SOME OPTIMAL BIVARIATE BONFERRONI-TYPE BOUNDS

JANOS GALAMBOS AND YUAN XU

(Communicated by George C. Papanicolaou)

ABSTRACT. Let $A_1, A_2, \dots, A_n$ and $B_1, B_2, \dots, B_m$ be two sets of events on a probability space. Let X_n and Y_m be the number of those A_j and B_s , respectively, that occur. Let $S_{k,t}$ be the (k, t) th binomial moment of the vector (X_n, Y_m) . We establish optimal bounds on $P(X_n \geq 1, Y_m \geq 1)$ by means of linear combinations of $S_{1,1}, S_{2,1}, S_{1,2}$, and $S_{2,2}$. Optimal lower bounds are also determined when only $S_{1,1}, S_{2,1}$, and $S_{1,2}$ are utilized.

1. INTRODUCTION

Let $A_1, A_2, \dots, A_n$ and $B_1, B_2, \dots, B_m$ be two sets of events on a probability space. Let $X = X_n$ and $Y = Y_m$ be the number of those A_j and B_s , respectively, that occur. Set $S_{0,0} = 1$ and for integers $k \geq 1$ and $t \geq 1$,

$$(1) \quad S_{k,t} = \sum P(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k} \cap B_{j_1} \cap B_{j_2} \cap \dots \cap B_{j_t}),$$

where the summation is over all subscripts $1 \leq i_1 < i_2 < \dots < i_k \leq n$ and $1 \leq j_1 < j_2 < \dots < j_t \leq m$. It can easily be verified, by turning to indicators, that $S_{k,t}$, $k \geq 1$, $t \geq 1$, is a binomial moment, namely,

$$(2) \quad S_{k,t} = E \left[\binom{X}{k} \binom{Y}{t} \right].$$

In order to prove (2) take the indicators of the intersections appearing in (1). The expectation of their sum $J_{k,t}$ evidently is $S_{k,t}$, while $J_{k,t} = \binom{X}{k} \binom{Y}{t}$.

While the univariate case (having just one block of events) has a vast literature, in the bivariate setting very little is known in terms of estimating $P(X_n \geq 1, Y_m \geq 1)$ by means of linear combinations of (some) $S_{k,t}$. In fact, the only results known to us are the identity in Fréchet's book [2], its extensions to inequalities by Éva Galambos [3] and Meyer [8], and the recent studies by Lee [7] (for applications of the inequalities of Éva Galambos and Meyer, see [5, Chapter 5]). None of these works deals with the question of optimality of a given inequality. Our aim is to establish optimal inequalities in terms of $S_{k,t}$, $1 \leq k, t \leq 2$, thus extending the univariate result of Dawson and Sankoff [1]

Received by the editors June 2, 1991.

1991 *Mathematics Subject Classification*. Primary 60E15; Secondary 60C05.

Key words and phrases. Two sets of events, bivariate binomial moments, Bonferroni-type bounds, optimal.

(for an optimality study of the inequality of Dawson and Sankoff, see [4]). Once such optimal inequalities are available, one can use iteration methods on the line of Galambos and Xu [6] for generating further inequalities, now involving $S_{k,t}$ with arbitrary values of k and t .

The actual results are formulated in the next section. We conclude this section by pointing out that one can obtain bounds on $P(X_n \geq 1, Y_m \geq 1)$ in terms of linear combinations of the $S_{k,t}$ by an appeal to univariate cases upon observing that

$$P(X_n \geq 1, Y_m \geq 1) = 1 - P(X_n = 0) - P(Y_m = 0) + P(X_n + Y_m = 0).$$

Such inequalities will not be optimal, however, because in order to get an upper bound, say, for the left-hand side, a combination of upper and lower bounds are needed for the three terms on the right-hand side. In addition, in estimating $P(X_n = 0)$ and $P(Y_m = 0)$, the joint binomial moments $S_{k,t}$, $k \geq 1$, $t \geq 1$, are not used but instead the individual binomial moments of X_n and Y_m , which we denote by $S_{k,0}$ and $S_{0,t}$, respectively (formally, both (1) and (2) apply). Hence, the present study of optimality is opening a new direction for Bonferroni-type inequalities.

2. THE RESULTS

Our results are contained in the following three theorems; there, constant coefficients mean that they do not depend on the events A_j and B_s but may depend on parameters including m and n . We also wish to emphasize that when we speak of bounds on $P(X_n \geq 1, Y_m \geq 1)$ we mean inequalities that are universally valid, that is, on an arbitrary probability space and for all possible choices of the events A_j and B_s .

