A Unified Approach to Jordan's Formula, Inclusion-exclusion Principle and Bonferroni Inequalities

Jay Cheng
Department of Electrical Engineering
National Tsing Hua University, Hsinchu, Taiwan 30013, R.O.C.
jcheng@ee.nthu.edu.tw

Abstract

For n given events in a probability space and $1 \leq r \leq n$, we first give a Bonferroni-type formula and a class of *sharp* Bonferroni-type inequalities for the probability that exactly r out of the n given events occur and for the probability that at least r out of the n given events occur. Then we show that our Bonferroni-type formulae and inequalities serve as a unified menthod for obtaining the well-known Jordan's formula, the generalized inclusion-exclusion principle, and the classical Bonferroni inequalities.

Keywords. Bonferroni inequalities, inclusion-exclusion principle, Jordan's formula.

1 Introduction

Let $(\Omega, \mathcal{A}, P)$ be a probability space, where Ω is the sample space underlying the probability space, $\mathcal{A}$ is a σ-algebra of subsets of the sample space Ω, and P is a probability measure on the σ-algebra $\mathcal{A}$. Let $A_1, A_2, \ldots, A_n$ be events in $\mathcal{A}$. Let $[n] = \{1, 2, \ldots, n\}$ and let $A_I = \cap_{i \in I} A_i$ be the event that the events A_i, $i \in I$, occur for all $I \subseteq [n]$ such that $I \neq \emptyset$. Let S_k be the kth *binomial moment* of the events $A_1, A_2, \ldots, A_n$ that is given by [1]

$$S_k = \sum_{\substack{I \subseteq [n] \\ |I| = k}} P(A_I), \text{ for } 1 \leq k \leq n. \tag{1}$$

Let E_r and F_r be the events that exactly r events and at least r events, respectively, out of the n events $A_1, A_2, \ldots, A_n$ occur for $r \geq 1$. Note that the events $E_1, E_2, \ldots, E_n$ are pairwise disjoint and $F_r = \cup_{\ell=r}^{n} E_\ell$ for $1 \leq r \leq n$. For $r > n$, it is clear that $P(E_r) = P(F_r) = 0$. For $1 \leq r \leq n$, the probability $P(E_r)$ of the event E_r can be expressed in terms of the binomial moments $S_r, S_{r+1}, \ldots, S_n$ as given in the well-known Jordan's formula (1867) as follows [2]:

$$P(E_r) = \sum_{k=r}^{n} (-1)^{k-r} \binom{k}{r} S_k, \text{ for } 1 \leq r \leq n. \tag{2}$$

For $1 \leq r \leq n$, the probability $P(F_r)$ of the event F_r can also be expressed in terms of the binomial moments $S_r, S_{r+1}, \ldots, S_n$ as given in the generalized inclusion-exclusion principle as follows [3]:

$$P(F_r) = \sum_{k=r}^{n} (-1)^{k-r} \binom{k-1}{r-1} S_k, \text{ for } 1 \leq r \leq n. \tag{3}$$

For the special case that $r = 1$, the generalized inclusion-exclusion principle in (3) reduces to the classical inclusion-exclusion principle for the probability of the union $\cup_{i=1}^{n} A_i$ of the events $A_1, A_2, \ldots, A_n$ (note that

$F_1 = \cup_{i=1}^n A_i$) as follows [4]–[6]:

$$P(\cup_{i=1}^n A_i) = \sum_{k=1}^{n} (-1)^{k-1} S_k. \tag{4}$$

Note that the classical inclusion-exclusion principle in (4) originates from the idea of Abraham de Moivre (1718) [4], and is also known as Da Silva's formula (1854) [5] or Sylvester's formula (1883) [6].

