J. Korean Math. Soc. **33** (1996), No. 4, pp. 793–799

OPTIMAL BIVARIATE BONFERRONI–TYPE INEQUALITIES VIA TAKING AVERAGE

MIN-YOUNG LEE

1. Introduction

Let $A_1, A_2, \cdots, A_m$ be a sequence of events on a given probability space and let $X_m(A)$ be the number of those $A's$ which occur. Put $S_0 = 1$ and

$$S_k = \sum P(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}), \, (1 \le k)$$

where the summation is over all subscripts satisfying $1 \le i_1 < i_2 < \cdots < i_k \le m$.

For convenience in some formulae we adopt the convention $S_k = 0$ if $k > m$. By turing to indicator variables we immediately find that

$$S_k = E\left[\binom{X_m(A)}{k}\right] \qquad (0 \le k \le m).$$

Inequalities of the form $\sum_{k=0}^{m} c_k S_k \le P(X_m(A) \ge r) \le \sum_{k=0}^{m} d_k S_k$, where $c_k = c_k(r,m)$ and $d_k = d_k(r,m)$ are constant, possible zero, are called Bonferroni-type inequalities.

In this univariate case, we have proved that
(1)

$$P(X_m(A) \le 1) \le S_1 - \sum_{i<j\le i+2} P(A_i \cap A_j) + \sum_{i=1}^{m-2} P(A_i \cap A_{i+1} \cap A_{i+2}).$$

Received July 10, 1995. Revised February 20, 1996.

1991 AMS Subject Classification: 60E15, 60C05.

Key words: Bonferroni-type inequality, the method of indicators.

This paper was supported by NON DIRECTED RESEARCH FUND, Korea Research Foundation, 1994.

Taking the averages of the above upper bounds over $i = 1, 2, \cdots, m$ of (1), we get the following Bonferroni-type inequality (B-T-I).

$$P(X_m(A) \geq 1) \leq S_1 - \frac{(2m-3)}{\binom{m}{2}} S_2 + \frac{(m-2)}{\binom{m}{3}} S_3$$

This inequality is known that it is the best possible upper bound in terms of S_1, S_2 and S_3 [see Kwerel(1975)].

Now we can extend the univariate B-T-I into the bivariate one. Let $A_1, A_2, \cdots, A_m$ and $B_1, B_2, \cdots, B_n$ be two sequences of events on the same probability space. Let $X = X_m(A)$ and $Y = Y_n(B)$, respectively, denote the numbers of those A_i and B_j which occur. Put $S_{0,0} = 1$ and, for integers r and t, set

$$(2) \qquad S_{r,t} = \sum \sum P(A_{i_1} A_{i_2} \cdots A_{i_r} B_{j_1} B_{j_2} \cdots B_{j_t})$$

where the summation is over all subscripts satisfying $1 \leq i_1 < i_2 < \cdots < i_r \leq m$ and $1 \leq j_1 < j_2 < \cdots < j_t \leq n$, $0 \leq r \leq m$ and $0 \leq t \leq n$ (We abbreviate $A \cap B$ as AB and an empty intersection is the sample space). We can easily prove (2) by using the method of indicators. The number $S_{r,t}$, $1 \leq r$, $1 \leq t$, is called the binomial moment of the vector (X,Y) because

$$(3) \qquad S_{r,t} = E\left[\binom{X}{r}\binom{Y}{t}\right].$$

Although the univariate case has been studied by many authors, in the bivariate case little information is know. In fact, Eva, Galambos [1965] and Meyer [1969] introduced the classical bivariate inequalities and Lee [1992], Galambos and Lee [1992] presented its extensions by using the method of indicators and combinations, and examples of optimal inequalities are shown by Galambos and Xu [1993].

We are interested in bivariate B-T-I which mean bounds by linear combinations of the binomial moments $S_{r,t}$. In particular, we want to establish upper bound of $y_{1,1} = P(X \geq 1 , Y \geq 1)$ which appears in many problems in statatics (See for example Galambos and Lee [1994]). Galambos and Xu have proved that

$$(4) \qquad P(X \geq 1, Y \geq 1) = y_{1,1} \leq S_{1,1} - \frac{2}{m} S_{2,1} - \frac{2}{n} S_{1,2} - \frac{4}{mn} S_{2,2}$$

which insists the best upper bound among all upper bounds of the form $d_1 S_{1,1} + d_2 S_{2,1} + d_3 S_{1,2} + d_4 S_{2,2}$.

