discussion of the two simplest cases: doubly-indexed arrays where the two indices are crossed and nested, respectively. With simple crossed arrays, our assumptions correspond to

row-column exchangeability, Aldous (1981, 1984), and we explain the relation of our results to those concerning the standard two-way random effects models. When the two indices are nested, we dub our assumption *between/within exchangeability*, and in that context illustrate our results with a third and fourth-order generalised k -statistic. Following these simple examples we turn in Section 3 to organising the preliminary algebraic results mentioned above; Section 4 is then devoted to the statement and proof of the main result and the derivation of some of its more interesting corollaries.

2. Two simple examples

Let us suppose that $X = (X_{ij} : i \in I, j \in J)$ is a doubly-indexed array of random variables defined on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ whose joint distribution is invariant under independent transformation of i and j by permutations of finite support, i.e.

$$(2.1) \quad (X_{ij} : i \in I, j \in J) \stackrel{\mathcal{D}}{=} (X_{igjh} : i \in I, j \in J)$$

where g and h are finitary permutations of I and J respectively, and where ig [respectively jh] denotes the image of i [respectively j] under g [respectively h]. In all that follows $I = J = \{1, 2, 3, \dots\}$, and the group of finitary permutations of I and J will be denoted by S_∞ , the restricted infinite symmetric group. Aldous (1981, 1984) gives many interesting results concerning such arrays, which he calls row-column exchangeable (RCE), but makes no explicit use of doubly-indexed reversed martingales (RMGs) in his work. We illustrate our later results by deriving some simple facts concerning RCE arrays using such RMGs.

Fix $m \geq 2$ and $n \geq 2$, let S_m [respectively S_n] denote the subgroups of S_∞ acting on I [respectively J] which leave all indices $i > m$ [respectively $j > n$] fixed, and write $G_{m,n} = S_m \times S_n$. Associated with $G_{m,n}$ is the sub- σ -field $\mathcal{G}_{m,n}$ of all events determined by (X_{ij}) which are invariant under $G_{m,n}$; for example, $\sum_1^m \sum_1^n X_{ij}$ is $\mathcal{G}_{m,n}$ -measurable. Standard arguments show that if $\mathbb{E}|X_{11}| < \infty$, then

$$(2.2) \quad \mathbb{E}\{X_{11} | \mathcal{G}_{m,n}\} = \frac{1}{mn} \sum_1^m \sum_1^n X_{ij} = X_{..}^{(m,n)}.$$

Since the $\mathcal{G}_{m,n}$ decrease as m, n increase, one would hope to conclude that $X_{..}^{(m,n)}$ converges a.s. and in L_1 to a limit; indeed this turns out to be the case, provided simply that $\mathbb{E}|X_{11}| \log^+ |X_{11}| < \infty$, but for less obvious reasons. As Dubins and Pitman (1979) show, the conclusion of a.s. convergence requires, in addition to the moment condition, the conditional independence of $\mathcal{G}_{m,1}$ and $\mathcal{G}_{1,n}$ given $\mathcal{G}_{m,n} = \mathcal{G}_{m,1} \cap \mathcal{G}_{1,n}$, and this is equivalent to requiring the permutability $G_{m,1}G_{1,n} = G_{1,n}G_{m,1}$ of the subgroups $G_{m,1}$ and $G_{1,n}$ of $S_\infty \times S_\infty$. It is easy to see that the

permutability condition is true in this case, and so the doubly-indexed RMG $(X_{..}^{(m,n)}; \mathcal{G}_{m,n})$ does converge a.s. (under the required moment condition). So also will the other RMGs which we shortly define.

Let us now consider some generalised k -statistics of order 2. (See III for a general discussion.) The following are proved in the same way as (2.2) by averaging over the orbits of $G_{m,n}$ and using the property of (RCE):

$$(2.3a) \quad \mathbb{E}\{X_{11}^2 | \mathcal{G}_{m,n}\} = \frac{1}{mn} \sum_i \sum_j X_{ij}^2,$$

$$(2.3b) \quad \mathbb{E}\{X_{11}X_{12} | \mathcal{G}_{m,n}\} = \frac{1}{mn(n-1)} \sum_i \sum_{j \neq j'} X_{ij}X_{ij'},$$

$$(2.3c) \quad \mathbb{E}\{X_{11}X_{21} | \mathcal{G}_{m,n}\} = \frac{1}{m(m-1)n} \sum_{i \neq i'} \sum_j X_{ij}X_{i'j},$$

$$(2.3d) \quad \mathbb{E}\{X_{11}X_{22} | \mathcal{G}_{m,n}\} = \frac{1}{m(m-1)n(n-1)} \sum_{i \neq i'} \sum_{j \neq j'} X_{ij}X_{i'j'},$$

and it follows from results in III (or is easily proved directly) that

$$(2.4) \quad \mathbb{E}\{X_{11}(X_{11} - X_{12} - X_{21} + X_{22}) | \mathcal{G}_{m,n}\} \\ = \frac{1}{(m-1)(n-1)} \sum_i \sum_j (X_{ij} - X_{i.} - X_{.j} - X_{..})^2,$$

where $X_{..} = X_{..}^{(m,n)}$, $X_{i.} = \frac{1}{n} \sum_j X_{ij}$ etc., and where all sums are taken over $i = 1, \dots, m$ and $j = 1, \dots, n$. The expression on the right-hand side of (2.4) is the familiar *rows.columns* or *interaction* mean-square associated with the anova of $(X_{ij}; i = 1, \dots, m; j = 1, \dots, n)$, and we denote it by $MS_{r,c}^{(m,n)}$. If $\mathbb{E}\{|X_{11}X_{ij}| \log^+ |X_{11}X_{ij}|\} < \infty$ for $i = 1, 2, j = 1, 2$, then it follows by what has already been stated that $MS_{r,c}^{(m,n)}$ converges a.s. as m, n increase.