To compute the probability $P(E_r)$ by using (2) or the probability $P(F_r)$ by using (3), we need to compute the binomial moments $S_r, S_{r+1}, \ldots, S_n$ as given in (1). However, since it is clear that the number of terms in (1) grows exponentially with n, in practice such an approach for the computation of $P(E_r)$ or $P(F_r)$ may not be feasible as n becomes large. In many realistic applications, a few binomial moments are either known/given or can be computed/estimated from historical data. By using the binomial moments available, say $S_r, S_{r+1}, \ldots, S_m$ are available, where $1 \leq r \leq m \leq n$, the classical Bonferroni inequalities (1936) [7] (also see [1]) give upper and lower bounds for $P(E_r)$ and $P(F_r)$ in terms of $S_r, S_{r+1}, \ldots, S_m$ as follows:

$$P(E_r) \begin{cases} \leq \sum_{k=r}^{m} (-1)^{k-r} \binom{k}{r} S_k, & \text{if } m-r \text{ is even,} \\ \geq \sum_{k=r}^{m} (-1)^{k-r} \binom{k}{r} S_k, & \text{if } m-r \text{ is odd,} \end{cases} \tag{5}$$

and

$$P(F_r) \begin{cases} \leq \sum_{k=r}^{m} (-1)^{k-r} \binom{k-1}{r-1} S_k, & \text{if } m-r \text{ is even,} \\ \geq \sum_{k=r}^{m} (-1)^{k-r} \binom{k-1}{r-1} S_k, & \text{if } m-r \text{ is odd.} \end{cases} \tag{6}$$

Note that for the special case that $r = 1$, the Bonferroni inequalities in (6) reduce to the classical inclusion-exclusion inequalities for the probability $P(\cup_{i=1}^n A_i)$ of the union $\cup_{i=1}^n A_i$ of the events $A_1, A_2, \ldots, A_n$ [4]–[6].

In this paper, we present a unified approach to Jordan's formula in (2), the generalized inclusion-exclusion principle in (3), and the classical Bonferroni inequalities in (5) and (6). We first give a Bonferroni-type formula for the probability $P(E_r)$, and show that it subsumes Jordan's formula in (2) as a special case. Our formula for $P(E_r)$ leads to a class of sharp Bonferroni-type inequalities for $P(E_r)$, which improve on the Bonferroni inequalities in (5). Furthermore, our sharp inequalities for $P(E_r)$ not only subsume the Bonferroni inequalities in (5) as a special case, but also give rise to a necessary and sufficient condition for the inequalities in (5) to hold with equality. Then we give a Bonferroni-type formula for the probability $P(F_r)$ that subsumes the generalized inclusion-exclusion principle in (3) as a special case. We also give a class of sharp Bonferroni-type inequalities for $P(F_r)$, which not only improve on and subsume as a special case the Bonferroni inequalities in (6), but also give rise to a necessary and sufficient condition for the inequalities in (6) to hold with equality.

2 The Unified Approach

In the following theorem, we give our Bonferroni-type formula and a class of sharp Bonferroni-type inequalities for the probability $P(E_r)$ of the event E_r.

Theorem 1 *Suppose $1 \leq r \leq m \leq n$.*

(i) The probability $P(E_r)$ of the event E_r that exactly r events out of the n events $A_1, A_2, \ldots, A_n$ occur can be expressed in terms of $S_r, S_{r+1}, \ldots, S_m$ and $P(E_{m+1}), P(E_{m+2}), \ldots, P(E_n)$ as given in the following Bonferroni-type formula:

$$P(E_r) = \sum_{k=r}^{m} (-1)^{k-r} \binom{k}{r} S_k + (-1)^{m-r+1} \sum_{\ell=m+1}^{n} \binom{\ell}{r} \binom{\ell-r-1}{m-r} P(E_\ell). \tag{7}$$

(ii) Let $[m+1,n]=\{m+1,m+2,\ldots,n\}$. For each $J\subseteq[m+1,n]$, the probability $P(E_r)$ of the event E_r can be bounded as given in the following Bonferroni-type inequalities:

$$P(E_r)\begin{cases}\le \sum_{k=r}^{m}(-1)^{k-r}\binom{k}{r}S_k-\sum_{\ell\in J}\binom{\ell}{r}\binom{\ell-r-1}{m-r}P(E_\ell), & \text{if } m-r \text{ is even},\\ \ge \sum_{k=r}^{m}(-1)^{k-r}\binom{k}{r}S_k+\sum_{\ell\in J}\binom{\ell}{r}\binom{\ell-r-1}{m-r}P(E_\ell), & \text{if } m-r \text{ is odd},\end{cases} \tag{8}$$

where the inequalities in (8) hold with equality if and only if $P(E_\ell)=0$ for all $\ell\in[m+1,n]\backslash J$.

