ON BONFERRONI-TYPE INEQUALITIES OF THE SAME DEGREE FOR THE PROBABILITY OF A UNION

by

Milton Sobel[§] and V. R. R. Uppuluri[§§]

University of Minnesota Oak Ridge National Lab.

Technical Report No. 156

April 1971

University of Minnesota

Minneapolis, Minnesota

[§]Research supported by NSF Grant GP-11021.

[§§]Visiting Professor, University of Minnesota, partially supported by the U.S. Atomic Energy commission under contract with the Union Carbide Corporation.

1. Introduction and Summary.

Let A_i $(i = 1,2,\ldots, k)$ denote a finite set of events associated with the probability space $(\Omega, \mathfrak{F}, P)$ and let $\chi_i(\omega)$ denote the indicator random variable of A_i. Then $\max_{i=1,2,\ldots,k} \chi_i(\omega)$ is the indicator random variable of the set $\bigcup_{i=1}^{k} A_i$. In [4], Kounias found bounds for $P\{\bigcup_{i=1}^{k} A_i\}$ in terms of $P\{A_i\}$ and $P\{A_iA_j\}$. Kounias strengthens the Bonferroni inequalities (cf. Feller [2]) by giving lower and upper bounds both of the second degree $(\nu = 2)$. In this paper we extend his result by considering both lower and upper bounds on $P\{\bigcup_{i=1}^{k} A_i\}$ for any fixed degree $\nu(\nu = 1,2,3,\ldots,k)$ and indicate some applications to Dirichlet integrals.

In the case of exchangeable random variables $\chi_i(\omega)$ we let $P_\alpha = P\{A_{j_1} A_{j_2} \ldots A_{j_\alpha}\}$. We obtain for any odd degree $\nu \leq k$

$$(1.1) \qquad \sum_{\alpha=1}^{\nu-1} (-1)^{\alpha-1}\binom{k}{\alpha}P_\alpha + \binom{k-1}{\nu-1}P_\nu \leq P\{\bigcup_{i=1}^{k} A_i\} \leq \sum_{\alpha=1}^{\nu} (-1)^{\alpha-1}\binom{k}{\alpha}P_\alpha$$

and for any even degree $\nu \leq k$

$$(1.2) \qquad \sum_{\alpha=1}^{\nu} (-1)^{\alpha-1}\binom{k}{\alpha}P_\alpha \leq P\{\bigcup_{i=1}^{k} A_i\} \leq \sum_{\alpha=1}^{\nu-1} (-1)^{\alpha-1}\binom{k}{\alpha}P_\alpha - \binom{k-1}{\nu-1}P_\nu .$$

One can easily construct examples where these bounds are attained. In fact, for $\nu = k$ the left and right sides are equal in both (1.1) and (1.2) so that equality is then attained. It should be noted that the left sides of (1.1) and (1.2) can be negative (in which case we use zero for the lower bound) and that the right sides can exceed 1 (in which case we use 1 for the upper bound).

Some applications of these bounds to Dirichlet integrals are considered in sections 4 and 5. These bounds are applicable for any problem involving a union of events and, in particular to problems involving Dirichlet integrals. For some further references where these integrals arise the reader may refer to the references in [7].

2. Derivation of the Bounds.

In order to express the results in a compact notation we will define a Bonferroni indicator random variable and an operation denoted by $*$.

Definition:

Let $B_{r,j}$ with $j \leq k$ and $1 \leq r \leq k$ denote the Bonferroni function of degree r on the j sets $A_{i_1}, A_{i_2}, \ldots, A_{i_j}$ defined by

$$(2.1) \qquad B_{r,j} = \sum_{\alpha_1=1}^{j} \chi_{\alpha_1} - \sum_{\alpha_1<\alpha_2} \chi_{\alpha_1}\chi_{\alpha_2} + \ldots + (-1)^{r-1} \sum_{\alpha_1<\alpha_2<\ldots<\alpha_r} \chi_{\alpha_1}\chi_{\alpha_2}\cdots\chi_{\alpha_r}$$

where $\chi_\alpha = \chi_\alpha(\omega)$ and $1 \leq \alpha_i \leq j$ $(i = 1,2,\ldots, r)$. We define $B_{0,j}$ to be identically zero.

