Statistics & Probability Letters 3 (1985) 121–125
North-Holland June 1985

ITERATING BONFERRONI BOUNDS

Fred M. HOPPE

Department of Statistics, The University of Michigan, Ann Arbor, Michigan 48109-1027, USA

Received January 1985

Abstract: A general method is presented for generating improved Bonferroni bounds whenever an improvement on Booles inequality holds.

AMS 1980 Subject Classification: 62E99.

Keywords: Bonferroni inequalities.

1. Introduction

If $A = \cup_1^n A_i$ where $\{A_i\}$ are arbitrary events then by the principle of inclusion–exclusion

$$P(A) = S_1 - S_2 + S_3 - \cdots + (-1)^{n-1} S_n$$

where $S_k = \Sigma P(A_{i_1} A_{i_2} \cdots A_{i_k})$ and the sum is taken over all distinct subscripts $1 \leqslant i_1 < i_2 < \ldots < i_k \leqslant n$. For each m the sum of the first m odd (even) terms provide upper (lower) Bonferroni bounds

$$P(A) \leqslant S_1, \quad P(A) \geqslant S_1 - S_2, \quad P(A) \leqslant S_1 - S_2 + S_3, \quad \ldots .$$

The usual proof (as presented in Feller (1968)) uses a combinatorial identity and gives little insight into the essential simplicity of these inequalities.

We present an elementary proof which has two appealing properties. First, each bound is obtained from its predecessor and ultimately, therefore, it is the first (also called Boole's) inequality

$$P(A) \leqslant \Sigma P(A_i) \tag{1}$$

which underlies all the bounds. Secondly, if an upper bound improving on (1) is available then by the same method tighter successive lower and upper bounds may be obtained. This technique is illustrated by parlaying an upper bound of Worsley (1982) into a lower bound. This lower bound is then calculated for a runs problem in Schwager (1984) in which the usual lower Bonferroni bounds fail to perform adequately.

2. Methodology

We begin with the equation

$$P(A) = P(A_1) + P(A_2 \bar{A}_1) + \cdots + P(A_n \bar{A}_1 \bar{A}_2 \cdots \bar{A}_{n-1}).$$

By expanding the general term and using (1) we get

$$\begin{aligned} P(A_i \bar{A}_1 \cdots \bar{A}_{i-1}) &= P(A_i) - P(A_i A_1 \cup A_i A_2 \cup \cdots \cup A_i A_{i-1}) \\ &\geqslant P(A_i) - [P(A_i A_1) + \cdots + P(A_i A_{i-1})] \end{aligned} \tag{2}$$

and then collecting terms we are led to

$$P(A) \geqslant \Sigma P(A_i) - \sum_{j<i} P(A_i A_j), \tag{3}$$

the second Bonferroni inequality. For the next one we again expand the general term as above but this time we use (3), which we have just proven, to derive the second line of (2). This gives

$$P(A_i \bar{A}_1 \cdots \bar{A}_{i-1}) \leqslant P(A_i) - \left[\sum_{j<i} P(A_i A_j) - \sum_{k<j<i} P(A_i A_j A_k) \right]$$

which results in

$$P(A) \leqslant \Sigma P(A_i) - \sum_{j<i} P(A_i A_j) + \sum_{k<j<i} P(A_i A_j A_k).$$

In this fashion each bound is obtained from the preceding one.

3. An improved lower bound

Worsley (1982) presents the following sharpening of (1). Represent the events $\{A_i\}$ as vertices $\{v_i\}$ of a graph G where v_i and v_j are connected with an edge e_{ij} if and only if $A_i A_j \neq \emptyset$. Let H be a subgraph of G. Then

$$P(A) \leqslant \Sigma P(A_i) - \sum_{\{e_{ij} \in H\}} P(A_i A_j) \tag{4}$$

if and only if H is a tree. If G is a tree and $H = G$ then there is equality in (4).