Remark 2 (i) Jordan's formula in (2) is a special case of our Bonferroni-type formula in (7). To see this, consider the special case that $m=n-1$. Then we see from $P(E_n)=P(A_{[n]})=S_n$ that (7) reduces to (2) as follows:

$$P(E_r)=\sum_{k=r}^{n-1}(-1)^{k-r}\binom{k}{r}S_k+(-1)^{n-r}\binom{n}{r}P(E_n)=\sum_{k=r}^{n}(-1)^{k-r}\binom{k}{r}S_k.$$

Alternatively, consider the special case that $A_i=\emptyset$ for all $m+1\le i\le n$. Then we have $P(E_\ell)=0$ for all $m+1\le\ell\le n$, and hence (7) reduces to

$$P(E_r)=\sum_{k=r}^{m}(-1)^{k-r}\binom{k}{r}S_k. \tag{9}$$

By removing all the terms in the S_k's in (9) involving $A_{m+1},A_{m+2},\ldots,A_n$ that have no contributions to the values of the S_k's in (9), we obtain (2) (with n in (2) replaced by m).

(ii) It is easy to see from $P(E_\ell)\ge$ for all $1\le\ell\le n$ that our Bonferroni-type inequalities in (8) improve on the classical Bonferroni inequalities in (5). Furthermore, for the special case that $J=\emptyset$, our inequalities in (8) reduce to the inequalities in (5), and the inequalities in (5) hold with equality if and only if $P(E_\ell)=0$ for all $m+1\le\ell\le n$, or, equivalently, $P(F_{m+1})=0$.

(Proof of Theorem 1) (i) Let $B_J=(\cap_{j\in J}A_j)\cap(\cap_{j\in[n]\backslash J}A_j^c)$ be the event that exactly the events $A_j,\ j\in J$, among the events $A_1,A_2,\ldots,A_n$ occur for all $J\subseteq[n]$. It is clear that the events $B_J,\ J\subseteq[n]$, are pairwise disjoint and $\cup_{J\subseteq[n]}B_J=\Omega$. It is also clear that $A_I=\cup_{I\subseteq J\subseteq[n]}B_J=\cup_{\ell=|I|}^{n}\cup_{I\subseteq J\subseteq[n],|J|=\ell}B_J$ for all $I\subseteq[n]$ such that $I\ne\emptyset$, and $E_\ell=\cup_{J\subseteq[n],|J|=\ell}B_J$ for all $1\le\ell\le n$.

In our proof, we will use the following identity from [1] (which can be obtained by first writing $\binom{k}{r}\binom{\ell}{k}=\binom{\ell}{r}\binom{\ell-r}{k-r}$, and then using binomial theorem for the case that $m=\ell$, and writing $\binom{\ell-r}{k-r}=\binom{\ell-r-1}{k-r-1}+\binom{\ell-r-1}{k-r}$ for $r+1\le k\le m$ and telescoping for the case that $m\le\ell-1$):

$$\sum_{k=r}^{m}(-1)^{k-r}\binom{k}{r}\binom{\ell}{k}=\begin{cases}0, & \text{if } r<m=\ell,\\ 1, & \text{if } r=m=\ell,\\ (-1)^{m-r}\binom{\ell}{r}\binom{\ell-r-1}{m-r}, & \text{if } m\le\ell-1.\end{cases} \tag{10}$$

Now we obtain (7) as follows:

$$\begin{aligned}&\sum_{k=r}^{m}(-1)^{k-r}\binom{k}{r}S_k\\ &=\sum_{k=r}^{m}(-1)^{k-r}\binom{k}{r}\sum_{\substack{I\subseteq[n]\\|I|=k}}\left(\sum_{\ell=k}^{n}\sum_{\substack{I\subseteq J\subseteq[n]\\|J|=\ell}}P(B_J)\right)\end{aligned}$$

$$= \sum_{\ell=r}^{m} \sum_{\substack{J\subseteq[n]\\|J|=\ell}} \left(\sum_{k=r}^{\ell} (-1)^{k-r} \binom{k}{r} \binom{\ell}{k} \right) P(B_J) + \sum_{\ell=m+1}^{n} \sum_{\substack{J\subseteq[n]\\|J|=\ell}} \left(\sum_{k=r}^{m} (-1)^{k-r} \binom{k}{r} \binom{\ell}{k} \right) P(B_J) \tag{11}$$

