Permutation Hypertrees In Probability

Fred M. Hoppe[1]
Department of Mathematics and Statistics
McMaster University
Hamilton, Ontario L8S 4K1
CANADA

David A. Grable
Department of Algebra, Combinatorics, and Analysis
Auburn University
Auburn, Alabama 36849
U.S.A.

Abstract. We examine properties of a class of hypertrees, occurring in probability, which are described by sequences of subscripts.

1. Introduction.

Tomescu [6] shows how to improve the Bonferroni inequalities for the probability of a union of events by including a term determined by a hypertree (defined in the next section). Specifically, if $\{A_1, \dots, A_n\}$ are arbitrary events and $S_{j,n} = \Sigma P(A_{i_1} \dots A_{i_j})$, where $A_{i_1} \dots A_{i_j}$ is the notation for the intersection of these events, and the sum is taken over all subsets $1 \leq i_1 < \dots < i_j \leq n$ of size j then:

$$P\left(\bigcup_{i=1}^{n} A_i\right) \leq \sum_{j=1}^{r} (-1)^{j-1} S_{j,n} - T_{r+1,n}, \quad r \text{ odd}; \tag{1.1a}$$

$$P\left(\bigcup_{i=1}^{n} A_i\right) \geq \sum_{j=1}^{r} (-1)^{j-1} S_{j,n} + T_{r+1,n}, \quad r \text{ even}; \tag{1.1b}$$

where $T_{r+1,n} = \Sigma P(A_{i_1} \dots A_{i_{r+1}})$, the summation being over all the edges $\{i_1, \dots, i_{r+1}\}$ of any hypertree of order n and degree $r+1$. Referring to a probability term or bound as being of degree d if it involves intersections of at most d of the $\{A_1, \dots, A_n\}$ then these bounds are of degree $r+1$. When $r = 1$ (1.1a) becomes the degree two upper bound

$$P\left(\bigcup_{i=1}^{n} A_i\right) \leq S_{1,n} - \sum_{\tau} P(A_i A_j), \tag{1.2}$$

[1] Research supported by NSERC of Canada Operating Grant.

τ being any tree with vertices $\{1, 2, \ldots, n\}$, an inequality due to Hunter [4].

A degree three lower bound

$$P\left(\bigcup_{i=1}^{n} A_i\right) \geq S_{1,n} - S_{2,n} + \sum_{i=3}^{n}\sum_{j=2}^{i-1} P(A_{j-1}A_jA_i), \tag{1.3}$$

was earlier obtained by Hoppe [2] by a method in which an existing upper (lower) bound is "iterated" into a lower (upper) bound. (1.3) is of the form (1.1b) with $r = 2$ because the collection of triples $\{(j-1, j, i): 2 \leq j \leq i-1, 3 \leq i \leq n\}$ form the edges of a 3-hypertree, and, in fact, this method is closely related to Tomescu's inductive proof of (1.1) since a hypertree emerges as the object to which a tree iterates.

Using iteration beginning with the inequality

$$P\left(\bigcup_{i=1}^{n} A_i\right) \geq \max_{1 \leq i \leq n} P(A_i)$$

and then optimizing over all permutations of subscripts Seneta [5] obtained

$$P\left(\bigcup_{i=1}^{n} A_i\right) \leq S_{1,n} - \max_{\Pi} \sum_{i=2}^{n} \max_{1 \leq k \leq i-1} P(A_iA_k) \tag{1.2p}$$

where Π is the set of all permutations of subscripts of $A_1, \ldots, A_n$. Since, subject to some permutation, the edges of any tree can be described as a collection of pairs $\{(i, k); i \geq 2, 1 \leq k \leq i-1\}$ this formulation produces the optimal bound in the class (1.2) by finding a maximal spanning tree using permutation of subscripts to generate all trees.

