The Annals of Probability
1977, Vol. 5, No. 4, 577–581

BONFERRONI INEQUALITIES

BY JANOS GALAMBOS

Temple University

Let $A_1, A_2, \cdots, A_n$ be events on a probability space. Let $S_{k,n}$ be the kth binomial moment of the number m_n of those A's which occur. An estimate on the distribution $y_t = P(m_n \geqq t)$ by a linear combination of $S_{1,n}$, $S_{2,n}, \cdots, S_{n,n}$ is called a Bonferroni inequality. We present for proving Bonferroni inequalities a method which makes use of the following two facts: the sequence y_t is decreasing and $S_{k,n}$ is a linear combination of the y_t. By this method, we significantly simplify a recent proof for the sharpest possible lower bound on y_1 in terms of $S_{1,n}$ and $S_{2,n}$. In addition, we obtain an extension of known bounds on y_t in the spirit of a recent extension of the method of inclusion and exclusion.

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

$$S_{k,n} = \sum_{1 \leqq i_1 < \cdots < i_k \leqq n} P(A_{i_1} \cdots A_{i_k}) . \tag{1}$$

With real numbers c_j and d_j, the inequalities

$$\sum_{j=0}^{n} c_j S_{j,n} \leqq P(m_n \geqq t) \leqq \sum_{j=0}^{n} d_j S_{j,n} \tag{2}$$

are called Bonferroni inequalities. It is not a restriction that, in both sums above, j runs from 0 to n, since c_j and d_j may take the value zero. Here c_j and d_j are preassigned; that is, they are not functions of $S_{k,n}$, $k \geqq 1$. They may, however, depend on t and n.

In a recent paper, Galambos (1975) pointed out that Bonferroni inequalities can be proved by assuming that the A's are either exchangeable or even that they are independent. Here we present a further simple method for proving (2). With this method, we shall reobtain the best known forms of (2) as well as extend some of them. The proofs for the known cases are considerably shorter than the original ones were. In addition, our method of proof will show that (2) is a special case of simple nonprobabilistic inequalities. This explains why Bonferroni inequalities are either valid for arbitrary events or fail for very simple structures (such as independent events or exchangeable ones).

The method of proof is based on the following observation. It is well known that $S_{k,n}$ of (1) is the kth binomial moment of m_n. Hence, putting

$$y_t = P(m_n \geqq t) , \tag{3}$$

we get

$$S_{k,n} = \sum_{t=k}^{n} \binom{t-1}{k-1} y_t , \qquad k \geqq 1 . \tag{4}$$

Received August 23, 1976.

AMS 1970 *subject classifications.* Primary 60C05; Secondary 60E05.

Key words and phrases. Events, number of occurrences, binomial moments, Bonferroni inequalities, best linear bounds, method of inclusion and exclusion, distribution of order statistics, dependent samples.

Here $y_1 \geqq y_2 \geqq \cdots \geqq y_n \geqq 0$. It turns out that (2) is a set of inequalities for decreasing sequences y_j, $1 \leqq j \leqq n$, where $S_{k,n}$ is defined by (4), but the probabilistic meaning (3) of y_t is not significant.

2. The inequalities. We prove the following two theorems. The first one is known, its original proof however is significantly simplified here. The inequalities of the second theorem are new.

THEOREM 1. *Assume that the inequality*

$$P(m_n \geqq 1) \geqq aS_{1,n} + bS_{2,n} \tag{5}$$

holds for an arbitrary sequence $A_1, A_2, \cdots, A_n$ *of events. Then there is a uniformly better bound than* (5) *unless the coefficients a and b satisfy*

$$a = -kb = 2/(k+1)\,, \qquad k > 0 \ \textit{integer}. \tag{6}$$

With the coefficients in (6), (5) *always holds. Hence, maximizing in k leads to*

$$P(m_n \geqq 1) \geqq 2S_{1,n}/(k^* + 1) - 2S_{2,n}/k^*(k^* + 1)\,, \tag{7}$$

where $k^* - 1$ *is the integer part of* $2S_{2,n}/S_{1,n}$.