It is possible to prove that, provided the relevant moments exist, we have

$$\text{Var}\{MS_{r,c}^{(m,n)}\} = A + O\left(\frac{1}{m}\right) + O\left(\frac{1}{n}\right), \quad m, n \rightarrow \infty,$$

where $A = f(2^2, 2^2) - [f(2, 2)]^2$ is a function of the generalised cumulants $f(2^2, 2^2)$ and $f(2, 2)$; the details will be given later. Here we simply note that if the array (X_{ij}) is given by a general additive model

$$(2.5) \quad X_{ij} = \mu + r_i + c_j + (r.c)_{ij},$$

where $\mu, (r_i), (c_j)$ and $((r.c)_{ij})$ are independent sets of independent effects with variances $\sigma_\mu^2, \sigma_r^2, \sigma_c^2$ and $\sigma_{r,c}^2$ respectively, then

$$(2.6) \quad \lim_{m,n} MS_{r,c}^{(m,n)} = \sigma_{r,c}^2 \quad \text{a.s.}$$

There are two ways to see that (2.6) follows from (2.5). The first is to prove that the expression A in the formula for $\text{Var}\{\text{MS}_{ir.c}^{(m,n)}\}$ vanishes under (2.5), whilst the second uses a more familiar argument to prove that $\cap\{\mathcal{G}_{m,n}: m \geq 1, n \geq 1\}$ is trivial when the same assumption is supplemented by $\sigma_\mu^2 = 0$. For in this case the left-hand side of (2.6) is necessarily a.s. constant, and it is well known that $\mathbb{E}\{\text{MS}_{r.c}^{(m,n)}\} = \sigma_{r.c}^2$. (It is not hard to show that the assumption $\sigma_\mu^2 = 0$ can also be dropped.)

The argument leading to (2.2) and (2.3) also gives

$$\begin{aligned} &\mathbb{E}\{(X_{12} - X_{22})X_{11} | \mathcal{G}_{m,n}\} \\ &= \frac{1}{mn(n-1)} \sum_i \sum_{j \neq j'} X_{ij}X_{ij'} - \frac{1}{m(m-1)n(n-1)} \sum_{i \neq i'} \sum_{j \neq j'} X_{ij}X_{i'j'}, \end{aligned}$$

and the general results of III imply that this simplifies to

$$\begin{aligned} &\frac{1}{n} \left[\frac{n}{m-1} \sum_i (X_{i.} - X_{..})^2 - \frac{1}{(m-1)(n-1)} \sum_i \sum_j (X_{ij} - X_{i.} - X_{.j} + X_{..})^2 \right] \\ &= \frac{1}{n} [\text{MS}_r^{(m,n)} - \text{MS}_{r.c}^{(m,n)}], \end{aligned}$$

where $\text{MS}_r^{(m,n)}$ is the mean-square for rows in the anova of (X_{ij}) . As before, this expression converges a.s. under the appropriate ($\mathcal{L} \log \mathcal{L}$) moment conditions, and, again as before, if (X_{ij}) arises as in (2.5), then

$$\lim_{m,n} \frac{1}{n} [\text{MS}_r^{(m,n)} - \text{MS}_{r.c}^{(m,n)}] = \sigma_r^2 \quad \text{a.s.}$$

A similar argument gives the a.s. convergence of the appropriate linear combination of mean squares to σ_c^2 , and we have obtained the strong consistency of the usual anova estimators of components of variance for the two-way random effects model (2.5).

We turn now to a brief discussion of what we term *between/within exchangeability* (BWE), where an array $X = (X_{ij})$ satisfies BWE if, for every finitary permutation g of I , and for every map f from I to the set of all finitary permutations of J which sends all but finitely many elements of I to the identity permutation, we have

$$(2.7) \quad (X_{ij}: i \in I, j \in J) \stackrel{\mathcal{D}}{=} (X_{igjif}: i \in I, j \in J),$$

where ig is the image of i under g , and where jif is the image of j under the permutation if . Informally, the distribution of X is invariant under the permutation of finitely many first indices and, independently within each of finitely many

first indices, of the permutation of finitely many second indices. The group which permutes the pairs of indices is known as the restricted permutation wreath product of the groups S_∞ on J with S_∞ on I , and will be denoted by $S_\infty \text{ wr } S_\infty$, the restricted nature being taken for granted.

For $m \geq 1$ and $n \geq 1$, let S_m denote the subgroup of S_∞ acting on I which leaves each index $i > m$ fixed, and let $S_n^{(\infty)}$ denote the subgroup of $S_\infty \text{ wr } S_\infty$ which, independently within each of finitely many first indices i , permutes only the second indices $j = 1, \dots, n$. The group $G_{m,n} = S_m \cdot S_n^{(\infty)}$ is the permutation wreath product of S_n acting on $j = 1, \dots, n$ (within finitely many i) with S_m acting on $i = 1, \dots, m$. As before let $\mathcal{G}_{m,n}$ be the sub- σ -field of all events determined by (X_{ij}) which are invariant under $G_{m,n}$. Because $G_{m,n}$ is not compact, we cannot make use of the results of Dubins and Pitman (1979) and must argue directly to prove that $\mathcal{G}_{m,1}$ and $\mathcal{G}_{1,n}$ are conditionally independent given $\mathcal{G}_{m,n}$.

