Available online at www.sciencedirect.com

SCIENCE DIRECT®

Discrete Mathematics 306 (2006) 526–530

DISCRETE MATHEMATICS

www.elsevier.com/locate/disc

Communication

Improving probability bounds by optimization over subsets☆

Fred M. Hoppe

Department of Mathematics and Statistics, McMaster University, 1280 Main St. W., Hamilton, Ont., L8S 4K1 Canada

Received 21 August 2005; received in revised form 19 December 2005; accepted 10 January 2006

Communicated by A. Rosa

Abstract

The simple device of maximization over subsets of events can provide substantial improvement over the Dawson–Sankoff degree two lower bound on the probability of a union of events and can also exceed a sharper bound that uses individual and pairwise joint event probabilities developed by Kuai, Alajaji, and Takahara. In each of their examples, the maximized bound achieves the exact probability of the union using a subset of events containing no redundant events of the original set of events.

Keywords: Binomial moments; Bonferroni bound; Dawson–Sankoff bound; Lower probability bound; Redundant events

1. Introduction

One of the outstanding second degree lower bounds on the probability of a union of events $\{A_i,\ i=1,\ldots,n\}$ is due to Dawson and Sankoff [1].

Theorem 1 (*Dawson and Sankoff [1]*).

$$P\left(\bigcup_{i=1}^{n} A_i\right) \geqslant \frac{\theta S_{1,n}^2}{2S_{2,n}+(2-\theta)S_{1,n}} + \frac{(1-\theta)S_{1,n}^2}{2S_{2,n}+(1-\theta)S_{1,n}}, \tag{1}$$

where

$$S_{1,n} = \sum_{i=1}^{n} P(A_i), \qquad S_{2,n} = \sum_{i=2}^{n}\sum_{j=1}^{i-1} P(A_i \cap A_j)$$

are the binomial moments

$$S_{j,n} = E\left[\binom{N}{j}\right] = \sum P(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_j}).$$

☆ Support by NSERC Discovery Grant is gratefully acknowledged.
E-mail address: hoppe@mcmaster.ca.

0012-365X/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.disc.2006.01.009

N is the random variable counting the number of events that occur, the summation is over all integers $1 \leqslant i_1 < i_2 < \cdots < i_j \leqslant n$, and θ is the constant

$$\theta = \frac{2S_{2,n}}{S_{1,n}} - \left\lfloor \frac{2S_{2,n}}{S_{1,n}} \right\rfloor$$

with $\lfloor x \rfloor$ the largest integer in x. The dependence of θ on n, $S_{1,n}$, $S_{2,n}$ is omitted from the notation.

Bound (1) is optimal in the sense that if $g(S_{1,n}, S_{2,n})$ is any other universal lower bound on $P\left(\bigcup_{i=1}^{n} A_i\right)$ that depends on n events only through $S_{1,n}$, $S_{2,n}$ then

$$\frac{\theta S_{1,n}^2}{2S_{2,n} + (2-\theta)S_{1,n}} + \frac{(1-\theta)S_{1,n}^2}{2S_{2,n} + (1-\theta)S_{1,n}} \geqslant g(S_{1,n}, S_{2,n}).$$

Optimality is proven by exhibiting, for any fixed, feasible values of $\{n, S_{1,n}, S_{2,n}\}$ a sample space and set of events $\{A_i^\star\}$ having the same values for the binomial moments as $S_{1,n}$, $S_{2,n}$ and achieving equality in (1), which therefore gives the minimum probability of any union of events consistent with these $\{n, S_{1,n}, S_{2,n}\}$. This type of optimality is called Fréchet optimality since the idea can be traced back to Fréchet [3] (also Section 7 of [6,7]). Proof of optimality is in [1] although not stated explicitly, and also the structure of the $\{A_i^\star\}$ can be read from the proof of minimality there.

Recently, Kuai et al. [5] obtained the following interesting bound that parallels the structure of (1), but in place of $S_{1,n}$, $S_{2,n}$ uses the individual and pairwise joint event probabilities

$$P(A_i),\ P(A_i \cap A_i), \quad 1 \leqslant i < j \leqslant n. \tag{2}$$

Theorem 2 (*Kuai et al. [5]*).

$$P\left(\bigcup_{i=i}^{n} A_i\right) \geqslant \sum_{i=1}^{n} \frac{\theta_i P(A_i)^2}{\sum_{j=1}^{n} P(A_i \cap A_j) + (1-\theta_i)P(A_i)} + \frac{(1-\theta_i)P(A_i)^2}{\sum_{j=1}^{n} P(A_i \cap A_j) - \theta_i P(A_i)}, \tag{3}$$

where

$$\theta_i = \frac{\sum_{j:j\neq i} P(A_i \cap A_j)}{P(A_i)} - \left\lfloor \frac{\sum_{j:j\neq i} P(A_i \cap A_j)}{P(A_i)} \right\rfloor.$$

They prove that (3) is always at least as large as (1) and illustrate improvement with numerical examples. They also prove that (3) improves the following bound of de Caen [2] that also uses (2).