THEOREM 2. *Let* $1 \leqq t \leqq n$, $0 \leqq j \leqq \frac{1}{2}(n-t) - 1$ *and* $0 \leqq r \leqq \frac{1}{2}(n-t-1)$ *be integers. Then*

$$\sum_{k=0}^{2j+1} (-1)^k \binom{k+t-1}{t-1} S_{k+t,n} + \frac{2j+2}{n-t} \binom{2j+t+1}{t-1} S_{2j+t+2,n}$$
$$\leqq P(m_n \geqq t) \leqq \sum_{k=0}^{2r} (-1)^k \binom{k+t-1}{t-1} S_{k+t,n} - \frac{2r+1}{n-t} \binom{2r+t}{t-1} S_{2r+t+1,n}\,.$$

Theorems of the above nature are very valuable for estimating the distribution of order statistics for dependent systems, when the joint distribution of the observations are not known in all dimensions. Indeed, if $X_1, X_2, \cdots, X_n$ are random variables and $A_j = A_j(x) = \{X_j \geqq x\}$, then

$$P(m_n \geqq t) = 1 - P(X_{n-t+1:n} < x)\,,$$

where $X_{r:n}$ is the rth order statistic of the X_j's. On the other hand, $S_{k,n}$ is expressible in terms of the k-dimensional distributions of the X_s, $1 \leqq s \leqq n$.

Theorem 1 is due to Kwerel (1975), but his proof is substantially reduced here. The actual estimate (7), without establishing its extremal property, was found earlier by Dawson and Sankoff (1967). Theorem 2 extends the so-called Jordan inequalities in the same manner in which Sobel and Uppuluri (1972) extended the method of inclusion and exclusion for exchangeable events, which was later shown by Galambos (1975) to hold for arbitrary events.

3. Proofs. The proof will not make use of the probabilistic nature of the inequalities of Theorems 1 and 2, as was pointed out in the introduction. Only the facts, that the sequence $y_t \geqq 0$ of (3) is decreasing and that (4) holds, will be used. Hence, the proof itself is of interest.

PROOF OF THEOREM 1. Substituting (3) and (4) in (5), we get

$$y_1 \geqq \sum_{j=1}^{n} \{a + b(j-1)\} y_j = a y_1 + \sum_{j=2}^{n} \{a + b(j-1)\} y_j$$

or

$$(1-a) y_1 \geqq \sum_{j=2}^{n} \{a + b(j-1)\} y_j \,. \tag{8}$$

Since (5) is assumed to hold for arbitrary events, the y_j are variables. We thus get the following estimates on a and b by different choices of the y's. The choice $y_1 = 1$ and $y_j = 0$ for $j \geqq 2$ yields the restriction $1 - a \geqq 0$ for a. On the other hand, with $y_j = 1$ for all j, (8) becomes

$$1 - an \geqq \tfrac{1}{2} bn(n-1)\,, \qquad n = 1, 2, \cdots .$$

This is possible only if $b < 0$. Thus $a \leqq 0$ would evidently imply that the bound in (5) can be improved. Consequently, $a > 0$, and thus, together with the preceding estimates, $0 < a \leqq 1$ and $b < 0$. Hence, the right-hand side of (8) can easily be split into positive and negative terms. Let the integer k be defined as follows. Let $k = n$, if all the coefficients $a + b(j-1)$ are positive, and if all of these coefficients are negative then $k = 2$. Otherwise k is defined as the unique integer for which

$$a + b(k-1) \geqq 0 \qquad \text{and} \qquad a + bk < 0 \,. \tag{9}$$

Another special choice, $y_1 = y_2 = \cdots = y_k = 1$ and $y_{k+1} = 0$, in (8) results in the inequality

$$1 - a \geqq a(k-1) + \tfrac{1}{2} bk(k-1)\,, \tag{10}$$

that is,

$$b \leqq 2(1-ak)/k(k-1)\,. \tag{10a}$$

Now, since the y's are monotonically decreasing, the validity of (10) implies (8) for arbitrary y's because of the special meaning of k. Hence, the best value of b is provided by equality in (10a), that is

$$b = 2(1-ak)/k(k-1)\,, \tag{11}$$

where k is defined in (9). Combining (9) and (11), we get that in the optimal case, k is the integer defined by

$$2/(k+1) < a \leqq 2/k \,. \tag{12}$$

But, since with the value of b in (11),

$$aS_{1,n} + bS_{2,n} = a(S_{1,n} - 2S_{2,n}/(k-1)) + 2\{k(k-1)\}^{-1} S_{2,n}\,,$$

the best choice of a is one of the end points in (12), depending whether the coefficient of a on the right-hand side of the equation above is negative or positive for a given value of k. But, whichever end point is to be taken in (12) for a, together with (11), it leads to the form (6) for a and b. Since, at each step of the preceding arguments, we took care that (5) should apply, we actually proved so far that, for any integer k satisfying $1 \leqq k \leqq n$, the inequality

$$y_1 \geqq \frac{2}{k+1} S_{1,n} - \frac{2}{k(k+1)} S_{2,n} = f(k) , \quad \text{say,}$$