Definition:

Let $B_{r,i} * B_{s,j}$ denote the function defined by

$$(2.2) \qquad B_{r,i} * B_{s,j} = B_{r,i} + B_{s,j} - B_{r,i}B_{s,j}.$$

It is implicitly assumed in this definition that the i sets in $B_{r,i}$ do not include any of the j sets in $B_{s,j}$ and vice versa i.e., the index sets are disjoint, although the sets need not be. It is easily verified that $B_{2,2} = B_{1,1} * B_{1,1}$ and we sometimes denote the latter by $B_{1,1}^{*2}$.

It is easily verified that the $*$ product is both associative and commutative. It should be noted that in general $B_{r,i} * B_{s,j} \neq B_{r+s,i+j}$.

We now state and prove a lemma on certain monotonicities among these Bonferroni functions which will help us to deduce (1.1) and (1.2) when the indicator random variables are exchangeable.

Lemma:

For any fixed $k \geq 2$ with $0 \leq \nu - 1 \leq k$ and for any partition of the index set $(1, 2,\ldots, k)$ into 2 parts of sizes 1 and $k - 1$, we have

(2.3) $$B_{\nu-1,k} \leq B_{1,1} * B_{\nu-1,k-1} \leq B_{k,k} \leq B_{\nu,k} \quad \text{for } \nu \text{ odd,}$$

(2.4) $$B_{\nu-1,k} \geq B_{1,1} * B_{\nu-1,k-1} \geq B_{k,k} \geq B_{\nu,k} \quad \text{for } \nu \text{ even.}$$

Proof:

By direct substitution of (2.1) into the second expressio-. -f (2.3) we have for any ν

(2.5) $$B_{1,1} * B_{\nu-1,k-1} = B_{\nu-1,k} + (-1)^{\nu-1} X_1 \sum_{1<\alpha_1<\ldots<\alpha_{\nu-1}} X_{\alpha_1} \cdots X_{\alpha_{\nu-1}}$$

which proves the first inequality in each of (2.3) and (2.4). To prove the second inequalities in each of (2.3) and (2.4) we iterate the inequality implied by (2.5) for ν odd and even obtaining, respectively,

(2.6) $$B_{\nu-1,k} \leq B_{1,1} * B_{\nu-1,k-1} \leq B_{2,2} * B_{\nu-1,k-2} \leq \cdots \leq B_{k,k} \leq 1,$$

(2.7) $$B_{\nu-1,k} \geq B_{1,1} * B_{\nu-1,k-1} \geq B_{2,2} * B_{\nu-1,k-2} \geq \cdots \geq B_{k,k} \geq 0.$$

The remaining inequality in each of (2.3) and (2.4) is the well-known Bonferroni bound; this completes the proof of the lemma. The first inequality of (2.7) with $\nu = 2$, namely $B_{1,k} \geq B_{1,1} * B_{1,k-1}$, is the inequality that appears in Kounias [4]. Of course, this can be further improved, possibly by using $B_{3,k}$, but for the fixed degree 2 neither the Bonferroni bound $B_{3,k}$ nor the improved bound in (2.7), $B_{2,2} * B_{1,k-2}$, are allowed, since both are of degree 3.