Theorem 1. *Let a and b be integers with $1 \leq a \leq n-1$ and $1 \leq b \leq m-1$. Set*

$$L_{a,b} = \frac{4S_{1,1}}{(a+1)(b+1)} - \frac{4S_{2,1}}{a(a+1)(b+1)} - \frac{4S_{1,2}}{(a+1)(b+1)} + \frac{4S_{2,2}}{a(a+1)b(b+1)}.$$

We then have for $(a, b) \in D = \{(u, v) : \text{either } u \geq n/2 \text{ or } v \geq m/2\}$

$$(3) \quad P(X_n \geq 1, Y_m \geq 1) \geq L_{a,b}.$$

Furthermore, if, with some constants c_j , $1 \leq j \leq 4$,

$$(4) \quad P(X_n \geq 1, Y_m \geq 1) \geq c_1 S_{1,1} + c_2 S_{2,1} + c_3 S_{1,2} + c_4 S_{2,2}$$

and the right-hand side is not of the form of an $L_{a,b}$, even though $L_{a,b}$ satisfies $L_{a,b} \leq L_{u,v}$ for some $(u, v) \in D$, then there is a pair (a, b) of integers such that

$$(5) \quad L_{a,b} \geq c_1 S_{1,1} + c_2 S_{2,1} + c_3 S_{1,2} + c_4 S_{2,2}$$

and, of course, (3) applies.

Theorem 2. *Let*

$$U_{n,m} = S_{1,1} - (2/n)S_{2,1} - (2/m)S_{1,2} + (4/nm)S_{2,2}.$$

Then

$$(6) \quad P(X_n \geq 1, Y_m \geq 1) \leq U_{n,m}$$

and $U_{n,m}$ is the best upper bound among all upper bounds of the form $d_1S_{1,1} + d_2S_{2,1} + d_3S_{1,2} + d_4S_{2,2}$, where the coefficients d_j , $1 \leq j \leq 4$, are constants.

The nature of the optimal bounds is completely different if we seek these bounds among the linear combinations of $S_{1,1}$, $S_{1,2}$, and $S_{2,1}$ only.

Theorem 3. Let u and v be integers with $2 \leq u \leq n-1$ and $2 \leq v \leq m-1$. Then the optimal lower bounds among the linear combinations of $S_{1,1}$, $S_{1,2}$, and $S_{2,1}$ are given by the family of inequalities

$$(7) \quad P(X_n \geq 1, Y_m \geq 1) \geq \frac{1}{uv} \left(3 - \frac{1}{u} - \frac{1}{v} \right) S_{1,1} - \frac{2S_{2,1}}{u^2v} - \frac{2S_{1,2}}{uv^2}.$$

Remarks. Theorem 1 becomes the result of Dawson and Sankoff [1] and Galambos [4] if we choose each B_j as the sample space Ω and $b = m-1$. In addition, inequality (6) is due to Lee [7] but without its optimality property.

3. PROOFS

Proof of Theorem 1. We turn to indicators. Let

$$(8) \quad I = I(X \geq 1, Y \geq 1) = \begin{cases} 1 & \text{if } X \geq 1 \text{ and } Y \geq 1, \\ 0 & \text{otherwise} \end{cases}$$

and

$$(9) \quad L_{a,b}(X, Y) = \frac{4XY}{(a+1)(b+1)} - \frac{2X(X-1)Y}{a(a+1)(b+1)} - \frac{2XY(Y-1)}{(a+1)b(b+1)} + \frac{X(X-1)Y(Y-1)}{a(a+1)b(b+1)}.$$

Then $E(I) = P(X_n \geq 1, Y_m \geq 1)$ and $E(L_{a,b}(X, Y)) = L_{a,b}$, and thus in order to prove (3), it suffices to prove

$$(10) \quad L_{a,b}(X, Y) \leq I.$$

Now, since both sides of (10) equal zero if either $X = 0$ or $Y = 0$, (10) is entailed by the inequality

$$(11) \quad L_{a,b}(X, Y) \leq 1 \quad \text{if } X \geq 1 \text{ and } Y \geq 1, (a, b) \in D,$$

which can also be written as

$$(11a) \quad \frac{X(2a+1-X)}{a(a+1)} \frac{Y(2b+1-Y)}{b(b+1)} \leq 1, \quad X \geq 1, Y \geq 1, (a, b) \in D.$$

The two numerators in (11a) are parabolas whose maxima, for integer variables X and Y , are at $X = a$ and $Y = b$, respectively, at which values the left-hand side of (11a) becomes one. This proves (11a) and thus (3) as well.