The argument goes like this. We prove that

$$(2.8) \quad \mathbb{E}_{m,1} \{ \mathbb{E}_{1,n} \{ t \} \} = \mathbb{E}_{m,n} \{ t \} = \mathbb{E}_{1,n} \{ \mathbb{E}_{m,1} \{ t \} \}$$

for functions t of the form $t_1 \circ X_{i_1 j_1} \cdots t_b \circ X_{i_b j_b}$, where $b \geq 1$, where t_1, \dots, t_b are arbitrary bounded functions, where $(i_1, j_1), \dots, (i_b, j_b)$ are an arbitrary set of pairs of indices $1 \leq i_1, \dots, i_b \leq m$, $1 \leq j_1, \dots, j_b \leq n$, and where $\mathbb{E}_{m,n}$ is an abbreviation for $\mathbb{E} \{ \cdot | \mathcal{G}_{m,n} \}$. A standard monotone convergence argument then implies the conditional independence result asserted. Now (2.8) is proved directly by averaging over the relevant orbits of $(\mathbf{i}, \mathbf{j}) = ((i_1, j_1), \dots, (i_b, j_b))$ under S_m , $S_n^{(\infty)}$, and $G_{m,n}$, respectively. Referring to Praeger *et al.* (1985) and Speed (1984) for notation and further details, we note that these orbits depend on the partition $\pi(1)$ of $\mathbf{b} = \{1, \dots, b\}$ induced by the equalities and inequalities of i_1, \dots, i_b , the partition $\pi(2)$ defined similarly by j_1, \dots, j_b , and the pair $(\pi(2), \pi(1)) \in \text{Hom}(2, \mathcal{P}(\mathbf{b}))$, respectively. The relevant facts we use are the following:

- (i) the S_m -orbit of (\mathbf{i}, \mathbf{j}) has cardinality $(m)_{\pi(1)}$;
- (ii) the $S_n^{(\infty)}$ -orbit of $(\mathbf{i}', \mathbf{j})$ has cardinality $(n)_{(\pi(2), \pi(1))}$ for every element $(\mathbf{i}', \mathbf{j})$ in the same S_m -orbit as (\mathbf{i}, \mathbf{j}) , and distinct elements of this S_m -orbit belong to disjoint $S_n^{(\infty)}$ -orbits;
- (iii) the $G_{m,n}$ -orbit of (\mathbf{i}, \mathbf{j}) has cardinality $(m)_{\pi(1)} (n)_{(\pi(2), \pi(1))}$ and can be viewed as the disjoint union of the $S_n^{(\infty)}$ -orbits of elements $(\mathbf{i}', \mathbf{j})$ belonging to the S_m -orbit of (\mathbf{i}, \mathbf{j}) .

Armed with these facts, (2.8) follows easily. For we see that

$$\mathbb{E}_{m,1} \{ t_1 \circ X_{i_1 j_1} \cdots t_b \circ X_{i_b j_b} \} = \frac{1}{(m)_{\pi(1)}} \sum^{(1)} t_1 \circ X_{i_1 j_1} \cdots t_b \circ X_{i_b j_b},$$

the sum $\Sigma^{(1)}$ being taken over all $(m)_{\pi(1)}$ elements in the S_m -orbit of (\mathbf{i}, \mathbf{j}) , whilst by linearity we have

$$\begin{aligned} \mathbb{E}_{1,n} \left\{ \mathbb{E}_{m,1} \left\{ t_1 \circ X_{i_1 j_1} \cdots t_b \circ X_{i_b j_b} \right\} \right\} \\ = \frac{1}{(m)_{\pi(1)}} \Sigma^{(1)} \mathbb{E}_{1,n} \left\{ t_1 \circ X_{i_1 j_1} \cdots t_b \circ X_{i_b j_b} \right\} \\ = \frac{1}{(m)_{\pi(1)}} \Sigma^{(1)} \frac{1}{(n)_{(\pi(2), \pi(1))}} \Sigma^{(2)} t_1 \circ X_{i_1 j_1} \cdots t_b \circ X_{i_b j_b}, \end{aligned}$$

the sum $\Sigma^{(2)}$ being taken over all the $(n)_{(\pi(2), \pi(1))}$ elements of the $S_n^{(\infty)}$ -orbit of $(\mathbf{i}', \mathbf{j})$. By the remarks (i), (ii) and (iii), this is seen to be just $\mathbb{E}_{m,n} \{ t_1 \circ X_{i_1 j_1} \cdots t_b \circ X_{i_b j_b} \}$. Thus (2.8) is proved, and we can use the convergence theorem for doubly-indexed RMGs. Note that the same argument involving the explicit averaging over orbits could have been used equally well for $S_m \times S_n$.

Having established the conditions on our σ -fields which are necessary to invoke the convergence theorem, we pass quickly to some illustrations which mirror and amplify the examples given for the simple crossed structure. For example, identities in III yield

$$(2.9) \quad \mathbb{E} \{ X_{11}(X_{11} - X_{12}) | \mathcal{G}_{m,n} \} = \frac{1}{m(n-1)} \sum_i \sum_j (X_{ij} - X_i)^2.$$

This is the familiar *within class* mean square which we denote by $MS_w^{(m,n)}$. Similarly,

$$(2.10) \quad \mathbb{E} \{ X_{11}(X_{12} - X_{22}) | \mathcal{G}_{m,n} \} = \frac{1}{n} [MS_b^{(m,n)} - MS_w^{(m,n)}],$$

where

$$MS_b^{(m,n)} = \frac{n}{m-1} \sum_i (X_{i.} - X_{..})^2$$

is the *between class* mean square from the usual anova. Two higher-order expressions which arise in the same way are

$$(2.11) \quad \mathbb{E} \{ X_{11}(X_{11}^2 - 3X_{12}^2 + 2X_{12}X_{13}) | \mathcal{G}_{m,n} \} = \frac{n}{m(n-1)(n-2)} \sum_i \sum_j (X_{ij} - X_i)^3,$$

which we have termed the *within class skewmulance*, $S_w^{(m,n)}$ say, and

$$(2.12) \quad \begin{aligned} \mathbb{E} \{ X_{11}(X_{11}^3 - 4X_{11}^2 X_{12} - 3X_{11} X_{12}^2 + 12X_{11} X_{12} X_{13} - 6X_{12} X_{13} X_{14}) | \mathcal{G}_{m,n} \} \\ = \frac{n}{m(n)_4} \left\{ n(n+1) \sum_i \sum_j (X_{ij} - X_i)^4 - 3(n-1) \sum_i \left[\sum_j (X_{ij} - X_i)^2 \right]^2 \right\}, \end{aligned}$$

which we call the *within class kurtosance*, $K_w^{(m,n)}$ say.