is valid for arbitrary decreasing sequence $\{y_j\}$ with $y_n \geqq 0$. The proof will therefore be completed if we determine the maximum of $f(k)$ for given $S_{1,n}$ and $S_{2,n}$, where $1 \leqq k \leqq n$. If we first disregard the restriction on k, then the maximum of $f(k)$ can easily be found by observing that $f(k)$ either decreases for all values of $k \geqq 1$, or it increases for $1 \leqq k \leqq k_0$ with some $k_0 > 1$ and decreases for all $k > k_0$. Hence, in either case, $f(k)$ takes its maximum for the least $k \geqq 1$ such that $f(k+1) - f(k)$ is negative. It is immediate that this condition occurs at $k = k^*$, where $k^* - 1$ is the integer part of $2S_{2,n}/S_{1,n}$. If we show that this k^* satisfies $1 \leqq k^* \leqq n$, the theorem will then be established. However, for $y_1 \geqq y_2 \geqq \cdots \geqq y_n \geqq 0$,

$$\begin{aligned}(n-1)S_{1,n} - 2S_{2,n} &= \textstyle\sum_{j=1}^{n} (n+1-2j)y_j \\ &= \textstyle\sum' (n+1-2j)(y_j - y_{n-j+1}) \geqq 0 ,\end{aligned}$$

where $\sum'$ signifies summation over j, from 1 to $\frac{1}{2}n$ or $\frac{1}{2}(n+1)$ according as n is even or odd. The inequality above is equivalent to $0 \leqq k^* - 1 \leqq n - 1$, and the proof is thus completed.

Proof of Theorem 2. We again start from the relation (4). Let a be any integer satisfying $a \geqq 0$ and $a + t \leqq n$. Then

$$\begin{aligned}\textstyle\sum_{k=0}^{a} (-1)^k \binom{k+t-1}{t-1} S_{k+t,n} &= \textstyle\sum_{k=0}^{a} (-1)^k \binom{k+t-1}{t-1} \sum_{s=k+t}^{n} \binom{s-1}{k+t-1} y_s \\ &= \textstyle\sum_{s=t}^{n} y_s \sum_{k=0}^{T} (-1)^k \binom{k+t-1}{t-1}\binom{s-1}{k+t-1} ,\end{aligned}$$

where $T = \min(a, s-t)$. Thus, since

$$\binom{k+t-1}{t-1}\binom{s-1}{k+t-1} = \binom{s-1}{t-1}\binom{s-t}{k} ,$$

$$\textstyle\sum_{k=0}^{a} (-1)^k \binom{k+t-1}{t-1} S_{k+t,n} = y_t + \sum_{s=t+1}^{n} \binom{s-1}{t-1} y_s \sum_{k=0}^{T} (-1)^k \binom{s-t}{k} . \tag{13}$$

We now apply the identity

$$\textstyle\sum_{k=0}^{T} (-1)^k \binom{s-t}{k} = (-1)^T \binom{s-t-1}{T} ,$$

which is well known and easy to prove by induction over T. Hence, in view of the special value of T and by (13),

$$\textstyle\sum_{k=0}^{a} (-1)^k \binom{k+t-1}{t-1} S_{k+t,n} = y_t + (-1)^a \sum_{s=t+a+1}^{n} \binom{s-1}{t-1}\binom{s-t-1}{a} y_s . \tag{14}$$

Hence, the left-hand side is smaller or larger than y_t according as a is odd or even, respectively. (14) thus implies the Jordan formulas. For the improvement of our theorem over these formulas, we have to estimate the term occurring on the right-hand side of (14) after y_t and to compare it with $S_{a+t+1,n}$. More precisely, the inequalities of Theorem 2 follow from (14), if we show that

$$\textstyle\sum_{s=t+a+1}^{n} \binom{s-1}{t-1}\binom{s-t-1}{a} y_s \geqq \frac{a+1}{n-t} \binom{a+t}{t-1} S_{a+t+1,n} . \tag{15}$$

However, (15) is immediate by an appeal to (4). As a matter of fact, after applying (4), if we write all binomial coefficients in terms of factorials, (15) reduces to the inequality $s \leqq n$, which is evident. This completes the proof of Theorem 2.

REFERENCES

[1] DAWSON, D. A. and SANKOFF, D. (1967). An inequality for probabilities. *Proc. Amer. Math. Soc.* **18** 504–507.

[2] GALAMBOS, J. (1975). Methods for proving Bonferroni-type inequalities. *J. London Math. Soc.* (2) **9** 561–564.

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

[4] SOBEL, M. and UPPULURI, V. R. R. (1972). On Bonferroni-type inequalities of the same degree for the probability of unions and intersections. *Ann. Math. Statist.* **43** 1549–1558.

DEPARTMENT OF MATHEMATICS
TEMPLE UNIVERSITY
PHILADELPHIA, PENNSYLVANIA 19122