Next, we prove the claimed optimality of (3). Assume that (4) holds but (5) fails for all a and b . That is, if we put

$$(12) \quad T_c(X, Y) = c_1XY + c_2X(X-1)Y/2 + c_3XY(Y-1)/2 + c_4X(X-1)Y(Y-1)/4,$$

then by (4) and (8),

$$(13) \quad T_c(X, Y) \leq 1 \quad \text{for } X \geq 1, Y \geq 1,$$

and the failure of (5) for all a and b means that there is a vector (X^0, Y^0) such that

$$(14) \quad L_{a,b}(X^0, Y^0) < T_c(X^0, Y^0) \quad \text{for all } a \text{ and } b.$$

Now, as we observed at (11a),

$$L_{X^0,b}(X^0, Y) = Y(2b + 1 - Y)/b(b + 1).$$

Thus, by choosing $a = X^0$ (or $a = X^0 - 1$ if $X^0 = n$) in (14), the combination of (13) and (14) contradicts the univariate result on optimality by Galambos [4] (see also [5, pp. 25–26]). This completes the proof of Theorem 1.

Proof of Theorem 2. Once more, we use indicator variables. Let

$$U_{n,m}(X, Y) = XY - \frac{X(X-1)Y}{n} - \frac{XY(Y-1)}{m} + \frac{X(X-1)Y(Y-1)}{nm}.$$

Since $U_{n,m}(0, Y) = U_{n,m}(X, 0) = 0$, for (6) we have to prove that

$$(15) \quad U_{n,m}(X, Y) \geq 1 \quad \text{for } 1 \leq X \leq n, \quad 1 \leq Y \leq m.$$

However,

$$(16) \quad \begin{aligned} U_{n,m}(X, Y) &= \left[X - \frac{X(X-1)}{n} \right] \left[Y - \frac{Y(Y-1)}{m} \right] \\ &= \frac{X(n+1-X)}{n} \frac{Y(m+1-Y)}{m} \end{aligned}$$

and the parabolas in the numerator evidently satisfy

$$\begin{aligned} X(n+1-X) &\geq n \quad \text{if } 1 \leq X \leq n, \\ Y(m+1-Y) &\geq m \quad \text{if } 1 \leq Y \leq m. \end{aligned}$$

This proves (15), which in turn entails (6).

For the optimality statement expressed in (6) we use a direct argument. Let $T_d(X, Y)$ signify the expression in (12) with d_j replacing c_j , $1 \leq j \leq 4$. Then we are seeking those coefficients d_j for which the right-hand side of the inequality

$$(17) \quad 1 \leq T_d(X, Y), \quad 1 \leq X \leq n, \quad 1 \leq Y \leq m, \quad X \text{ and } Y \text{ integers},$$

is as close to one as possible. We must have

$$1 \leq T_d(1, 1) = d_1$$

and

$$1 \leq T_d(X, 1) = d_1 X + d_2 X(X-1)/2 \quad \text{for all } 1 \leq X \leq n.$$

Upon rewriting the second inequality as

$$(d_2/2)X^2 + (d_1 - d_2/2)X - 1 \geq 0 \quad \text{for } 1 \leq X \leq n,$$

we observe that the quadratic expression on the left-hand side is (-1) at $X = 0$, must be nonnegative for $1 \leq X \leq n$, and, in the optimal case, we want to keep it as small as possible. Out of all parabolas, including the degenerated case into a straight line, we must have one with negative leading coefficient ($d_2 < 0$) and that is zero both at $X = 1$ and $X = n$. Therefore, $d_1 = 1$ and $d_2 = -2/n$. By interchanging the roles of X and Y , we get $d_3 = -2/m$. By these values,

(17) in the form $1 \leq T_d(n, m)$ results in the inequality $d_4 \geq 4/nm$, which for the optimal case yields $d_4 = 4/nm$. The proof is complete.