From (4) Worsley obtains bounds derived by Kounias (1968) and Kwerel (1975) as well as the interesting corollary

$$P(A) \leqslant \sum_{i=1}^{n} P(A_i) - \sum_{i=1}^{n-1} P(A_i A_{i+1}). \tag{5}$$

We now present an improved lower bound based on (5).

Theorem

$$P(A) \geqslant \sum_{i=1}^{n} P(A_i) - \sum_{m<i} P(A_i A_m) + \sum_{i=3}^{n} \sum_{m=1}^{i-2} P(A_m A_{m+1} A_i). \tag{6}$$

Proof. Consider what happens if (5) is used in place of (1) in the argument of Section 2. We get

$$\begin{aligned} P(A_i \bar{A}_1 \bar{A}_2 \ldots \bar{A}_{i-1}) &= P(A_i) - [P(A_i A_1 \cup \cdots \cup A_i A_{i-1})] \\ &\geqslant P(A_i) - \sum_{m=1}^{i-1} P(A_i A_m) + \sum_{m=1}^{i-2} P(A_m A_{m+1} A_i) \end{aligned}$$

which results in (6).

It is possible to obtain lower bounds based on (4) and on the upper bounds of Kounias and Kwerel but we have chosen to focus on (5) for simplicity of expression.

4. Numerical example

Schwager (1984) considers a sequence $\{X_1, X_2, \ldots, X_n\}$ of independent and identically distributed Bernoulli random variables where $p = P[X_1 = 1]$ and $1 - p = P[X_1 = 0]$. Given k let $N = n - k + 1$, $A_i = \{X_i = X_{i+1} = \cdots = X_{i+k-1} = 1\}$ and $A = \cup_{i=1}^{N} A_i$, the event that the sequence contains a string of at least k successive 1's. Schwager computes the first two upper bounds S_1, S_3, the first Worsley bound S_1^* (the right-hand side of (5)), the first two lower bounds S_2, S_4 and compares them with the exact values of $P(A)$ for $p = 0.5$, 0.6, $k = 10$, 15, 20 and $n = 100$, 300, 500, 1000, 1500. He shows that S_1^* approximates $P(A)$ very well in most of these cases and is always much better than S_1 or S_3. The bounds S_2 and S_4, however, are all poor and in many cases negative.

We have calculated the right-hand side of (6), denoted by S_2^*, and in all the above cases S_2^* offers improvement over the lower bounds similar to what S_1^* does for the upper bounds (as well as providing the

Table 1

n	S_2	S_2^*	Exact	S_1^*
$p = 0.5$, $k = 10$				
100	−0.00106	0.04338	0.04414	0.04492
300	−0.03530	0.12382	0.13351	0.14258
500	−0.10768	0.18519	0.21452	0.24023
1000	−0.45553	0.25517	0.38545	0.48437
1500	−1.04181	0.20593	0.51918	0.72852
$p = 0.5$, $k = 15$				
100	0.00006	0.00133	0.00133	0.00133
300	0.00003	0.00437	0.00437	0.00438
500	−0.00004	0.00738	0.00741	0.00743
1000	−0.00038	0.01484	0.01495	0.01506
1500	−0.00094	0.02219	0.02244	0.02269
$p = 0.5$, $k = 20$				
100	0.000002	0.00004	0.00004	0.00004
300	0.000002	0.00013	0.00013	0.00013
500	0.000002	0.00023	0.00023	0.00023
1000	0.000001	0.00047	0.00047	0.00047
1500	0.000001	0.00071	0.00071	0.00071
$p = 0.6$, $k = 10$				
100	−0.36660	0.17634	0.20491	0.22372
300	−2.28017	−0.13211	0.51606	0.70745
500	−5.65620	−0.49709	0.70545	1.19118
1000	−20.49455	−4.62941	0.91487	2.40051
1500	−44.47330	−12.41790	0.97540	3.60983
$p = 0.6$, $k = 15$				
100	−0.01898	0.01624	0.01637	0.01646
300	−0.07347	0.05084	0.05278	0.05407
500	−0.13681	0.08190	0.08785	0.09169
1000	−0.33383	0.14408	0.16994	0.18572
1500	−0.58612	0.18415	0.24465	0.27976
$p = 0.6$, $k = 20$				
100	−0.00135	0.00120	0.00121	0.00121
300	−0.00504	0.00412	0.00412	0.00413
500	−0.00880	0.00700	0.00703	0.00706
1000	−0.01841	0.01412	0.01427	0.01437
1500	−0.02836	0.02111	0.02145	0.02168

same degree of accuracy). The computations are presented in Table 1 which repeats the relevant parts of Table 1 of Schwager (1984) as well as the values of S_2^*.