More recently Hoppe and Seneta [3], refining [2], developed a direct approach, not requiring induction or iteration, by first establishing an identity for the probability of a union

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{j=1}^{r} (-1)^{j-1} S_{j,n} + (-1)^r \Delta_{r+1,n}, \quad r \geq 0 \tag{1.4}$$

where

$$\Delta_{r+1,n} = \sum_{2 \leq i_1 < i_2 \ldots < i_r \leq n} P\left(\bigcup_{k=1}^{i_1-1} A_k A_{i_1} A_{i_2} \ldots A_{i_r}\right)$$

and then bounding each $P(\cup A_k A_{i_1} A_{i_2} \ldots A_{i_r})$ by the probability of a single term in the union, say $P(A_k A_{i_1} A_{i_2} \ldots A_{i_r})$ for some $1 \leq k \leq i_1 - 1$. The events

A_i are given an arbitrary labelling from which ensue the following degree $r+1$ bounds:

$$P\left(\bigcup_{i=1}^{n} A_i\right) \leq \sum_{j=1}^{r}(-1)^{j-1} S_{j,n} - D_{r+1,n}, \quad r \text{ odd}; \tag{1.5a}$$

$$P\left(\bigcup_{i=1}^{n} A_i\right) \geq \sum_{j=1}^{r}(-1)^{j-1} S_{j,n} - D_{r+1,n}, \quad r \text{ even}; \tag{1.5b}$$

where

$$D_{r+1,n} = \max_{\Pi} \sum_{2 \leq i_1 < i_2 \ldots < i_r \leq n} \max_{1 \leq k \leq i_1 - 1} P(A_k A_{i_1} A_{i_2} \ldots A_{i_r}). \tag{1.6}$$

When $r \geq 2$ the collection of subscripts permitted on the right side of (1.6) has the structure

$$(k, i_1, \ldots, i_r): \quad 2 \leq i_1 < \ldots < i_r \leq n, \ 1 \leq k \leq i_1 - 1 \tag{1.7}$$

subject to some permutation. Since all trees can be generated by permutation of subscripts as in (1.2p) it might be expected that r-hypertrees ($r \geq 3$), which are defined recursively and have a structure not as visually transparent as that of trees, could also be described concretely with the representation (1.7). However, (1.7) does not generate all hypertrees. It does describe a large class of r-hypertrees having an interesting structure some of whose properties are presented here.

2. Recursive definition of hypertrees.

Let $X = \{x_1, \ldots, x_n\}$ be a set of vertices and $E = \{E_1, \ldots, E_m\}$ a collection of subsets of X whose members are called edges. The pair (X, E) is called an r-uniform hypergraph if, for each $1 \leq i \leq m$, E_i contains exactly r vertices.

Definition: [6] An r-hypertree $\tau_n^r = (X, E)$ of order n and degree r is an r-uniform hypergraph having the following additional structure:

(a) if $r = 2$ then τ_n^r is a tree;
(b) if $r \geq 3$ and $r = n$ then τ_n^r has only one edge $\{x_1, \ldots, x_n\}$;
(c) if $r \geq 3$ and $r \leq n-1$ then there is some point, x, called a terminal vertex, for which:
  (i) if $E_1, \ldots, E_m$ are all the edges containing x then $E_1 - \{x\}, \ldots, E_m - \{x\}$ are the edges of an $(r-1)$-hypertree with vertex set $X - \{x\}$;
  (ii) if $F_1, \ldots, F_k$ are all the edges not containing x then they form the edges of an r-hypertree on $X - \{x\}$.

The definition is recursive, an r-hypertree of order n being defined in terms of an $(r-1)$-hypertree of order $n-1$ and an r-hypertree of order $n-1$. Each of these hypertrees in turn is constructed from hypertrees of order $n-2$. Thus, at this stage four hypertrees are required. The procedure continues, each hypertree generated requiring two for its construction, until either it has been reduced to an a-hypertree of order a $(3 \leq a \leq r)$ or a 2-hypertree (that is a tree) of order b $(2 \leq b \leq n-r+2)$. Each of these may readily be constructed and may be viewed as the ultimate basic building blocks.