Under the moment conditions necessary for the use of the a.s. convergence theorems, each of (2.9), (2.10), (2.11) and (2.12) converges a.s. as m, n increase, and it only remains to identify the limits wherever possible. The appropriate general additive model here is

$$X_{ij} = \mu + b_i + w_{ij},$$

where $\mu, (b_i)$ and (w_{ij}) are independent sets of independent effects with variances σ_μ^2, σ_b^2 and σ_w^2 , respectively. Naturally we find that under the appropriate moment conditions:

$$\begin{aligned} \lim_{m,n} MS_w^{(m,n)} &= \sigma_w^2 \quad \text{a.s.} \\ \lim_{m,n} MS_b^{(m,n)} &= \sigma_b^2 \quad \text{a.s.} \\ \lim_{m,n} S_w^{(m,n)} &= \kappa_3(w) \quad \text{a.s.} \\ \lim_{m,n} K_w^{(m,n)} &= \kappa_4(w) \quad \text{a.s.,} \end{aligned}$$

where $\kappa_3(w)$ and $\kappa_4(w)$ are the third and fourth cumulants of the within effects (w_{ij}) .

3. Group-theoretic preliminaries

In this section we organise the group-theoretic results which are necessary for us to be able to give a compact proof of our main convergence result. For general background we refer to earlier papers in this series.

The random arrays $y = (y_i)_{i \in I}$ which we consider will be labelled by the elements $\mathbf{i} = (i_p : p \in P)$ of the set $\mathbf{I} = \prod_{p \in P} \mathbf{I}_p$ where, for each element p of the partially ordered set (P, \leq) , the set $\mathbf{I}_p = \{1, 2, \dots\}$. The notion of *generalised wreath product* of permutation groups G_p acting on the components $\mathbf{I}_p, p \in P$, was defined and studied in Bailey *et al.* (1983) in the case $|\mathbf{I}_p| < \infty, p \in P$, and we refer to that paper for fuller details concerning what follows. The definition and main results given there extend naturally to the case where, for all $p \in P$, the groups G_p consist of all finitary permutations of a countably infinite set \mathbf{I}_p , and the maps $f_p : \prod\{\mathbf{I}_q : q > p\} \rightarrow G_p$ appearing in the definition send all but a finite number of elements of $\prod\{\mathbf{I}_q : q > p\}$ to the identity permutation 1_p of G_p . The group arising from the construction so restricted will be termed the *restricted generalised wreath product* of the (restricted) symmetric groups G_p acting on $\mathbf{I}_p, p \in P$. It will be denoted by $G = \prod^*\{G_p : p \in P\}$ and abbreviated to $G = \text{GW}(\mathbf{I})$ when no confusion will result.

Having defined the groups G permuting our underlying index set \mathbf{I} , our next task is to describe the orbits of the action of G on m -tuples \mathbf{I}^m . It is not hard to see that these may be described in exactly the same way as when $\mathbf{I}_p = \{1, \dots, n_p\}$

and $G_p = S_{n_p}$, the full symmetric group on $n_p < \infty$ symbols, $p \in P$, and so (see III, Section 2 for further details) they are labelled by the elements $\varphi = (\varphi(p) : p \in P)$ of $\text{Hom}(P, \mathcal{P}(\mathbf{m}))$, the lattice of all order-preserving maps from the nesting poset P to the lattice $\mathcal{P}(\mathbf{m})$ of all partitions of $\mathbf{m} = \{1, \dots, m\}$.

In Praeger *et al.* (1985) a discussion may be found of those subgroups $G^{(p)}$, $p \in P$, of the generalised wreath product G of finite groups acting on finite sets which correspond to moving only the coordinate i_p of \mathbf{i} labelled by $p \in P$, and that discussion applies without change to our context, where the groups G_p are infinite but consist of finitary permutations of a countable set.

LEMMA 3.1. *For any pair $r, s \in P$ we have $G^{(r)}G^{(s)} = G^{(s)}G^{(r)}$.*

PROOF. Take $f \in G^{(r)}$, $g \in G^{(s)}$, and $\mathbf{i} \in \mathbf{I}$. Then the definitions of Bailey *et al.* (1983) and Praeger *et al.* (1985) imply that $(\mathbf{i}gf)_r = i_r(\mathbf{i}g)\pi^r f_r$ and $(\mathbf{i}gf)_s = i_s(\mathbf{i}\pi^s)g_s$. Our aim is to define an $f' \in G^{(r)}$ such that $f'g = gf$.

Suppose that r and s are incomparable in (P, \leq) . Then $f' = f$ will do, for in this case $\mathbf{i}g\pi^r = \mathbf{i}\pi^r$, and it follows that $(\mathbf{i}gf)_r = (\mathbf{i}f)_r$ and $(\mathbf{i}gf)_s = (\mathbf{i}f)_s$, whence $gf = fg$.

On the other hand, suppose that $r < s$. Then it is also true that $(\mathbf{i}f)_p = i_p(\mathbf{i}\pi^p)f_p = i_p$ for all $p > s$, and so $\mathbf{i}f\pi^s = \mathbf{i}\pi^s$. Since f was an arbitrary (but fixed) element of $G^{(r)}$, this is in fact true for every element of $G^{(r)}$. Let us define $f' \in G^{(r)}$ as follows: put $f'_p = 1_p$, the identity element of G_p when $p \neq r$, and define $f'_r : \prod_{p>r} \mathbf{I}_p \rightarrow G_r$ by $\mathbf{i}\pi^r f'_r = (\mathbf{i}g)\pi^r f_r$. With this definition we find that

$$(\mathbf{i}f'g)_r = i_r(\mathbf{i}\pi^r)f'_r = i_r(\mathbf{i}g)\pi^r f_r = (\mathbf{i}gf)_r$$

and

$$(\mathbf{i}f'g)_s = i_s(\mathbf{i}f')\pi^s g_s = i_s(\mathbf{i}\pi^s)g_s,$$

i.e. $f'g = gf$.