For exchangeable indicator random variables we take expectations in (2.5) and obtain for any ν

(2.8) $$E\{B_{1,1}{*}\, B_{\nu-1,k-1}\} = \binom{k}{1}P_1 - \binom{k}{2}P_2 + \dots + (-1)^{\nu-2}\binom{k}{\nu-1}P_{\nu-1} + (-1)^{\nu-1}\binom{k-1}{\nu-1}P_\nu.$$

Hence from the last two inequalities in (2.3) and (2.4), we obtain

(2.9) $$\sum_{\alpha=1}^{\nu-1}(-1)^{\alpha-1}\binom{k}{\alpha}P_\alpha + \binom{k-1}{\nu-1}P_\nu \le P\{\bigcup_{i=1}^{k} A_i\} \le E\{B_{\nu,k}\} \quad (\text{for } \nu \text{ odd}),$$

(2.10) $$\sum_{\alpha=1}^{\nu-1}(-1)^{\alpha-1}\binom{k}{\alpha}P_\alpha - \binom{k-1}{\nu-1}P_\nu \ge P\{\bigcup_{i=1}^{k} A_i\} \ge E\{B_{\nu,k}\} \quad (\text{for } \nu \text{ even}),$$

which are equivalent to (1.1) and (1.2).

It is clear from (2.3) and (2.4) that our bounds are improvements on $E\{B_{\nu-1,k}\}$. This has been achieved by increasing the degree from $\nu - 1$ to ν for the new lower bound in (2.3) for ν odd and for the new upper bound in (2.4) for ν even.

3. Further Improvement of the Bounds.

From the inequalities (2.6) and (2.7) we can actually obtain bounds which are as good or better than those in (1.1) and (1.2). Since these are quite useful for small ν we write them explicitly for $\nu = 2, 3$ and 4 and then give a general expression at the end of this section.

To derive these improved bounds of common second degree we set $\nu = 3$ in (2.6) and $\nu = 2$ in (2.7) and obtain after taking expectations

(3.1) $$\mathrm{Max}(0,\ E\{B_{2,k}\}) \le P\{\bigcup_{i=1}^{k} A_i\} \le \mathrm{Min}(E\{B_{1,1}{*}\, B_{1,k-1}\},\ 1).$$

For the bounds of common third degree we set $\nu = 3$ in (2.6), $\nu = 4$ and $\nu = 2$ in (2.7) obtaining

(3.2) $$\mathrm{Max}(0,\ E\{B_{1,1}{*}\, B_{2,k-1}\}) \le P\{\bigcup_{i=1}^{k} A_i\} \le \mathrm{Min}(E\{B_{3,k}\},\ E\{B_{2,2}{*}\, B_{1,k-2}\},\ 1).$$

For the bounds of common fourth degree we set $\nu = 5$ and $\nu = 3$ in (2.6) and $\nu = 4$ and 2 in (2.7) obtaining

$$(3.3)\qquad \mathrm{Max}(0,\ E\{B_{4,k}\},\ E\{B_{2,2} * B_{2,k-2}\}) \le P\{\bigcup_{i=1}^{k} A_i\}$$

$$\le \mathrm{Min}(E\{B_{1,1} * B_{3,k-1}\},\ E\{B_{3,3} * B_{1,k-3}\},\ 1).$$

In (3.1) no new expressions arise and the result agrees with (1.2) for degree 2. In (3.2) two of the expressions already appear in (1.1) for degree 3; the new expression for the exchangeable case with any $k \ge 3$ is

$$(3.4)\qquad E\{B_{2,2} * B_{1,k-2}\} = kP_1 - (2k-3)P_2 + (k-2)P_3 .$$

In (3.3) two of the expressions already appear in (1.2) for degree 4; the two new expressions for the exchangeable case with any $k \ge 4$ are

$$(3.5)\qquad E\{B_{2,2} * B_{2,k-2}\} = kP_1 - \binom{k}{2}P_2 + (k-2)^2 P_3 - \binom{k-2}{2}P_4,$$

$$(3.6)\qquad E\{B_{3,3} * B_{1,k-3}\} = kP_1 - 3(k-2)P_2 + (3k-8)P_3 - (k-3)P_4.$$