Theorem 3 (*de Caen [2]*).

$$P\left(\bigcup_{i=1}^{n} A_i\right) \geqslant \sum_{i=1}^{n} \frac{P(A_i)^2}{P(A_i) + \sum_{j\neq i} P(A_i \cap A_j)}. \tag{4}$$

The purpose of this note is to observe that these three bounds can be further improved by maximization over subsets. The same systems of events used in [5] demonstrate the increase obtained. In all four cases, the exact probability of the union is obtained by this simple device.

2. Main result

For $m < n$

$$P\left(\bigcup_{i=1}^{n} A_i\right) \geqslant P\left(\bigcup_{i=1}^{m} A_i\right) \tag{5}$$

and the Dawson–Sankoff bound can be applied to the right-hand side of (5) which is based on fewer events. Since the Dawson–Sankoff bound is optimal, it would seem reasonable that the optimal bound based on fewer events should be no better than the optimal bound based on the full set of n events. Interestingly, computation showed this not to be the case, and it is therefore possible to obtain larger, hence better, Dawson–Sankoff bounds for the same union based on *fewer* events. The same strategy can be applied to (3) and (4).

What motivated this work was the observation that all the examples in [5] shared the property that the union of the n events $\{A_i\}$ could be obtained as a union of fewer events. In systems I, II, and III $A_3 \cup A_5 \cup A_6 \subseteq A_1 \cup A_2 \cup A_4$ while in system IV, $A_7 \subseteq A_1 \cup A_2 \cup A_5 \cup A_6$. Thus, events that were already included in a union of fewer events could be removed prior to computing any bounds.

Therefore, let I be a subset of $\{1, 2, \ldots, n\}$ and restrict attention to the events $\{A_i\}$ for $i \in I$. Define

$$S_1(I) = \sum_{i \in I} P(A_i), \quad S_2(I) = \sum \sum_{j < i \in I} P(A_i \cap A_j)$$

and

$$\theta(I) = \frac{2S_2(I)}{S_1(I)} - \left\lfloor \frac{2S_2(I)}{S_1(I)} \right\rfloor.$$

Theorem 4.

(i)

$$P\left(\bigcup_{i=1}^{n} A_i\right) \geqslant \max_I \left(\frac{\theta(I) S_1^2(I)}{2S_2(I) + (2 - \theta(I)) S_1(I)} + \frac{(1 - \theta(I)) S_1^2(I)}{2S_2(I) + (1 - \theta(I)) S_1(I)} \right), \tag{6}$$

(ii)

$$P\left(\bigcup_{i=1}^{n} A_i\right) \geqslant \max_I \left(\sum_{i \in I} \frac{\theta_i(I) P(A_i)^2}{\sum_{j \in I} P(A_i \cap A_j) + (1 - \theta_i(I)) P(A_i)} + \frac{(1 - \theta_i(I)) P(A_i)^2}{\sum_{j \in I} P(A_i \cap A_j) - \theta_i(I) P(A_i)} \right), \tag{7}$$

where for $i \in I$

$$\theta_i(I) = \frac{\sum_{j \in I, j \neq i} P(A_i \cap A_j)}{P(A_i)} - \left\lfloor \frac{\sum_{j \in I, j \neq i} P(A_i \cap A_j)}{P(A_i)} \right\rfloor,$$

(iii)

$$P\left(\bigcup_{i=1}^{n} A_i\right) \geqslant \max_I \sum_{i \in I} \frac{P(A_i)^2}{P(A_i) + \sum_{j \in I, j \neq i} P(A_i \cap A_j)}. \tag{8}$$

3. Examples

To examine the possible effect of maximization over subsets of events, we present in Table 1 the bounds considered in [5] together with the new bounds (6)–(8). (Note that the corresponding entries in the last row of Table 5 of [5] for system IV, including $P(\bigcup_{i=i}^{n} A_i)$, appear to be systematically too large by an amount 0.0002.) There are four systems of $n = 6$ events in the first three systems and $n = 7$ in the last one.

Observe that (3) is larger than (1) or (4) in all cases. However, when the Dawson–Sankoff bound is optimized over subsets $I \subseteq \{1, 2, \ldots, n\}$ in (6) it exceeds (3) in all four systems and in fact achieves the exact probability of the union.