Proof of Theorem 3. The indicator variable version of (7) is

$$\frac{1}{uv} \left(3 - \frac{1}{u} - \frac{1}{v} \right) XY - \frac{X(X-1)Y}{u^2v} - \frac{XY(Y-1)}{uv^2} \leq 1,$$

which can be simplified to

$$(18) \quad \frac{XY}{uv} \left(3 - \frac{X}{u} - \frac{Y}{v} \right) \leq 1, \quad \text{where } 1 \leq X \leq n \text{ and } 1 \leq Y \leq m.$$

In order to prove (18), we introduce the function

$$g(x, y) = \frac{x}{u} \frac{y}{v} \left(a - \frac{bx}{u} - \frac{cy}{v} \right), \quad a, b, c > 0,$$

and seek conditions on the coefficients a, b , and c entailing

$$(19) \quad g(X, Y) \leq 1 \quad \text{for } 1 \leq X \leq n, \quad 1 \leq Y \leq m.$$

Clearly, $g(0, y) = g(x, 0) = 0$ and $g(x, y) < 0$ for large x or y . Consequently, the maximum of $g(x, y)$ can be determined by setting the first partial derivatives equal to zero. We thus have that, at the maximum of $g(x, y)$,

$$\frac{y}{uv} \left(a - b \frac{x}{u} - c \frac{y}{v} \right) - \frac{xy}{uv} \frac{b}{u} = 0$$

and

$$\frac{x}{uv} \left(a - b \frac{x}{u} - c \frac{y}{v} \right) - \frac{xy}{uv} \frac{c}{v} = 0.$$

Upon multiplying the first equation by x and the second one by y and subtracting the resulting equations, we get that at the point of maximum of $g(x, y)$, $xb/u = yc/v$. By substituting this into either of the preceding equations, we get for the critical point $x_0 = au/3b$ and $y_0 = av/3c$. This yields the maximum $g(x_0, y_0) = (a/3)^3(1/bc)$. Hence (19) is valid whenever

$$(19a) \quad (a/3)^3 \leq bc.$$

In particular, we have (18) since there $a = 3$ and $b = c = 1$.

Finally, we prove the optimality of (7) among all lower bounds

$$P(X_n \geq 1, Y_m \geq 1) \geq aS_{1,1} - bS_{2,1} - cS_{1,2}.$$

If the parametric family (7) were not optimal, then we would have coefficients a, b , and c such that

$$aXY - bX(X-1)Y/2 - cXY(Y-1)/2 \leq 1 \quad (1 \leq X \leq n, \quad 1 \leq Y \leq m),$$

which we rearrange in the form

$$(20) \quad XY(A - BX - CY) \leq 1 \quad \text{for } 1 \leq X \leq n, \quad 1 \leq Y \leq m,$$

and, for some X_0 and Y_0 ,

$$(21) \quad \frac{X_0 Y_0}{uv} \left(3 - \frac{X_0}{u} - \frac{Y_0}{v} \right) < X_0 Y_0 (A - BX_0 - CY_0) \quad \text{for all } u, v.$$

Choose $u = X_0$ and $v = Y_0$ and get a contradiction between (20) and (21). The theorem is established.

REFERENCES

1. D. A. Dawson and D. Sankoff, *An inequality for probabilities*, Proc. Amer. Math. Soc. **18** (1967), 504–507.
2. M. Fréchet, *Les probabilités associées à un système d'événements compatibles et dépendants*, Hermann, Paris, 1940.
3. Éva Galambos, *Discussion of probabilistic inequalities by the method of A. Rényi*, Dissertation, L. Eötvös University, Budapest, 1965. (Hungarian)
4. J. Galambos, *Bonferroni inequalities*, Ann. Probab. **5** (1977), 577–581.
5. ———, *The asymptotic theory of extreme order statistics*, 2nd ed., Krieger, Malabar, FL, 1987.
6. J. Galambos and Y. Xu, *A new method for generating Bonferroni-type inequalities by iteration*, Math. Proc. Cambridge Philos. Soc. **107** (1990), 601–607.
7. M. Y. Lee, Personal communication, 1990.
8. R. M. Meyer, *Note on a multivariate form of Bonferroni's inequalities*, Ann. Math. Statist. **40** (1969), 692–693.

DEPARTMENT OF MATHEMATICS, TU 038-16, TEMPLE UNIVERSITY, PHILADELPHIA, PENNSYLVANIA 19122

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF TEXAS AT AUSTIN, AUSTIN, TEXAS 78712
Current address: Department of Mathematics, University of Oregon, Eugene, Oregon 97403