COROLLARY. *Let p_1, p_2, \dots, p_a be any enumeration of the $a = |P|$ elements of the poset P . Then we have*

$$G = G^{(p_1)}G^{(p_2)} \dots G^{(p_a)}.$$

Our next lemma concerns averaging of functions on a set over the orbits of a permutation group acting on that set. Let G be a group acting on a set \mathbf{I} ; in what follows G will be our (restricted) generalised wreath product and $\mathbf{I} = \prod\{\mathbf{I}_p : p \in P\}$, but for the present both can be arbitrary. The *orbit* of the point $\mathbf{i} \in \mathbf{I}$ is

$i^G = \{ig : g \in G\}$, and the stabilizer of i is $G_i = \{g \in G : ig = i\}$. The points of i^G are in one-to-one correspondence with the cosets of G_i and hence $|i^G| = |G : G_i|$; we suppose in what follows that all orbits are finite. I am very grateful to Don Taylor for providing the proofs of the lemma and corollary which follow.

LEMMA 3.2. *If S and T are subgroups of G such that $ST = TS$, then for any $i \in \mathbf{I}$ and $j \in i^{ST}$, we have $|i^S \cap j^T| = |i^S| |j^T| / |i^{ST}|$.*

PROOF. In proving this result we may assume that ST is faithful and transitive on I . Then each orbit of T has non-empty intersection with each orbit of S , and we may assume that $j = i$. Choose representative s_1, \dots, s_n for the cosets $s(S \cap T)$ of $S \cap T$ in S . These elements are also coset representatives for T in ST , and therefore $|ST|/|T| = |S|/|S \cap T|$.

Let $X = \{(s, t) \in S \times T | is = it\}$. If $s \in S \cap T$, $t \in T$ and $s_a t \in (ST)_i$, then $(s_a s, t^{-1}s) \in X$. We shall show that every element of X can be written in this form in exactly one way. Indeed, if $(s_1, t_1) \in X$, let $x = s_1 t_1^{-1}$ and write $x = s_a t$, where $t \in T$. Then $x \in (ST)_i$, $s_a^{-1} s_1 = t t_1 \in S \cap T$, and we have $s_1 = s_a (s_a^{-1} s_1)$ and $t_1 = t^{-1} (s_a^{-1} s_1)$, as required. It is clear that s_a , t and $s_a^{-1} s_1$ are uniquely determined by (s_1, t_1) . This shows that there is a one-to-one correspondence between X and $(ST)_i \times (S \cap T)$ and hence that $|X| = |(ST)_i| |S \cap T|$.

If $j \in i^S \cap i^T$, then $j = is = it$ for some $s \in S$ and $t \in T$, and we have $(s, t) \in X$. If we also have $(s_1, t_1) \in X$ and $j = is_1$, then $s_1 \in S_i s$ and $t_1 \in T_i t$. It follows that $|X|/|S_i| |T_i| = |i^S \cap i^T|$, and consequently

$$\begin{aligned} |i^S \cap i^T| &= |(ST)_i| |S \cap T| / |S_i| |T_i| \\ &= |(ST)_i| |S| |T| / |ST| |S_i| |T_i| \\ &= |i^S| |i^T| / |i^{ST}|. \end{aligned}$$

As a corollary we have the following discrete version of a result of Dubins and Pitman (1979).

COROLLARY. *If f is any function defined on \mathbf{I} , and if S and T are subgroups of G such that $ST = TS$, then*

$$|i^S|^{-1} \sum_{j \in i^S} |j^T|^{-1} \sum_{k \in j^T} f(k) = |i^{ST}|^{-1} \sum_{k \in i^{ST}} f(k).$$

PROOF. If O_1, O_2, \dots, O_d are the orbits of T that meet i^S , then

$$|i^S|^{-1} \sum_{j \in i^S} |j^T|^{-1} \sum_{k \in j^T} f(k) = |i^S|^{-1} \sum_{a=1}^d |i^S \cap O_a| |O_a|^{-1} \sum_{k \in O_a} f(k).$$

By the lemma, $|\mathbf{i}^S|^{-1}|\mathbf{i}^S \cap O_a||O_a|^{-1} = |\mathbf{i}^{ST}|^{-1}$ and so the above expression reduces to $|\mathbf{i}^{ST}|^{-1} \sum_{\mathbf{k} \in \mathbf{i}^{ST}} f(\mathbf{k})$.

A generalised form of the corollary is clearly valid when S, T, \dots, U are an arbitrary number of permutable sub-groups of G , and it is in this form that we use the result in Lemma 4.1 below.