S_2^* is computed as follows. $S_2^* = S_2 + \Delta_2$ where

$$\Delta_2 = \sum_{i=3}^{N} \sum_{m=1}^{i-2} P(A_m A_{m+1} A_i).$$

If $m + k < i$ then $P(A_m A_{m+1} A_i) = p^{2k+1}$ while if $m + k \geqslant i$ then $P(A_m A_{m+1} A_i) = p^{k+i-m}$. Thus

$$\Delta = \sum_{i=3}^{k+1} \sum_{m=1}^{i-2} P(A_m A_{m+1} A_i) + \sum_{i=k+2}^{N} \sum_{m=1}^{i-k-1} P(A_m A_{m+1} A_i) + \sum_{i=k+2}^{N} \sum_{m=i-k}^{i-2} P(A_m A_{m+1} A_i)$$
$$= \sum_{i=3}^{k+1} \sum_{m=1}^{i-2} p^{i+k-m} + \sum_{i=k+2}^{N} \sum_{m=1}^{i-k-1} p^{2k+1} + \sum_{i=k+2}^{N} \sum_{m=i-k}^{i-2} p^{k+i-m} = a + b + c$$

where

$$a = p^k\left[(k-1)p^2 - (p^3 - p^{k+2})/(1-p)\right]/(1-p), \quad b = p^{2k+1}(N-k-1)(N-k)/2$$

and

$$c = (p^{k+2} - p^{2k+1})(N-k-1)/(1-p).$$

These complicated expressions fail to expose what S_2^* really is. Observe that

$$\Delta_2 = \sum_{i=3}^{N} \sum_{m=1}^{i-2} P(A_m A_{m+1} A_i) = p \sum_{i=3}^{N} \sum_{m=1}^{i-1} P(A_{m+1} A_i)$$
$$= p \sum_{i=2}^{N-1} \sum_{m=1}^{i-1} P(A_m A_i) \quad \text{(by stationarity)}$$
$$\cong p \sum_{i=2}^{N} \sum_{m=1}^{i-1} P(A_m A_i) \quad \text{(approximately)}$$
$$= p(S_1 - S_2).$$

Thus $S_2^* = S_2 + p(S_1 - S_2) = pS_1 + (1-p)S_2$, a linear combination of S_1 and S_2.

It is possible to derive a second upper bound by the method of Section 3 namely

$$P(A) \leqslant S_3^* = S_3 - \Delta_3$$

where

$$\Delta_3 = \sum_{r=4}^{N} \sum_{i=3}^{r-1} \sum_{m=1}^{i-2} P(A_m A_{m+1} A_i A_r).$$

As in the previous paragraph we can show that

$$\Delta_3 \cong p(S_3 - S_2)$$

and $S_3^* \cong pS_2 + (1-p)S_3$. Unfortunately this bound does not improve on S_1^*.

References

Feller, W. (1968), *An Introduction to Probability Theory and its Applications, Vol. I* (John Wiley & Sons, New York, 3rd ed.).

Kounias, E.G. (1968), Bounds for the probability of a union, *Ann. Math. Statist.* **39**, 2154–2158.

Kwerel, S.M. (1975), Most stringent bounds on aggregated probabilities of partially specified dependent probability systems, *J. Amer. Statist. Assoc.* **70**, 472–479.

Schwager, S.J. (1984), Bonferroni sometimes loses, *Amer. Statist.* **38**, 192–197.

Worsley, K.J. (1982), An improved Bonferroni inequality and applications, *Biometrika* **69**, 297–302.