Bull. Korean Math. Soc. 32 (1995), No. 2, pp. 329–336

AN IMPROVED BONFERRONI–TYPE INEQUALITY

MIN-YOUNG LEE

1. Introduction

Let $A_1, A_2, \cdots, A_n$ be a sequence of events on a given probability space and let m_n be the number of those $A's$ which occur. Put $S_{0,n} = 1$ and

$$S_{k,n} = \sum P(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}), \quad (1 \leq k)$$

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

Kounias (1968) has proved that

$$P(\cup_{i=1}^n A_i) \leq \sum_{i=1}^n P(A_i) - \max_j \sum_{i \neq j} P(A_i \cap A_j) \tag{1}$$

which improves on the simple Bonferroni upper bound of $\sum P(A_i)$. Margolin and Maurer (1976) generalizes this result by using more than just $\sum P(A_i)$ from the classical estimates. Hunter (1976), whose result is reobtained in the paper of Worsley (1982), presents an improved upper bound which is constructed by edges on a graph.

When one compares (1) with the lower bound

$$S_{1,n} - S_{2,n} \leq P(\cup_{i=1}^n A_i)$$

we see that the real meaning of (1) is that if not too many terms of intersections of pairs are subtracted from $S_{1,n}$, it still remains an upper bound, but when all intersections of pairs are subtracted, we get a lower

Received April 27, 1994.
1991 AMS Subject Classification: Primary 60E15; Secondary 60C05.
Key words: Bonferroni-type inequality, the method of indicators.
This paper was supported by NON DIRECTED RESEARCH FUND, KOREA.

bound. The question thus arises that if we subtract more that Kounias does but much less than $S_{2,n}$, how many intersections of three events will compensate for this in order to get an upper bound again. The classical upper bound of degree three is

$$P(\cup_{i=1}^{n} A_i) \leq S_{1,n} - S_{2,n} + S_{3,n}$$

and my idea is to reduce the number of terms both in $S_{2,n}$ and $S_{3,n}$ without violating its being an upper bound. For a related idea, see the graph-dependent models of Renyi (1961) and Galambos (1966). It is well demonstrated in the literature that the classical Bonferroni bounds are sometimes of little value exactly becauce of the large number of terms in $S_{k,n}$, $2 \leq k$. For such cases, bounds with limited number of terms can be of value. In this direction, I prove the inequality of the theorem that follows.

2. An improved bound

The upper bound is improved by the following result.

THEOREM 1. *For integers $3 \leq n$ and $1 \leq i \leq n-2$*

$$(2)\quad P(m_n \geq 1) \leq S_{1,n} - \sum_{i<j\leq i+2} P(A_i \cap A_j) + \sum_{i=1}^{n-2} P(A_i \cap A_{i+1} \cap A_{i+2}).$$

Proof. We use the method of indicators. That is, let $I(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k})$ be 1 if $A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}$ occurs or 0, otherwise. Then

$$I(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}) = I(A_{i_1})I(A_{i_2}) \cdots I(A_{i_k})$$

and

$$E[I(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k})] = P(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k})$$

Futhermore, the indicator variable $I(m_n \geq 1)$ is 1 if $m_n \geq 1$ and 0 if $m_n = 0$. Note also that $\sum_{i=1}^{n} I(A_i) = m_n$ and $S_{1,n} = E[m_n]$.

We thus have to prove

$$(3)\qquad m_n - \sum_{i<j\le i+2} I(A_i)I(A_j) + \sum_{i=1}^{n-2} I(A_i)I(A_i+1)I(A_i+2) \ge 1$$

if $m_n \ge 1$ and the left hand side of (3) is greater than zero or equal to zero if $m_n = 0$. The latter case is evident, having zero on both sides. Also, if $m_n = 1$ both sides of (3) equal 1. Hence, for the sequel we may assume $m_n \ge 2$.

Next, we place the events $A_1, A_2, \cdots, A_n$ at every sample point into blocks which consist of events of the kind $A_j \cap A_{j+1} \cap \cdots \cap A_{j+k_j}$, which is a full block if neither A_{j-1} nor A_{j+k_j+1} occurs. Assume that in this way, at a given sample point, we have t blocks. We distinguish three cases.

(i) First case. For all $j, k_j \ge 2$; that is, every full block has at least two events. One can express both $\sum_{i<j\le i+2} I(A_i)I(A_j)$ and $\sum_{i=1}^{n-2} I(A_i)I(A_{i+1})I(A_{i+2})$ by means of blocks; that is, if the t blocks have length $k_j, i \le j \le t$, then the above sums equal

$$(4)\quad \sum_{j=1}^{t}[2(k_j-2)+1] + \begin{pmatrix} 0 \\ 1 \\ 2 \\ \vdots \\ t-1 \end{pmatrix} \qquad \text{and} \qquad \sum_{j=1}^{t}(k_j-2), \qquad \text{respectively}$$

where $\begin{pmatrix} 0 \\ 1 \\ 2 \\ \vdots \\ t-1 \end{pmatrix}$ denotes the number $\sum_{j=1}^{t} L_d^j$, L_d^j being 1 if $d = 2$ and 0 if $d > 2$ and d is the difference between last number of j-th block and first number of next one.