It is interesting to note that for any Bonferroni product of two B's as in (3.4), (3.5) and (3.6) the sum of the coefficients is identically one. In fact, if Q denotes the operation of setting all χ_i equal to 1

$$(3.7)\qquad Q\{B_{r,r} * B_{s,j}\} = Q\{1 - (1-B_{r,r})(1-B_{s,j})\} = 1 - Q(1-B_{r,r})Q(1-B_{s,j}) = 1$$

since $Q\{1 - B_{r,r}\} = \sum_{\alpha=0}^{r} (-1)^{\alpha}\binom{r}{\alpha} = 0.$

More generally, it can be shown (the proof is omitted) that in symbolic notation we can write the $\nu^{\underline{th}}$ degree bounds as

$$(3.8)\qquad E\{B_{j,j} * B_{\nu-j,k-j}\} = 1 - (1-P)^j \sum_{i=0}^{\nu-j}(-1)^i\binom{k-j}{i}P^i,$$

where, after expanding, we replace P^{α} by P_{α} for all α. For any fixed degree ν this gives a lower (upper) bound if $\nu - j$ is even (odd). As j varies we obtain all of our bounds.

In the next section we shall illustrate these bounds by means of a numerical example involving Dirichlet integrals.

4. Application to Dirichlet Integrals.

Suppose n balls are dropped independently into $k + 1$ cells, k of which have a common single-trial cell probability $p \leq 1/k$ (and one with probability $1 - kp$). Let $\text{Min}(k, n)$ denote the observed minimum of the k cell frequencies in this multinomial problem. Then it is shown in [6] that for $r \geq 1$ and $n \geq kr$

$$(4.1) \qquad P\{\text{Min}(k, n) \geq r\} = I_p^{(k)}(r, n),$$

where

$$(4.2) \qquad I_p^{(k)}(r, n) = \frac{\Gamma(n+1)}{\Gamma^k(r)\Gamma(n-kr+1)} \int_0^p \cdots \int_0^p \left(1 - \sum_{i=1}^{k} x_i\right)^{n-kr} \prod_{\alpha=1}^{k} x_\alpha^{r-1} dx_\alpha .$$

For $n < kr$ we define $I_p^{(k)}(r, n)$ to be zero. Similarly we let $\text{Max}(k, n)$ denote the observed maximum of the k cell frequencies. In [7] the properties of this type I-Dirichlet integral (4.2) have been investigated and tables have been computed in [3]; for $k = 1$ it reduces to the incomplete beta function, $I_p^{(1)}(r, n) = I_p(r, n-r+1)$.

If we let the set A_i denote the event that the frequency in the i^{th} cell is at least r, then $\bigcup_{i=1}^{k} A_i$ denotes the event that $\text{Max}(k, n) \geq r$ and the event $A_{\alpha_1} A_{\alpha_2} \cdots A_{\alpha_j}$ (for any collection of size j) denotes the event that $\text{Min}(k, n) \geq r$. Hence by using the inclusion-exclusion principle, it is easy to see that

$$(4.3) \qquad P\{\text{Max}(k, n) \geq r\} = E\{B_{k,k}\} = \sum_{\alpha=1}^{k} (-1)^{\alpha-1} \binom{k}{\alpha} P\{\text{Min}(\alpha, n) \geq r\}$$

$$= \sum_{\alpha=1}^{k} (-1)^{\alpha-1} \binom{k}{\alpha} I_p^{(\alpha)}(r, n).$$

If $n < kr$ then some of the terms on the right side of (4.3) will vanish so that we need only sum up to $[n/r]$, the integer part of n/r. The distribution of $\text{Max}(k, n)$ can be obtained from (4.3) using tables for $I_p^{(\alpha)}(r, n)$ $(\alpha \leq k)$. However for higher values of k some of these tables are not available and it becomes useful to obtain bounds for the left side of (4.3), preferably of the same degree. Thus we utilize the results of sections 2 and 3 as illustrated below.