4. The main result

In this section the integer $m \geq 1$ is fixed, and $X = (X_i)_{i \in \mathbf{I}}$ is an array of random m -vectors labelled by the index set $\mathbf{I} = \prod \mathbf{I}_p$. Where necessary we will denote the components of X_i by $(X_i(l) : l \in \mathbf{m})$. Our random array will satisfy the following *generalised exchangeability assumption* (GE): the joint distribution of X is invariant under the action of the (restricted) generalised wreath product group $G = \text{GW}(\mathbf{I})$ of the (restricted) symmetric groups on the $\mathbf{I}_p, p \in P$, i.e.

$$(4.1) \quad (X_i : i \in \mathbf{I}) \stackrel{\mathcal{D}}{=} (X_{ig} : i \in \mathbf{I}), \quad g \in G.$$

To enable us to formulate our limit theorem we must focus on the finite subarrays $X^{(\mathbf{n})} = (X_i : i \in \mathbf{I}^{(\mathbf{n})})$ of X obtained by *truncating* \mathbf{I} to $\mathbf{I}^{(\mathbf{n})} = \prod_{p \in P} \{1, \dots, n_p\}, 1 \leq n_p < \infty, p \in P$. In terms of $X^{(\mathbf{n})}$ we define the random variables

$$a_\pi(X^{(\mathbf{n})}) = [A_\pi^{(\mathbf{n})} | X^{(\mathbf{n})}]$$

and

$$k_\pi(X^{(\mathbf{n})}) = [F_\pi^{(\mathbf{n})} | X^{(\mathbf{n})}], \quad \pi \in \text{Hom}(P, \mathcal{P}(\mathbf{m})).$$

Here $A_\pi^{(\mathbf{n})}$ and $F_\pi^{(\mathbf{n})}$ are the association and generalised k -statistic tensors defined in III, G -symmetric functions of X which will turn out to be RMGs multi-indexed by $\mathbf{n} = (n_p : p \in P)$.

For a given $p \in P$ and integer $n_p > 1$, we define $G_{n_p}^{(p)}$ to be the subgroup of $G^{(p)}$ consisting of all permutations of \mathbf{I} which move only the coordinate i_p of $\mathbf{i} = (i_p : p \in P)$, and which leave all values $i_p > n_p$ fixed. Similarly we define $\mathcal{G}_{n_p}^{(p)}$ to be the sub- σ -field of our basic probability space consisting of all events generated by $X = (X_i)$ which are invariant under $G_{n_p}^{(p)}$. Finally, for $\mathbf{n} = (n_p : p \in P)$ write $\mathcal{F}_\mathbf{n}$ for $\bigcap_{p \in P} \mathcal{G}_{n_p}^{(p)}$, the σ -field of all events defined by X which are invariant under $G^{(\mathbf{n})}$, the permutable product of all the $G_{n_p}^{(p)}$.

LEMMA 4.1. *The sub- σ -fields $\{\mathcal{G}_{n_p}^{(p)} : p \in P\}$ are mutually conditionally independent given $\mathcal{F}_\mathbf{n}$.*

PROOF. The proof has three steps which we now outline. Firstly, we note that for $p \in P$, $\mathbb{E}_p\{t\} = \mathbb{E}\{t | \mathcal{G}_{n_p}^{(p)}\}$ is obtained by averaging t over the appropriate orbit when t has the form

$$(4.2) \quad (t_1 \circ X_i)(t_2 \circ X_j) \cdots (t_b \circ X_k),$$

where $i, j, \dots, k \in \mathbf{I}$, and where t_1, t_2, \dots, t_b are bounded real-valued functions. Secondly, we use the generalised form of the corollary to Lemma 3.2 above to deduce that the conditional expectation operators $\{\mathbb{E}_p : p \in P\}$ all commute in their action on the set \mathcal{C} of all t of the form (4.2), i.e.

$$(4.3) \quad \mathbb{E}_p \mathbb{E}_q \cdots \mathbb{E}_r \{t\} = \mathbb{E}\{t | \mathcal{F}_n\}$$

for every enumeration p, q, \dots, r of P . And thirdly, we observe that the set \mathcal{H} of all bounded measurable functions t of X which satisfy (4.3) for every enumeration p, q, \dots, r of P satisfies the conditions of Meyer (1966, I T20), whilst the set \mathcal{C} is contained in \mathcal{H} and is closed under multiplication.

Turning now to the details, we begin by noting that if t has the form (4.2), the average giving $\mathbb{E}_p t$ will be over the $G_{n_p}^{(p)}$ -orbit labelled by the partition of $b = \{1, 2, \dots, b\}$ defined by the equalities and inequalities amongst $\{i_p, j_p, \dots, k_p\}$. Denoting this partition by $\sigma(p)$, we have, cf. III and Speed (1984),

$$\mathbb{E}_p t = \frac{1}{(n_p)_{(\sigma(p), \wedge\{\sigma(q) : q > p\})}} \sum (t_1 \circ X_{i'}) (t_2 \circ X_{j'}) \cdots (t_b \circ X_{k'})$$

the sum being taken over all k -tuples i', j', \dots, k' from $\mathbf{I}^{(n)}$ in the $G_{n_p}^{(p)}$ -orbit which contains i, j, \dots, k .

The final two steps really need little elaboration, and it follows from Meyer (1966, I T20) that \mathcal{H} contains all measurable function of X .

COROLLARY. For t of the form (4.2) and for $1 \leq i_p, j_p, \dots, k_p \leq n_p, p \in P$, we have

$$\mathbb{E}\{t | \mathcal{F}_n\} = \frac{1}{(n)_\pi} \sum (t_1 \circ X_{i'}) (t_2 \circ X_{j'}) \cdots (t_b \circ X_{k'}),$$

where $\pi(p) = \wedge\{\sigma(q) : q \geq p\}, p \in P$, where the sum is taken over all (i', j', \dots, k') in the same G -orbit as (i, j, \dots, k) , and where

$$(n)_\pi = \prod_{p \in P} (n_p)_{(\pi(p), \wedge\{\pi(q) : q > p\})}$$

PROOF. This is immediate from the proof of the lemma and from the formula for $(n)_\pi$ given in Speed (1984).