Since $\sum_{j=1}^{t} k_j = m_n$, by (4), the left hand side of (3) becomes

$$t - \begin{pmatrix} 0 \\ 1 \\ 2 \\ \vdots \\ t-1 \end{pmatrix} \geq 1$$

Hence, we get (3).

(ii) Second case. There is only one j with $k_j = 1$; that is, only in the r-th block, say, only one event occurs, and for $j \neq r, \quad k_j \geq 2$. We now have

$$\begin{aligned} &\sum_{i<j\leq i+2} I(A_i)I(A_j) \\ (5) \qquad &= \sum_{j=1}^{r-1}[2(k_j-2)+1] + \begin{pmatrix} 0 \\ 1 \\ 2 \\ \vdots \\ r-2 \end{pmatrix} + \binom{0}{1} + \binom{0}{1} \\ &\quad + \sum_{j=r+1}^{t} [2(k_j-2)+1] + \begin{pmatrix} 0 \\ 1 \\ 2 \\ \vdots \\ t-r-2 \end{pmatrix} \\ &= \sum_{j=1}^{r-1}[2(k_j-2)+1] + \sum_{j=r+1}^{t} [2(k_j-2)+1] + \begin{pmatrix} 0 \\ 1 \\ 2 \\ \vdots \\ t-1 \end{pmatrix} \end{aligned}$$

and

$$(6) \qquad \sum_{i=1}^{n-2} I(A_i)I(A_{i+1})I(A_{i+2}) = \sum_{j=1}^{r-1}(k_j-2) + \sum_{j=r+1}^{t}(k_j-2)$$

Since $\sum_{j=1}^{r-1} k_j + 1 + \sum_{j=r+1}^{t} k_j = m_n$, in view of (5) and (6), the left hand side of (3) is

$$t - \begin{pmatrix} 0 \\ 1 \\ 2 \\ \vdots \\ t-1 \end{pmatrix} \geq 1.$$

Once again, (3) obtains.

(iii) Third case. There exist more than one j with $k_j = 1$; that is, there are several blocks which have only one event. In the same manner as in (ii), except that several terms contribute $\binom{0}{1}$, we get (3). This completes the proof.

Taking average which over $i = 1, 2, \cdots, n$ of (2), We get the following Bonferroni-type inequality.

THEOREM 2. *Let n be integers with $n \geq 3$. Then*

$$P(m_n \geq 1) \leq S_{1,n} - \frac{(2n-3)}{\binom{n}{2}} S_{2,n} + \frac{(n-2)}{\binom{n}{3}} S_{3,n}$$

This inequality is known that it is the best possible upper bound in terms of $S_{1,n}$, $S_{2,n}$ and $S_{3,n}$ (see kwerel (1975))

Also, its simple proof appers in Galambos and Xu (1990) in consequence of the iteration method.

3. Numerical Examples

EXAMPLE 1. Let X_j be the time to failure of the j-th component of a piece of equipment. Assume that each X_j is a unit exponential variate; that is, for each j,

$$P(X_j < x) = 1 - e^{-x}, \qquad (x > 0).$$

Consider a group of five components, X_1, X_2, X_3, X_4, X_5. We assume we just know the following information.

(a) X_i and X_{i+1} are dependent; that is, X_1 and X_2 are dependent, so are X_2 and X_3, X_3 and X_4 and, finally, X_4 and X_5.

(b) X_{i+1} is dependent on both X_i and X_{i+2}; that is, X_2 is dependent on both X_1 and X_3; X_3 is dependent on both X_2 and X_4; X_4 is dependent on both X_3 and X_5.

No other information is available on the interdependence of the components. We also specify the bivariate and the trivariate distributions of the X_j. For simplicity, let the bivariate and the trivariate distributions for all dependent components specified in (a) and (b) be the same. Let

$$\begin{aligned} P(X_1 < x, X_2 < y) &= P(X_2 < x, X_3 < y) = P(X_3 < x, X_4 < y) \\ &= P(X_4 < x, X_5 < y) = P(X_1 < x, X_3 < y) \\ &= P(X_2 < x, X_4 < y) = P(X_3 < x, X_5 < y) \\ &= (1 - e^{-x})(1 - e^{-y})(1 - \frac{1}{2}e^{-x-y}) \end{aligned}$$

$$\begin{aligned} P(X_1 < x, X_2 < y, X_3 < z) &= P(X_2 < x, X_3 < y, X_4 < z) \\ &= P(X_3 < x, X_4 < y, X_5 < z) \\ &= (1 - e^{-x})(1 - e^{-y})(1 - e^{-z})(1 - \frac{1}{3}e^{-x-y-z}). \end{aligned}$$