Suppose, for example, we wish to find lower and upper bounds of degree 3 for $P\{Max(k, n) \geq r\}$ when $k = 10$, $r = 2$ and $n = 8$. Using (3.2) and (3.4) for the upper bound we obtain .98715 since

$$(4.4) \qquad E\{B_{3,10}\} = .98715 \ ; \ E\{B_{22} * B_{1,8}\} > 1.$$

For the lower bound we use (3.2) or (1.1) and obtain

$$(4.5) \qquad E\{B_{1,1} * B_{2,9}\} = 10P_1 - 45P_2 + 36P_3 = .85132.$$

The exact value, using (4.3) and tabulated values in [3], is .98186. The comparable bounds given by Mallows [5] are

$$(4.6) \qquad Max(1 - e^{-kP_1}, kP_1 - \binom{k}{2}P_2) \leq P\{Max(k, n) \geq r\} \leq Min(kP_1, 1)$$

and for the above example with $k = 10$, $r = 2$, $n = 8$ this gives

$$(4.7) \qquad Max(.84572, .79311) \leq P\{Max(10, 8) \geq 2\} \leq Min(1.86895, 1),$$

so that both of our bounds for the third degree are closer to the exact value than the corresponding bounds in [5].

For many values of k, r, and n it should be noted that 'degeneracies' will enter into our problem in different possible ways. For example, if $n \geq kr$ then

$$(4.8) \qquad P\{Max(k, n) \geq r\} = E\{B_{k,k}\} = 1,$$

since for any distribution of n balls in k cells the maximum frequency is at least $[n/k] \geq r$. Another type of 'degeneracy' is when $E\{B_{\nu,k}\}$ already gives the exact answer. For example, if $[n/r] \geq \nu$ then $I_p^{(j)}(r, n) = 0$ for $j > \nu$ and we seen from (4.3) that

$$(4.9) \qquad P\{\mathrm{Max}(k, n) \geq r\} = E\{B_{k,k}\} = E\{B_{\nu,k}\},$$

so that the $\nu^{\underline{th}}$ degree bounds give exact answers. In this sense the bounds in (1.1) and (1.2) cannot be further sharpened.

5. Other Applications to Dirichlet Integrals.

In the previous section we developed bounds of the $\nu^{\underline{th}}$ degree for the $P\{\mathrm{Max}(k, n) \geq r\}$ in terms of $P\{\mathrm{Min}(i, n) \geq r\}$ for $1 \leq i \leq \nu$. However, if k is large, tables of $P\{\mathrm{Min}(k, n) \geq r\}$ itself are not available and it is useful to have $\nu^{\underline{th}}$ order bounds for this quantity (for any k) in terms of $P\{\mathrm{Min}(i, n) \geq r\}$ with $1 \leq i \leq \nu$. We first develop these bounds in the context of exchangeable random variables and then specialize to the case of multinomial probabilities.

Let $\widetilde{A}_i$ denote the complement of A_i $(i = 1,2,\ldots, k)$ and let $\widetilde{B}_{r,j}$ with $j \leq k$ and $1 \leq r \leq k$ denote the same Bonferroni function as in (2.1) for the j sets $\widetilde{A}_{i_1}, \widetilde{A}_{i_2}, \ldots, \widetilde{A}_{i_j}$. We now use (3.2) with all B's replaced by $\widetilde{B}$'s to obtain, after taking expectations, for the case $\nu = 3$

$$(5.1) \qquad \mathrm{Max}(0, E\{\widetilde{B}_{1,1} * \widetilde{B}_{2,k-1}\}) \leq P\{\bigcup_{i=1}^{k} \widetilde{A}_i\} \leq \mathrm{Min}(E\{\widetilde{B}_{3,k}\}, E\{\widetilde{B}_{2,2} * \widetilde{B}_{1,k-2}\}, 1).$$