When we compare (4) with the classical lower bound

(5)
$$S_{1,1} - S_{1,2} - S_{2,1} \leq P\big((\cup_{i=1}^{m} A_i) \cap (\cup_{j=1}^{n} B_j)\big) = P(X \geq 1, Y \geq 1) = y_{1,1}$$

We see that if we subtract from $S_{1,1}$ all intersections of triples (i.e., $S_{1,2} + S_{2,1}$) we get a lower bound but if only a percentage of these triples are subtracted we get an upper bound (note that $S_{2,2}$ is negative in (4)). This suggests that if, instead of $S_{1,2}$ and $S_{2,1}$, we subtract a restricted number of terms of the form $P(A_i B_j B_k)$ and $P(A_i A_j B_k)$ we shall have an upper bound. The question is that what kind of sums of $P(A_i B_j B_k)$ and $P(A_i A_j B_k)$ will guarantee to have a universal upper bound when these sums are subtracted from $S_{1,1}$. An answer to this question is contained in Theorem 1 of the next section.

2. The Inequlities

We establish some new inequalities in which our ideas are to prove upper bounds using some of terms instead of all terms $S_{1,2}$, $S_{2,1}$ and $S_{2,2}$. Also, we prove optimal upper bounds on $y_{1,1}$ by taking average of new ones.

THEOREM 1.. *Let i and j be integers with $1 \leq i \leq m, 1 \leq j \leq n$. Then*

(6)
$$y_{1,1} \leq S_{1,1} - \max\Big\{ \sum_{i=1}^{m-1} \sum_{j=1}^{n} P(A_i A_{i+1} B_j) + \sum_{j=1}^{n-1} P(A_k B_j B_{j+1}), \sum_{i=1}^{m} \sum_{j=1}^{n-1} P(A_i B_j B_{j+1}) + \sum_{i=1}^{m-1} P(A_i A_{i+1} B_k) \Big\}$$

where A_k and B_k are arbitrary fixed events.

Taking the averages of the above upper bounds over $i = 1, 2, ..., m$, $j = 1, 2, ..., n$, we get theorem 2.

THEOREM 2. *For positive integers m, n,*

$$(7)\quad y_{1,1} \le \min\left(S_{1,1} - \frac{2}{m}S_{2,1} - \frac{2}{mn}S_{1,2}\,,\ S_{1,1} - \frac{2}{n}S_{1,2} - \frac{2}{mn}S_{2,1}\right)$$

This inequality is a new one and better than the inequality in Lee [1992] in terms of the binomial moments $S_{1,1}, S_{1,2}$ and $S_{2,1}$; that is, for integers m and n, $2 \le m,\ 2 \le n$,

$$y_{1,1} \le \min\left(S_{1,1} - \frac{2}{m}S_{2,1}, S_{1,1} - \frac{2}{n}S_{1,2}\right).$$

3. Proofs

Proof of Theorem 1. We use the method of indicators. Let

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

and by using binomial moment of (3) and indicators, the right hand side of (6) becomes

(8)

$$E\Bigg[\binom{X}{1}\binom{Y}{1} - \max\Bigg\{\sum_{i=1}^{m-1}\sum_{j=1}^{n} I(A_i)I(A_{i+1})I(B_j) + \sum_{j=1}^{n-1} I(A_k)I(B_j)I(B_{j+1}),$$
$$\sum_{i=1}^{m}\sum_{j=1}^{m-1} I(A_i)I(B_j)I(B_{j+1}) + \sum_{i=1}^{m-1} I(A_i)I(A_{i+1})I(B_k)\Bigg\}\Bigg].$$

Then $E[I(X \ge 1, Y \ge 1)] = P(X \ge 1, Y \ge 1)$, and thus in order to prove (8), it suffices to show that

(9)

$$I(X \ge 1)I(Y \ge 1)$$
$$\le XY - \max\Bigg\{\sum_{i=1}^{m-1}\sum_{j=1}^{n} I(A_i)I(A_{i+1})I(B_j) + \sum_{j=1}^{n-1} I(A_k)I(B_j)I(B_{j+1}),$$
$$\sum_{i=1}^{m}\sum_{j=1}^{n-1} I(A_i)I(B_j)I(B_{j+1}) + \sum_{i=1}^{m-1} I(A_i)I(A_{i+1})I(B_k)\Bigg\}$$

Note that both sides of (9) are zero if either X or Y equals zero, hence, in proving (9) we may assume that $X \geq 1$ and $Y \geq 1$, in which the left hand side of (9) is identically one. Thus, we have to prove that

(10)
$U(X, Y) =$ the right hand side of (9) ≥ 1 for $1 \leq X \leq m$, $1 \leq Y \leq n$.