Now let us return to the random variables $a_\pi(X^{(n)})$ and $k_\pi(X^{(n)})$ which for convenience we write as $a_n(\pi)$ and $k_n(\pi)$ respectively. From the definition of $a_n(\pi)$ we have

$$(\mathbf{n})_\pi^{-1} a_n(\pi) = (\mathbf{n})_\pi^{-1} \sum X_{h(1)}(1) \cdots X_{h(m)}(m)$$

the sum being taken over all maps $h: \mathbf{m} \rightarrow \prod_{p \in P} \{1, \dots, n_p\} = \mathbf{I}^{(n)}$ with $\varphi^h = \pi$, (see Praeger *et al.* (1985)), i.e. over all h in the orbit of the (permutable) product $G^{(n)}$ of all the groups $G_{n_p}^{(p)}$, $p \in P$. Thus we have

LEMMA 4.2. *For any $h: \mathbf{m} \rightarrow \mathbf{I}^{(n)}$ with $\varphi^h = \pi$ we have*

$$(\mathbf{n})_\pi^{-1} a_n(\pi) = \mathbb{E} \{ X_{h(1)}(1) \cdots X_{h(m)}(m) \mid \mathcal{F}_n \} \quad \text{a.s.}$$

COROLLARY. (i) $\{(\mathbf{n})_\pi^{-1} a_n(\pi); \mathcal{F}_n\}$ is a RMG.

(ii) $\{k_n(\pi); \mathcal{F}_n\}$ is a RMG.

PROOF. It is clear that the sub- σ -fields \mathcal{F}_n decrease as \mathbf{n} increases, i.e. as $n_p \uparrow \infty$, $p \in P$, and so (i) is true. The generalised k -statistics $k_n(\pi)$ are just linear combinations of the random variables $(\mathbf{n})_\sigma^{-1} a_n(\sigma)$ for $\sigma \leq \pi$, and so (ii) follows.

As explained in Section 2 above, we get a.s. convergence of multi-indexed RMGs provided that they satisfy a moment condition and their defining sub- σ -fields satisfy the appropriate conditional independence property. In our case the following moment condition is adequate: for $\sigma \in \text{Hom}(P, \mathcal{P}(\mathbf{m}))$, and for $h: \mathbf{m} \rightarrow \mathbf{I}^{(n)}$ with $\varphi^h = \sigma$, we have

$$(4.4)_\sigma \quad X_{h(1)}(1) \cdots X_{h(m)}(m) \in \mathcal{L}(\log \mathcal{L})^{a-1},$$

where $a = |P|$.

THEOREM 4.3. *Suppose that the random array X satisfies (4.1). If X also satisfies (4.4) $_\sigma$, then $(\mathbf{n})_\sigma^{-1} a_n(\sigma)$ converges a.s. as \mathbf{n} increases. Further, if X satisfies (4.4) $_\sigma$ for all $\sigma \leq \pi$, with $\pi \in \text{Hom}(P, \mathcal{P}(\mathbf{m}))$, then $k_n(\pi)$ converges a.s. as \mathbf{n} increases.*

PROOF. By virtue of Lemma 4.1 this is essentially the main result of Gut (1976). See also Dubins and Pitman (1979, page 306).

It follows from the theorem that

$$\lim_{\mathbf{n}} (\mathbf{n})_\sigma^{-1} a_n(\sigma) = \mathbb{E} \{ X_{h(1)}(1) \cdots X_{h(m)}(m) \mid \mathcal{F}_\infty \} \quad \text{a.s.}$$

for any $h: \mathbf{m} \rightarrow \mathbf{I}^{(m)} = \{1, \dots, m\}^a$ with $\varphi^h = \sigma$, and a similar result holds for $\lim_{\mathbf{n}} k_n(\pi)$. This naturally leads to the question: under what conditions is it true that $\mathcal{F}_\infty = \bigcap_{\mathbf{n}} \mathcal{F}_n$ is trivial? Let us suppose that the array X arises as a

generalised additive model, i.e. that for each $\mathbf{i} \in \mathbf{I}$ we have

$$(4.5) \quad X_{\mathbf{i}} = \sum_{a \in F(P)} \varepsilon_{\mathbf{i}}(a),$$

where $F(P)$ is the distributive lattice of all filters of the nesting poset P , where, for $a \in F(P)$, the random variables $\varepsilon_{\mathbf{i}}(a)$ are mutually independent and identically distributed and depend only on $(i_p, p \in a)$, and where the sets of ε corresponding to distinct $a \in P$ are mutually independent.

PROPOSITION 4.4. *If the array X arises as a generalised additive model, and if $\varepsilon(\emptyset)$ is a.s. constant, then $\bigcap_{\mathbf{n}} \mathcal{F}_{\mathbf{n}}$ is trivial.*

PROOF. Let $\mathcal{R}_{\mathbf{n}}, \mathcal{S}_{\mathbf{n}}$ and $\mathcal{T}_{\mathbf{n}}$ denote the σ -fields generated by the r.v.s $\varepsilon_{\mathbf{i}}(a)$ for all $\emptyset \neq a \in F(P)$, and for all $\mathbf{i} \in \mathbf{I}$ subject to $i_p \leq n_p$ ($p \in P$), $n_p < i_p \leq 2n_p$ ($p \in P$), and $i_p > n_p$ ($p \in P$), respectively. We prove the proposition in three steps (cf. Meyer (1966, VIII §1)), showing in turn that (i) $\mathcal{T}_{\infty} = \bigcap_{\mathbf{n}} \mathcal{T}_{\mathbf{n}} \subseteq \mathcal{F}_{\infty}$; (ii) $\mathcal{F}_{\infty} \subseteq \mathcal{T}_{\infty}$ a.s. and (iii) \mathcal{T}_{∞} is trivial.

Firstly, note that $\mathcal{T}_{\mathbf{n}}$ is generated by events invariant under $G^{(n)}$, and so, by Meyer I T19, all sets in $\mathcal{T}_{\mathbf{n}}$ have this property. Now take $A \in \mathcal{T}_{\infty}$ and $g \in G$. By the construction of the restricted generalised wreath product group G , there exists \mathbf{n} such that $g \in G^{(n)}$, and so, since $A \in \mathcal{T}_{\infty} \subseteq \mathcal{T}_{\mathbf{n}}$, we have $g^{-1}A = A$; it follows that $A \in \mathcal{F}_{\infty}$, and (i) is proved.