We count the number required in general. At each step in the recursive construction the order of the trees involved is dropped, and the degree is either maintained or reduced by one. In the steps to a τ_a^a beginning with a τ_n^r the penultimate hypertree must be a τ_{a+1}^a. This requires $r-a$ reductions of degree and $n-(a+1)$ reductions of order. Since a reduction of degree always results in a reduction of order, we need count the number of reductions of order which also reduce degree. There are clearly $\binom{n-a-1}{r-a}$ such selections. To arrive at a τ_b^2 hypertree the penultimate hypertree must be τ_{b+1}^3 which requires $r-3$ reductions of degree and $n-(b+1)$ reductions of order, for a total of $\binom{n-b-1}{r-3}$ possible choices. These binomial coefficients count the maximum number of hypertrees of the specified types needed for the construction of a general τ_n^r. Of course, fewer may be needed if the same lower order hypertrees are used more than once. For instance when $n = 10$ and $r = 5$, in general as many as 56 individual hypertrees may need to be constructed to build up a general τ_{10}^5, although as few as eight of these may suffice (by duplication) to construct various special cases. These are enumerated below according to type and multiplicity. The multiplicity refers to the maximum number of hypertrees, of the degree and order specified, required for the general case, calculated by the binomial coefficients above.

hypertree	τ_5^5	τ_4^4	τ_3^3	τ_7^2	τ_6^2	τ_5^2	τ_4^2	τ_3^2
multiplicity	1	5	15	1	3	6	10	15.

If one is interested in optimizing the bounds (1.1) it is a tedious task to enumerate all hypertrees.

3. Permutation hypertrees.

In this section we formalize some properties of the collections of subscripts (1.7) defined in [3]. Let I_{r-1} denote the collection of all $(r-1)$ tuples of integers $\mathbf{i} = (i_1, \ldots, i_{r-1}), 2 \leq i_1 < \ldots < i_{r-1} \leq n$. Let $c\colon I_{r-1} \to \{1, 2, \ldots, n-r+1\}$ be any function satisfying $1 \leq c(\mathbf{i}) \leq i_1 - 1$. Let $\pi = (\pi_1, \pi_2, \ldots, \pi_n)$ be any permutation of $(1, 2, \ldots, n)$ and define the collection of ordered r-tuples

$$E_{c,\pi} = \left\{ \left(\pi_{i_0}, \pi_{i_1}, \ldots, \pi_{i_{r-1}}\right) : \quad \mathbf{i} \in I_{r-1} \quad \text{and} \quad i_0 = c(\mathbf{i}) \right\}.$$

We identify each $(\pi_{i_0}, \pi_{i_1}, \ldots, \pi_{i_{r-1}})$ as an edge of an r-hypergraph whose vertices are $\{1, 2, \ldots, n\}$. When π is the identity the set $E_{c,\pi}$ is a typical collection of subscripts in (1.7) used to define $D_{r+1,n}$. Use of a general π merely allows an arbitrary labelling of the events $\{A_i\}$, or equivalently, a permutation of subscripts.

Theorem 1. *For each c and π the pair $(X, E_{c,\pi})$ is an r-hypertree on* $X = \{1,2,\dots,n\}$.

Proof: Without loss of generality let π be the identity. Suppose we have verified that $(X, E_{c,\pi})$ is an r-hypertree for those values of r and n satisfying $1 \leq n \leq N$, $1 \leq r \leq n$. Consider the vertex set $X = \{1,2,\dots,N+1\}$. If $r = N+1$ this theorem is trivially true. Assume then that $r \leq N$. Those subsets of X which do not contain the point $x = N+1$ are of the form $(i_0, i_1, \dots, i_{r-1})$ where $2 \leq i_1 < \dots < i_{r-1} \leq N$, $i_0 = c(\mathbf{i})$, and, thus, form the edges of an r-hypertree on $\{1,2,\dots,N)$; while those subsets of X which do contain x are of the form $(i_0, i_1, \dots, i_{r-2}, x)$ where $2 \leq i_1 < \dots < i_{r-2} \leq N$, $i_0 = c(\mathbf{i})$, and, thus, with x removed are the edges of an $(r-1)$-hypertree on $X - \{x\}$. Hence, x is a terminal vertex in the sense of part (c) in the definition of a hypertree, forcing $(X, E_{c,\pi})$ to have the recursive structure defining an r-hypertree on X. This completes the induction step and the proof. ∎