For the exchangeable case we now let $\widetilde{Q}_i$ denote $P\{\widetilde{A}_{\alpha_1} \widetilde{A}_{\alpha_2} \ldots \widetilde{A}_{\alpha_i}\}$ and obtain from (5.1) as in (3.3)

$$(5.2) \qquad \mathrm{Max}(0, k\widetilde{Q}_1 - \binom{k}{2}\widetilde{Q}_2 + \binom{k-1}{2}\widetilde{Q}_3) \leq P\{\bigcup_{i=1}^{k} \widetilde{A}_i\}$$

$$\leq \mathrm{Min}(k\widetilde{Q}_1 - \binom{k}{2}\widetilde{Q}_2 + \binom{k}{3}\widetilde{Q}_3, k\widetilde{Q}_1 - (2k-3)\widetilde{Q}_2 + (k-2)\widetilde{Q}_3, 1).$$

Using the inclusion-exclusion principle, we find that $\tilde{Q}_i$ for any i can be written as

$$(5.3) \qquad \tilde{Q}_i = \sum_{j=0}^{i} (-1)^j \binom{i}{j} P_j$$

where $P_j = P\{A_{\alpha_1} A_{\alpha_2} \cdots A_{\alpha_j}\}$ as in section 1 and $P_0 = 1$. Substituting these in (5.2) gives us for the third degree

$$(5.4) \qquad \text{Max}(0,\ 1 - \binom{k-2}{2}P_1 + (k-1)(k-3)P_2 - \binom{k-1}{2}P_3) \leq P\{\bigcup_{i=1}^{k} \tilde{A}_i\}$$

$$\leq \text{Min}(1 + \binom{k-1}{3} - k\binom{k-2}{2}P_1 + (k-3)\binom{k}{2}P_2 - \binom{k}{3}P_3,\ 1 + (k-3)P_2 - (k-2)P_3, 1).$$

For the multinomial case (where k cells have common probability p and one cell has probability $1 - kp$), $\tilde{A}_i$ denotes the event that the frequency of the i^{th} cell is less than r, i.e., $N_{i,n} < r$ and hence $\bigcup_{i=1}^{k} \tilde{A}_i$ denotes the event that $\text{Min}(k, n) < r$. Since $I_p^{(k)}(r, n) = P\{\text{Min}(k,n) \geq r\} = P_k$, we take the complement of all three expressions in (5.4) and obtain the final result for $k \geq 3$

$$(5.5) \qquad \text{Max}(\binom{k}{3}P_3 - (k-3)\binom{k}{2}P_2 + k\binom{k-2}{2}P_1 - \binom{k-1}{3},\ (k-2)P_3 - (k-3)P_2,\ 0) \leq I_p^{(k)}(r,n)$$

$$= P_k \leq \text{Min}(P_3,\ \binom{k-1}{2}P_3 - (k-1)(k-3)P_2 + \binom{k-2}{2}P_1),$$

where we have replaced the upper bound 1 by $P_3 \leq 1$. For $k = 3$ we obtain from (5.5) the "identity check" $P_3 \leq P_3 \leq P_3$ and in this sense the bounds are sharp.

For $\nu = 2$ and $\nu = 4$ we give the corresponding bounds for P_k in terms of $P_1, \ldots, P_\nu$ without derivation. For $\nu = 2$ and any $k \geq 2$

$$(5.6)\qquad \mathrm{Max}\{(k-1)P_2-(k-2)P_1,\ 0\} \le P_k \le \mathrm{Min}\{P_2,\ \binom{k}{2}P_2-k(k-2)P_1+\binom{k-1}{2}\}.$$