Next let $t = 1_A$ be the indicator function of an event $A \in \mathcal{F}_{\infty}$, and take $\varepsilon > 0$. By the convergence theorem for multi-indexed RMGs of Gut (1976) cited above, there exists N_1 sufficiently large that for $\mathbf{n} \geq N_1$,

$$(4.6) \quad \|\mathbb{E}\{t | \mathcal{T}_{\mathbf{n}}\} - \mathbb{E}\{t | \mathcal{T}_{\infty}\}\|_1 < \varepsilon.$$

It is also possible to choose N_2 sufficiently large that for $\mathbf{n} \geq N_2$,

$$(4.7) \quad \|t - \mathbb{E}\{t | \mathcal{R}_{\mathbf{n}}\}\|_1 < \varepsilon.$$

Now it is easy to see that G contains the permutation g of \mathbf{I} which interchanges 1 and $n_p + 1$, 2 and $n_p + 2, \dots, n_p$ and $2n_p$ and which leaves all other integers $i_p > 2n_p$ fixed, $p \in P$. Such a permutation leaves the joint distribution of X fixed, sends $\mathcal{R}_{\mathbf{n}}$ into $\mathcal{S}_{\mathbf{n}}$, and leaves t unchanged, whence

$$(4.8) \quad \|t - \mathbb{E}\{t | \mathcal{S}_{\mathbf{n}}\}\|_1 < \varepsilon.$$

Now $\mathcal{S}_{\mathbf{n}} \subseteq \mathcal{T}_{\mathbf{n}}$, and $\mathbb{E}\{\cdot | \mathcal{T}_{\mathbf{n}}\}$ is a contraction, so that

$$(4.9) \quad \|\mathbb{E}\{t | \mathcal{S}_{\mathbf{n}}\} - \mathbb{E}\{t | \mathcal{T}_{\mathbf{n}}\}\|_1 < \varepsilon.$$

But (4.6), (4.8) and (4.9) imply that for $\mathbf{n} \geq N_1 \vee N_2$,

$$\|t - \mathbb{E}\{t | \mathcal{T}_{\infty}\}\|_1 < 3\varepsilon$$

and this, since ε was arbitrary, implies that $t \in \mathcal{F}_{\infty}$ a.s.

Finally, we remark that a familiar argument giving the Kolmogorov 0-1 law, see e.g. Rényi (1970, page 280), implies that the σ -field $\mathcal{T}_\infty = \bigcap_n \mathcal{T}_n$ is trivial; \mathcal{T}_∞ is clearly a “tail” σ -field, and the fact that the r.v.s $\varepsilon_i(a)$ are multi-indexed is irrelevant, their independence being the key property.

The following corollary is an immediate consequence of the preceding result and the main theorem. Together with the corollary to Proposition 5.1 of III it can be seen to imply the a.s. convergence of all the standard unbiased estimates of components of variance to their expectations under a quite general additive (random effects) model. The notation $f(\pi)$ is defined in III.

COROLLARY. *If X arises as a generalised additive model, then for all $\pi \in \text{Hom}(P, \mathcal{P}(m))$, we have $\lim_n k_n(\pi) = f(\pi)$ a.s.*

6. Acknowledgements

Many thanks are due to Jim Pitman for a number of helpful discussions on the material in this paper and to Don Taylor for providing the proof of Lemma 3.2 and its corollary.

References

- David J. Aldous (1981), ‘Representations for partially exchangeable arrays of random variables’, *J. Multivar. Anal.* **11** 581–598.
- David J. Aldous (1985), ‘Exchangeability and related topics’, *École d’Été de Probabilités de Saint-Flour XIII* 1983, eds., D. J. Aldous, I. A. Ibragimov and J. Jacod, Lecture Notes in Mathematics 1117, Springer, New York, 1–198.
- R. A. Bailey, C. E. Praeger, C. A. Rowley and T. P. Speed (1983), ‘Generalised wreath products of permutation groups’, *Proc. London. Math. Soc.* (3) **47**, 69–82.
- Lester E. Dubins and Jim Pitman (1979), ‘A pointwise ergodic theorem for the group of rational rotations’, *Trans. Amer. Math. Soc.* **251**, 299–308.
- Allan Gut (1976), ‘Convergence of reversed martingales with multidimensional indices’, *Duke Math. J.* **43**, 269–275.
- John James Miller (1973), *Asymptotic properties and computation of maximum likelihood estimates in the mixed model of the analysis of variance* (Ph.D. Thesis, Stanford University).
- John J. Miller (1977), ‘Asymptotic properties of maximum likelihood estimates in the mixed model of the analysis of variance’, *Ann. Statist.* **5**, 746–762.
- P. A. Meyer (1966), *Probabilities and potentials* (Blaisdell, Waltham, Mass.)
- C. E. Praeger, C. A. Rowley and T. P. Speed (1985), ‘A note on generalised wreath product groups’, *J. Austral. Math. Soc. (Ser. A)* **39**, 415–420.

Alfred Rényi (1970), *Foundations of probability* (Holden-Day, San Francisco).

T. P. Speed (1984), 'On the Mobius function of $\text{Hom}(P, Q)$ ', *Bull. Austral. Math. Soc.* **29**, 39–46.

T. P. Speed (1986a), 'Cumulants and partition lattices II: Generalised k -statistics', *J. Austral. Math. Soc. (Ser. A)* **40**, 34–53.

T. P. Speed (1986b), 'Cumulants and partition lattices III: Multiply-indexed arrays', *J. Austral. Math. Soc. (Ser. A)* **40**, 161–182.

C.S.I.R.O.

Box 1965, G.P.O.

Canberra, A.C.T. 2601

Australia