Table 1
Comparison of bounds

System	n	$P\left(\bigcup_{i=i}^{n} A_i\right)$	(6)	(7)	(8)	DS (1)	Kuai et al. [5]	de Caen [2]
I	6	0.7890	0.7890	0.7890	0.7332	0.7007	0.7247	0.7087
II	6	0.6740	0.6740	0.6740	0.6296	0.6150	0.6227	0.6154
III	6	0.7890	0.7890	0.7890	0.7349	0.6933	0.7222	0.7048
IV	7	0.9687	0.9687	0.9687	0.8978	0.8879	0.8909	0.8757

The same must then be true for the optimized versions of the Kuai et al. bound. Moreover, in both (6) and (7) the optimized bounds are achieved for the same subsets I. These are:

system I: $I = \{1, 2, 4\}$, $I = \{1, 2, 5\}$,
system II: $I = \{1, 2, 4\}$, $I = \{1, 2, 3, 4\}$,
system III: $I = \{1, 2, 4\}$, $I = \{1, 2, 3, 4\}$,
system IV: $I = \{2, 3, 5, 6\}$, $I = \{2, 3, 4, 5, 6\}$.

So, in system I, $\bigcup_{i=1}^{6} A_i = A_1 \cup A_2 \cup A_4$ and the exact probability of the union can be obtained by applying either the Dawson–Sankoff or the Kuai et al. bound to either $A_1 \cup A_2 \cup A_4$ or $A_1 \cup A_2 \cup A_5$. Similarly, the union in the other systems can be obtained with fewer events and the resulting bounds are the exact probabilities. Thus, in systems II and III the exact probability can be obtained using either $A_1 \cup A_2 \cup A_4$ or $A_1 \cup A_2 \cup A_3 \cup A_4$. Although A_3 is redundant in both of these systems, its inclusion does not decrease the value of the two bounds. Finally, in system IV, the exact value can be achieved with $A_2 \cup A_3 \cup A_5 \cup A_6$ or $A_2 \cup A_3 \cup A_4 \cup A_5 \cup A_6$ with A_4 redundant.

4. Final remarks

For any set of events $A_1, A_2, \ldots, A_n$ for which equality holds in (1)—and on account of Fréchet optimality such a set of events exists—equality must also hold in (3), since (3) is always at least as large as (1). In all four of the examples considered from Kuai et al. [5], such equality is achieved using a subset which contains no redundant events of the original set of events. Inasmuch as equality in (1) and (3) is actually achieved with given sets, this suggests that these examples may be rather unrepresentative of a general situation, and raises obvious questions about their structure. More generally, there are questions about the effect on bounds of redundant events, and about the effect of additional nonredundant events. These will be addressed in a more extensive study.

We note that Kounias [4] has suggested maximizing the second-degree Bonferroni lower bound over subsets and actually cited [1] for a numerical comparison, but made no mention of maximizing it over subsets.

For the second-degree Bonferroni lower bound, inclusion of redundant events cannot improve the bound, as we now show.

Theorem 5. *Suppose* $A_n \subseteq \bigcup_{i=1}^{n-1} A_i$ *and is therefore redundant. Then*

$$S_{1,n} - S_{2,n} \leqslant S_{1,n-1} - S_{2,n-1}.$$

Proof. $A_n \subseteq \bigcup_{i=1}^{n-1} A_n \cap A_i$ and thus $P(A_n) \leqslant \sum_{i=1}^{n-1} P(A_n \cap A_i)$. As a result

$$\begin{aligned} S_{1,n} - S_{2,n} &= S_{1,n-1} + P(A_n) - \left(S_{2,n-1} + \sum_{i=1}^{n-1} P(A_n \cap A_i)\right) \\ &= S_{1,n-1} - S_{2,n-1} + \left(P(A_n) - \sum_{i=1}^{n-1} P(A_n \cap A_i)\right) \\ &\leqslant S_{1,n-1} - S_{2,n-1}. \qquad \square \end{aligned} \tag{9}$$

Acknowledgment

Thanks to a referee for the helpful suggestions.

References

[1] D.A. Dawson, D. Sankoff, An inequality for probabilities, Proc. Amer. Math. Soc. 18 (1967) 504–507.
[2] D. de Caen, A lower bound on the probability of a union, Discrete Math. 169 (1997) 217–220.
[3] M. Fréchet, Généralisations du théorème des probabilités totales, Fund. Math. 25 (1935) 379–387.
[4] E. Kounias, Bounds for the probability of a union, with applications, Ann. Math. Statist. 39 (1968) 2154–2158.
[5] H. Kuai, F. Alajaji, G. Takahara, A lower bound on the probability of a finite union of events, Discrete Math. 215 (2000) 147–158.
[6] E. Seneta, On the history of the strong law of large numbers and Boole's inequality, Historia Math. 19 (1992) 24–39.
[7] E. Seneta, T. Chen, On explicit and Fréchet-optimal lower bounds, J. Appl. Probab. 39 (2002) 81–90.