Definition: We call $(X, E_{c,\pi})$ a *permutation* hypertree and denote it by $\tau_{c,\pi}$.

Tomescu [6] proves the following properties, by induction.

(i) Every r-hypertree of order n has $\binom{n-1}{r-1}$ edges.

(ii) If $d(x)$ is the number of edges containing x then $d(x) \geq \binom{n-2}{r-2}$ and $d(x) = \binom{n-2}{r-2}$ if and only if x is a terminal vertex.

For permutation hypertrees these properties are intuitive. Assume without loss of generality that π is the identity. For $\tau_{c,\pi}$ property (i) is obvious because I_{r-1} has $\binom{n-1}{r-1}$ elements. As for (ii) if $x \in \{2,3,\dots,n\}$ then x lies in exactly $\binom{n-2}{r-2}$ elements of I_{r-1} while additionally x may be determined by the function c as a possible choice for i_0 and so x lies in at least $\binom{n-2}{r-2}$ edges. But if $x = 1$ then any $(r-1)$-tuple in I_{r-1} with $i_1 = 2$ forces $c(\mathbf{i}) = 1 = x$ and the corresponding edge must, therefore, contain x. So again $d(x) \geq \binom{n-2}{r-2}$.

Two other properties are:

(iii) $d(\pi_{n-r+2}) = d(\pi_{n-r+3}) = \dots = d(\pi_n) = \binom{n-2}{r-2}$ and so a permutation r-hypertree contains at least $r-1$ terminal vertices. This follows by noting that $c(\mathbf{i}) \notin \{n-r+2,\dots,n\}$ so vertex j $(n-r+2 \leq j \leq n)$ lies in an edge $(c(\mathbf{i}), \mathbf{i})$ if and only if j is one of the subscripts comprising the elements of $\mathbf{i}$. Removing j from the set $\{2,3,\dots,n\}$ leaves $n-2$ integers of which $r-2$ must be selected, in addition to j, to complete $\mathbf{i}$. Thus, $d(\pi_j) = \binom{n-2}{r-2}$.

(iv) Every 2-hypertree is a permutation hypertree and, hence, a permutation 2-hypertree contains at least two, not at least $r-1$ $(= 1)$ as when $r \geq 3$, terminal vertices.

Theorem 2. *For $r \geq 3$ and $n \geq r+2$ there exist permutation r-hypertrees with exactly s terminal vertices, for any $r-1 \leq s \leq n-1$.*