No further assumption is made. We would like to estimate $P(W_5 \geq x)$ where $W_5 = \min(X_1, X_2, X_3, X_4, X_5)$. We choose the events $A_j = (X_j < x)$ and then $(m_5 = 0) = (W_5 \geq x)$. For a numerical calculation, let us choose $x = 0.1$. We then estimate $P(W_5 \geq 0.1)$. We have

$$S_{1,5} = 5(1 - e^{-0.1}) = 0.47580$$

$$\sum_{i<j<=i+2} P(A_i \cap A_j) = 7[(1 - e^{-0.1})^2(1 - 1over2e^{-0.2})] = 0.03744$$

$$\begin{aligned} \sum P(A_i \cap A_{i+1} \cap A_{i+2}) &= 3[(1 - e^{-0.1})^3(1 - 1over3e^{-0.3})] \\ &= 0.00195. \end{aligned}$$

Theorem 1 now gives $P(\cup_{i=1}^5 A_i) = P(m_n \geq 1) \leq 0.44031$. As was pointed out, with the events $(X_j < 0.1)$

$$(7) \qquad P(W_5 \geq 0.1) = P(m_n = 0) = 1 - P(m_n \geq 1) \geq 0.55969$$

When we use the method of maximum spanning tree by Worsley by choosing $T = (c_{12}, c_{23}, \cdots, c_{n-1,n})$, we have

$$(8) \qquad P(\cup_{i=1}^{n} A_i) \leq \sum_{i=1}^{n} P(A_i) - \sum_{i=1}^{n-1} P(A_i \cap A_{i+1}).$$

This yields

$$P(m_n \geq 1) \leq 0.45441$$

which can be written as

$$(9) \qquad P(W_5 \geq 0.1) = P(m_n = 0) \geq 0.54559$$

It is, of course, not surprising that (7) was a better estimate than (9). While (9) is the best possible that can be obtained in terms of $\sum P(A_i)$ and $\sum P(A_i \cap A_{i+1})$, in (7) we made use of further information on the X_j. In order to give upper bounds for $P(W_5 \geq 0.1)$, we assume, in addition to those which led to (7), that every pair not in (a) is independent. Then we get lower bound of $P(\cup A_i)$ by the method of Margolin and Maurer; that is,

$$(10) \qquad P(\cup_{i=1}^{n} A_i) \geq S_{1,n} - S_{2,n} + \max_{(i \neq r \neq j, i<j)} P(A_i \cap A_j \cap A_r)$$

By (10), we get

$$P(\cup_{i=1}^{5} A_i) \geq 0.41314$$

which can be written as

$$P(W_5 \geq 0.1) \leq 0.58686.$$

EXAMPLE 2. Consider numerical example 3.1 [peak periods of a disease] in the paper of Worsley. We calculate the bivariate normal distribution $P(A_1 \cap A_3)$, $P(A_2 \cap A_4)$, $P(A_3 \cap A_5)$ and $P(A_4 \cap A_6)$ with covariance $\rho_{i,i+2} = \frac{1}{3}$ and the trivariate normal distribution $P(A_1 \cap A_2 \cap A_3)$, $P(A_2 \cap A_3 \cap A_4)$, $P(A_3 \cap A_4 \cap A_5)$, and $P(A_4 \cap A_5 \cap A_6)$ with $\rho_{i,i+1} = \frac{2}{3}$ and $\rho_{i,i+2} = \frac{1}{3}$. Then our new estimate (2) is better than (8) because the right hand side of (8) is greater than or equal to the right hand side of (2). Also, we can get better lower bound by (10) which clearly improves on the Bonferroni lower bound $S_{1,n} - S_{2,n}$.

References

1. Galambos, J., *On the sieve methods in probability theory I*, Studia Sci. Math, Hungar. **1** (1966), 39-50.
2. Galambos, J. and Xu, Y., *A new method for generating Bonferroni-type inequalities by interation*, Math. Proc. Cambridge Philas. Soc. **107** (1990), 601-607.
3. Hunter, D., *An upper bound for the probability of a union* J. Appl. Prob. **13** (1976), 597-603.
4. Kounias, E.G., *Bounds for the probability of a union of events, with applications*, Ann. Math. Statist **39** (1968), 2154-8.
5. Kwerel, S.M., *Bounds on the probability of the union and intersection of m events*, Adv. in Appl. Probab **7** (1975), 431-438.
6. Margolin, B. J. and Maurer, W., *Kolmogorov-Smirnov type for exponential data with unknown scale, and related problems*, Biometrika **63** (1976), 149-60.
7. Renyi, A., *A general method for proving theorems in probability theory and some of its applications*, Original in Hungarian. Translated into english in: Selected Papers of A. Renyi, **2** (1961); Akademiai Kiado, Budapest (1976), 581-602.
8. Sobel, M. and Uppuluri, V.R.R., *On Bonferroni-type inequalities of the same degree for the probability of unions and intersections*, Ann. Math. Statist. **43** (1972), 1549-58.
9. Worsley, K.J., *An improved Bonferroni inequality*, Biometrika **69** (1982), 297-302.

Department of Mathematics, Dankook University, Cheonan-si 330-714, Korea