We distinguish three cases.
(i) First case. There are integers X and Y with X=1, Y=1 ; that is, there are only two events A_i and B_j occur because the events A_i 's and B_j 's occur numercal order. Then this case is evident, having one on both sides of (10).
(ii) Second case. For integers p,q with $2 \leq p \leq m$, $2 \leq q \leq n$, $X = 1$ and $Y = q$ or $X = p$ and $Y = 1$; that is, there are the events that exactly one $A_i(B_j)$ and at least two more $B_j' s(A_i' s)$ occur. Then

$$u(1, q) = 1 \cdot q - \max\{0, (q-1) + 0\} = 1 \text{ and}$$
$$u(p, 1) = p \cdot 1 - \max\{(p-1) + 0, 0\} = 1.$$

Hence, we get (10).
(iii) Third case. For integers p,q with $2 \leq p \leq m, 2 \leq q \leq n, X = p$ and $Y = q$; that is, there are the events that at least two more $A_i' s \text{and} B_j' s$ occur. Then

$$u(p, q) = p \cdot q - \max\{(p-1)q + (q-1), p(q-1) + (p-1)\} = 1$$

Hence, We get (10). This completes the proof.

Proof of Theorem 2. We turn to indicators. Hence, to prove (7) it suffices to show that

$$1 \leq XY - \frac{X(X-1)Y}{m} - \frac{XY(Y-1)}{mn} \tag{11}$$
$$\text{if } 1 \leq X \leq m, \ 1 \leq Y \leq n \text{ and}$$

$$1 \leq XY - \frac{XY(Y-1)}{n} - \frac{X(X-1)Y}{mn} \tag{12}$$
$$\text{if } 1 \leq X \leq m, \ 1 \leq Y \leq n$$

Let $f(X,Y)$ = the right hand side of (11) $= \dfrac{XY(mn-n(X-1)-(Y-1))}{mn}$. Then the function $f(X,Y)$ is parabola whose minimum point, for integers X,Y with $1 \le X \le m,\ 1 \le Y \le n$, are at

$$C = \left\{(X,Y) = (1,1),(1,n),(m,1),(m,n),([\frac{m+1}{2}],1)([\frac{m+1}{2}]+1,1)\right\}$$

where the ordered pair $(\frac{m+1}{2},1)$ come from $\frac{\partial f(X,Y)}{\partial X} = 0$ and $[\frac{m+1}{2}]$ means the integer part of $\frac{m+1}{2}$. We know that $f((X,Y) \in C) \ge 1$ if $1 \le X \le m,\ 1 \le Y \le n$ which completes the proof of (11).

Let $g(X,Y)$ = the right hand side of (12) $= \dfrac{XY(mn-m(Y-1)-(X-1))}{mn}$. By the same way, $g((X,Y) \in D) \ge 1$ if $1 \le X \le m,\ 1 \le Y \le n$ where the set

$$D = \left\{(X,Y) = (1,1),(1,n),(m,1),(m,n),(1,[\frac{n+1}{2}]),(1,[\frac{n+1}{2}]+1)\right\}$$

which the ordered pair $(1,\frac{n+1}{2})$ come from $\frac{\partial g(X,Y)}{\partial Y} = 0$ and $[\frac{n+1}{2}]$ means the integer part of $\frac{n+1}{2}$. This completes the proof of (12).

References

1. Éva Galambos, *Discussion of probabilistic inequalities by the method of A. Renyi(Hungarian)*, Dissertation, L. Eotvos University, Budapest., 1965.
2. J. Galambos and M.-Y. Lee, *Extensions of some univariate Bonferroni-type inequalities to multivariate setting*, In Probability Theory and Applications, ed. J. Galambos and I. Katai, Kluwer, Dordrecht (1992), 143-154.
3. J. Galambos and M.-Y. Lee, *Further studies of Bonferroni-type inequalities*, Journal of applied probability **31A** (1994), 63-69.
4. J. Galambos and Y. XU, *Some optimal bivariate Bonferroni-type bounds*, Proc. Amer. Math. Soc. **117** (1993), 523-528.
5. M.-Y. Lee, *Bivariate Bonferroni inequalities*, Aequationes Maht. **44** (1992), 220-225.
6. S. M. Kwerel, *Most stringent bounds on aggregated probabilities of partially specified dependent systems*, J. Amer. Stat. Assoc. **70** (1975), 472-479.

7. R. M. Meyer, *Note on a multivariate form of Bonferroni's inequalities*, Ann. Math. Statist. **40** (1969), 692-693.

Department of Mathematics
Dankook University
Cheonan, Chungnam, Korea