Proof: Note the anomalous cases $r = 2$, and $n = r+1$ where $r \leq s \leq n-1$. The proof given, counting edges and vertices, explains why the pattern breaks down for these values. First let $s = r-1$ and without loss of generality let π be the identity. Introduce the term "facet" to refer to a subset of size $r-1$ in an r-uniform hypergraph. There are $\binom{n-1}{r-1}$ facets in I_{r-1}, represented as ordered sequences $\mathbf{i}$, to each of which is adjoined another vertex $c(\mathbf{i})$ to form all the edges $(c(\mathbf{i}), \mathbf{i})$ of a general permutation r-hypertree. Each of the vertices $\{2, 3, \dots, n\}$ appears in exactly $\binom{n-2}{r-2}$ of these facets. If $i_1 = 2$ then $c(\mathbf{i})$ must be 1 resulting in $\binom{n-2}{r-2}$ edges of the form $(1, c(\mathbf{i}))$ in each of which both vertices 1 and 2 appear. Thus, counting appearances of vertices in these $\binom{n-2}{r-2}$ edges and the remaining $\binom{n-1}{r-1} - \binom{n-2}{r-2} = \binom{n-2}{r-1}$ facets in I_{r-1} yet to be completed, we see that every vertex $\{1, 2, \dots, n\}$ has already appeared $\binom{n-2}{r-2}$ times. Whenever we adjoin a vertex to complete a facet into an edge, the vertex selected, as a consequence of appearing in more than the permitted $\binom{n-2}{r-2}$ edges cannot be a terminal vertex. As the $r-1$ points $\{n-r+2, \dots, n\}$ are terminal vertices and since $c(\mathbf{i}) \notin \{n-r+2, \dots, n\}$ then in order to achieve exactly $r-1$ terminal vertices each of the points $\{1, 2, \dots, n-r+1\}$ must be selected at least once as a value of $c(\mathbf{i})$, $\mathbf{i}$ ranging over all remaining $\binom{n-2}{r-1}$ facets.

When $r = 2$ there are $\binom{n-2}{r-1} = n-2$ facets and $n-r+1 = n-1$ vertices and so at least one vertex cannot be adjoined verifying that a tree must have at least 2 terminal vertices (which is well known and only included for completeness of the argument). There are $\binom{n-j}{r-1}$ facets to which j, $2 \leq j \leq n-r+1$, can be adjoined, and note that vertices 1 and 2 share the same such facets since we have considered those with $i_1 = 2$. This quantity exceeds one unless $n-j = r-1$ which, together with $2 \leq j \leq n-r+1$ forces $n = r+1$ (and, consequently, $j = 2$ as the only permissable value), again leaving only one facet for the two vertices $1, 2$. Thus, all permutation hypertrees with $n = r+1$ have at least r terminal vertices. In the remaining cases, that is $r \geq 3$ and $n \geq r+2$, there are always enough facets left to include each vertex $\{1, 2, \dots, n-r+1\}$ as a possible value for $c(\mathbf{i})$; consequently, it is possible to construct permutation r-hypertrees with exactly $r-1$ terminal vertices.

For general s we merely omit some vertices, say $\{n-s+1, \dots, n-r+1\}$, for definiteness, as possible values for $c(\mathbf{i})$ and so achieve exactly s terminal vertices for each $r-1 \leq s \leq n-1$. ∎

Theorem 3. *For $r \geq 3$ there exist hypertrees containing only one terminal vertex.*

Proof: Let $X = \{1,2,3,4,5,6\}$, $r = 3$, and $E = \{(1,2,3), (1,2,4), (1,2,5), (1,2,6), (1,3,5), (1,4,5), (2,3,6), (3,4,5), (3,4,6), (4,5,6)\}$. The edges containing the point 6 are $(1,2,6),(2,3,6),(3,4,6)$, and $(4,5,6)$. Removing 6 from each edge leaves $(1,2),(2,3),(3,4)$, and $(4,5)$. These are the edges of a tree on $\{1,2,3,4,5\}$. The edges not containing 6 are $(1,2,3)$, $(1,2,4)$, $(1,2,5)$, $(1,3,5)$, $(1,4,5)$, and $(3,4,5)$, and these form the edges of a permutation 3-hypertree on $\{1,2,3,4,5\}$ where $\pi = (21534)$ and $c((2,3)) = 1$, $c((2,4)) = 1, c((2,5)) = 1, c((3,4)) = 2, c((3,5)) = 1$, and $c((4,5)) = 3$. Hence, (X,E) is, by definition, a 3-hypertree with a terminal vertex 6. A check verifies that 6 is the only terminal vertex since all other vertices lie in at least five edges, while a terminal vertex must lie in exactly $\binom{n-1}{r-2} = \binom{4}{1}$ edges. ∎

Corollary. *For $r \geq 3$ there exist hypertrees which are not permutation hypertrees.*

4. Final comments.

There appears to be scant literature on hypertrees. Tomescu's recursive definition was designed specifically [personal communication] to extend Hunter's inequality (1.2) to higher degree bounds. Permutation hypertrees form a proper subclass for degree $r \geq 3$ and when $n \geq 6$, as shown by our example in Theorem 3 display an interesting contrast, when $r \geq 3$, between hypertrees, which can have only one terminal vertex, and permutation hypertrees which must have at least $r-1$.