For $\nu = 4$ and any $k \geq 4$

$$(5.7)\qquad \mathrm{Max}\{\binom{k-1}{3}P_4-(k-4)\binom{k-1}{2}P_3+(k-1)\binom{k-3}{2}P_2+\binom{k-2}{3}P_1,\ (k-3)P_4-(k-4)P_3,\ 0\}$$

$$\le P_k \le \mathrm{Min}\{P_4,\ \binom{k-2}{2}P_4-(k-2)(k-4)P_3+\binom{k-3}{2}P_2,\ \binom{k}{4}P_4-(k-4)\binom{k}{3}P_3$$

$$+\binom{k}{2}\binom{k-3}{2}P_2-k\binom{k-2}{3}P_1+\binom{k-1}{4}\}.$$

In a similar manner we can obtain bounds of the $\nu^{\underline{th}}$ degree for P_k in terms of $P_1, P_2, \ldots, P_\nu$ for any ν. In fact, a general expression (given without proof) which gives all of our bounds for any degree ν is given by

$$(5.8)\qquad \mathrm{Max}\{L_1,\ L_3, \ldots,\ L_{\nu_0}\} \le P_k \le \mathrm{Min}\{L_2,\ L_4, \ldots,\ L_{\nu_e}\},$$

where $\nu_0(\nu_e)$ is the largest odd (even) integer $\leq \nu$ and

$$(5.9)\qquad L_\beta = \sum_{\alpha=0}^{\beta} (-1)^\alpha \binom{k-\nu-1+\alpha}{\alpha}\binom{k-\nu+\beta}{\beta-\alpha} P_{\nu-\alpha}\ .$$

A device to express (5.9) more succinctly is to write the generating function of L_β

$$(5.10)\qquad L(x) = \sum_{\beta=0}^{\nu} L_\beta\, x^\beta$$

in symbolic notation. It can be shown (proof is omitted) that we can write $L(x)$ symbolically as

$$(5.11)\qquad L(x) = (-1)^\beta P^{\nu-\beta}(1-x)^{-(k-\nu)}(1-Px)^{k-\nu+\beta}$$

where, after obtaining the coefficient of x^β, we replace P^j by P_j. In the interesting special case $k - \nu = 1$ we have

(5.12) $$L_\beta = \sum_{\alpha=0}^{\beta} (-1)^{\beta+\alpha}\binom{\beta+1}{\alpha} P_{\nu-\beta+\alpha}$$

and hence from (5.8) we can write all of our bounds in the single symbolic form for all β $(0 \leq \beta \leq \nu)$

(5.13) $$P^{\nu-\beta}(1-P)^{\beta+1} = (-1)^{\beta}(L_\beta - P_k) \geq 0.$$

An immediate corollary of (5.13) for $\beta = 1$ is that the sequence $P_1, P_2, P_3, \ldots$ is convex with respect to the subscript since for any $\nu \leq k$

(5.14) $$P_{\nu-1} + P_{\nu+1} \geq 2P_\nu \text{ .}$$

It should be noted that all of the above inequalities (5.1) - (5.14) are valid for any set of exchangeable events, though we are discussing them in the context of the multinomial distribution.

To illustrate numerically the third degree bounds in (5.5), consider the multinomial example with $k = 10$, $r = 2$, $n = 40$ and $p = .09$. In this case the exact value of $P_{10} = I_{.09}^{(10)}(2, 40)$, available from [3], is .24434. Our lower bound is Max(.23809, .02826, 0) = .23809 and the upper bound is Min(.29487, .68801) = .29487.

This idea of bounding P_k can also be used for the case of a multinomial distribution with unequal cell probabilities $\underset{\sim}{p} = (p_1, p_2, \ldots, p_k \,;\, 1 - \sum_1^k p_i)$. Alam [1] has already shown the interesting property that

(5.15) $$I_{\underset{\sim}{p}}^{(k)}(r, n) \leq I_{\overline{p}}^{(k)}(r, n)$$

where $\overline{p} = (p_1 + p_2 + \ldots + p_k)/k$ and $I_{\underset{\sim}{p}}^{(k)}(r, n)$ is defined by

$$(5.16)\qquad I_{\underline{p}}^{(k)}(r, n) = \frac{\Gamma(n+1)}{[\Gamma(r)]^k\Gamma(n-kr+1)} \int_0^{p_1} \cdots \int_0^{p_k} (1 - \sum_{i=1}^{k} x_i)^{n-kr} \prod_{i=1}^{k} x_i^{r-1} dx_i .$$