There are other graph theoretic structures having tree-like properties, such as k-matroid trees, a generalization of hypertrees, for which properties (i) and (ii) hold [1].

In the probabilistic setting it is of interest to maximize (1.1). For degree three and higher the issue of finding an efficient algorithm is open for both hypertrees and permutation hypertrees.

For an r-uniform hypergraph H on $\{1,2,\ldots,n\}$ the degree of a subset S in H, denoted by $d_H(S)$, is the number of edges of H which contain S as a subset. An r-uniform hypergraph H is called prunable if there exists an ordering $E_1, E_2, \ldots, E_m$ of the edges of H such that there exists a facet $f_j \in E_j$ with

$$d_{H-\{E_1,\ldots,E_{j-1}\}}(f_j) = 1. \tag{4.1}$$

Theorem 4. *Every hypertree is prunable.*

Proof: When $r = 2$ this is a basic property of trees and when $n = r$ there is nothing to prove. Suppose the theorem is true for all $(r-1)$-hypertrees and all r-hypertrees of order $n-1$. Let x be a terminal vertex of an r-hypertree $H = (X,E)$ where $E_1,\ldots,E_m$ are all the edges containing x. By definition $E_1-\{x\},\ldots,E_m-\{x\}$ form all the edges of H^X, an $(r-1)$-hypertree of order $n-1$ with vertex set $X-\{x\}$, which, by induction, is prunable. Let $E_1-\{x\},\ldots,E_m-\{x\}$ be the order in which these edges must be pruned. Thus, there exist facets $f_1,\ldots,f_m$ where $f_j \in E_j-\{x\}$ whose degree in $H^X - \{E_1-\{x\},\ldots,E_{j-1}-\{x\}\}$ is

1. Thus, the degree of $f_j \cup \{x\}$ in $H - \{E_1, \ldots, E_{j-1}\}$ is 1. Consequently, the edges $E_1, \ldots, E_m$ may be pruned in that order from H leaving those edges not containing x which, being the edges of an r-hypertree of order $n-1$ (on vertex set $X - \{x\}$), may also be pruned, by induction. ∎

We close with an observation about the identity (1.4). It is not immediately obvious from [6] that the bounds (1.1) become identities when each of the events $A_1, \ldots, A_n$ is the sample space Ω. However, this is clear from (1.4) since each probability term then becomes unity reproducing the known combinatorial identity

$$\sum_{j=0}^{a}(-1)^j\binom{n}{j} = (-1)^a\binom{n-1}{a}$$

with a probabilistic interpretation.

Acknowledgement

Thanks to Alex Rosa for useful discussions.

References

1. D.A. Grable, *Two packing problems on k-matroid trees*, Auburn University (1989). (preprint).
2. F.M. Hoppe, *Iterating Bonferroni bounds*, Statist. and Probab. Lett. 3 (1985), 121–125.
3. F.M. Hoppe, and E. Seneta, *A Bonferroni-type identity and permutation bounds*, Internat. Statist. Rev. 58 (1991), 253-261.
4. D. Hunter, *An upper bound for the probability of a union*, J. Appl. Probab. 13 (1976), 597–603.
5. E. Seneta, *Degree, iteration, and permutation in improving Bonferroni-type bounds*, Austral. J. Statist. 30A (1988), 27–38.
6. I. Tomescu, *Hypertrees and Bonferroni inequalities*, J. Comb. Theor. B41 (1986), 209–217.