In getting an upper bound for the left-side of (5.15) it may happen (for example, when k is large) that the right side of (5.15) is unavailable. Then our bound for $I_p^{(k)}(r, n)$ given for $\nu = 3$ in (5.5) can then be used as an alternative upper bound for the left side of (5.15).

For the lower bound Alam [1] gives

$$(5.17)\qquad I_{\underline{p}}^{(k)}(r, n) \geq \frac{\Gamma(n+1)}{[\Gamma(r)]^k\Gamma(n-kr+1)} \int_0^{p^*} \cdots \int_0^{p^*} \int_0^{p^{**}} (1 - \sum_{1}^{k} x_i)^{n-kr} \prod_{i=1}^{k} x_i^{r-1} dx_i ,$$

where $p^* = \min(p_1, p_2, \ldots, p_k)$ and $(k-1)p^* + p^{**} = \sum_{i=1}^{k} p_i$. In this case our lower bound cannot be directly used for the right side of (5.17) since we do not have homogeneity in the k cell probabilities. A lower bound can, however, be obtained by writing

$$(5.18)\qquad P\{\bigcup_{i=1}^{k} A_i\} \leq P\{\bigcup_{i=1}^{k-1} A_i\} + P\{A_k\}$$

where A_i denotes the event that the frequency $N_{i,n}$ of the $i^{\underline{th}}$ cell is at least r. We apply this to the case of a multinomial distribution with $k - 1$ cells having probability p^* and the $k^{\underline{th}}$ cell having probability p^{**}, and obtain, by taking complements in (5.18),

$$(5.19)\qquad P\{\mathrm{Min}(k, n) \geq r\} \geq P\{\mathrm{Min}(k-1, n) \geq r\} + P\{N_{k,n} \geq r\} - 1$$
$$= I_{p^*}^{(k-1)}(r, n) + I_{p^{**}}^{(1)}(r, n) - 1.$$

Since the right side of (5.17) equals the left side of (5.19) we can use the right side of (5.19) as a lower bound for $I_{\underline{p}}^{(k)}(r, n)$ in (5.17) (when both members of (5.17) are not available). If k is large, so that $I_{p^*}^{(k-1)}(r, n)$ is also not available, then we can substitute the lower bounds from (5.5) into (5.19).

Acknowledgement.

The authors wish to thank Dr. G. P. Steck of Sandia Corporation for making available to us his tables [8].

REFERENCES

[1] Alam, K. (1970). Monotonicity properties of the multinomial distribution. Ann. Math. Statist. 41 315-317.

[2] Feller, W. (1968). An Introduction to Probability Theory and its Applications, Volume 1. Wiley, New York.

[3] Frankowski, K., Sobel, M., and Uppuluri, V. R. R. (1971). Tables of the Type I -Dirichlet integral, and applications. (In preparation.)

[4] Kounias, E. G. (1968). Bounds for the probability of a union, with applications. Ann. Math. Statist. 39 2154-2158.

[5] Mallows, C. L. (1968). An inequality involving multinomial probabilities. Biometrika 55 422-424.

[6] Olkin, I. and Sobel, M. (1965). Integral expressions for tail probabilities of the multinomial and negative multinomial distributions. Biometrika 52 167-179.

[7] Sobel, M. and Uppuluri, V. R. R. (1971). Some properties and applications of Dirichlet integrals. (In preparation.)

[8] Steck, G. P. (1960). Tables of the distribution of the maximum of r jointly multinomially distributed variables with all cell probabilities equal, $r = 2(1)50$, $N = 1(1)50$. Unpublished.