$$= \sum_{\substack{J\subseteq[n]\\|J|=r}} 1 \cdot P(B_J) + \sum_{\ell=r+1}^{m} \sum_{\substack{J\subseteq[n]\\|J|=\ell}} 0 \cdot P(B_J) + \sum_{\ell=m+1}^{n} \sum_{\substack{J\subseteq[n]\\|J|=\ell}} (-1)^{m-r} \binom{\ell}{r} \binom{\ell-r-1}{m-r} P(B_J)$$

$$= P(E_r) + (-1)^{m-r} \sum_{\ell=m+1}^{n} \binom{\ell}{r} \binom{\ell-r-1}{m-r} P(E_\ell),$$

where the first equality follows from (1) and $P(A_I) = P(\cup_{\ell=|I|}^{n} \cup_{I\subseteq J\subseteq[n],|J|=\ell} B_J) = \Sigma_{\ell=|I|}^{n} \Sigma_{I\subseteq J\subseteq[n],|J|=\ell} P(B_J)$ for all $I \subseteq [n]$ such that $I \neq \emptyset$, the third equality follows from (10), and the last equality follows from $P(E_\ell) = P(\cup_{J\subseteq[n],|J|=\ell} B_J) = \Sigma_{J\subseteq[n],|J|=\ell} P(B_J)$ for all $1 \leq \ell \leq n$.

(ii) Since $P(E_\ell) \geq$ for all $1 \leq \ell \leq n$, it is clear that (ii) follows from (i). ■

In the following theorem, we give our Bonferroni-type formula and a class of sharp Bonferroni-type inequalities for the probability $P(F_r)$ of the event F_r.

Theorem 3 *Suppose* $1 \leq r \leq m \leq n$.

(i) The probability $P(F_r)$ *of the event* F_r *that at least* r *events out of the* n *events* $A_1, A_2, \ldots, A_n$ *occur can be expressed in terms of* $S_r, S_{r+1}, \ldots, S_m$ *and* $P(E_{m+1}), P(E_{m+2}), \ldots, P(E_n)$ *as given in the following Bonferroni-type formula:*

$$P(F_r) = \sum_{k=r}^{m} (-1)^{k-r} \binom{k-1}{r-1} S_k + (-1)^{m-r+1} \sum_{\ell=m+1}^{n} \sum_{u=m}^{\ell-1} \binom{u}{r-1} \binom{u-r}{m-r} P(E_\ell). \tag{12}$$

(ii) For each $J \subseteq [m+1, n]$*, the probability* $P(F_r)$ *of the event* F_r *can be bounded as given in the following Bonferroni-type inequalities:s:*

$$P(F_r) \begin{cases} \leq \sum_{k=r}^{m} (-1)^{k-r} \binom{k-1}{r-1} S_k - \sum_{\ell\in J} \sum_{u=m}^{\ell-1} \binom{u}{r-1} \binom{u-r}{m-r} P(E_\ell), & \text{if } m-r \text{ is even,} \\ \geq \sum_{k=r}^{m} (-1)^{k-r} \binom{k-1}{r-1} S_k + \sum_{\ell\in J} \sum_{u=m}^{\ell-1} \binom{u}{r-1} \binom{u-r}{m-r} P(E_\ell), & \text{if } m-r \text{ is odd,} \end{cases} \tag{13}$$

where the inequalities in (13) hold with equality if and only if $P(E_\ell) = 0$ *for all* $\ell \in [m+1, n] \backslash J$.

We make the following remark that is similar to Remark 2:

Remark 4 (i) The generalized inclusion-exclusion principle in (3) is a special case of our Bonferroni-type formula in (12).

(ii) Our Bonferroni-type inequalities in (13) improve on the classical Bonferroni inequalities in (6). Furthermore, our inequalities in (13) subsume the inequalities in (6).as a special case, and the inequalities in (6) hold with equality if and only if $P(F_{m+1}) = 0$.

We need the following lemma for the proof of Theorem 3.

Lemma 5 *Suppose* $1 \leq r \leq m \leq \ell$. *Then we have*

$$\sum_{k=r}^{m} (-1)^{k-r} \binom{k-1}{r-1} \binom{\ell}{k} = 1 + (-1)^{m-r} \sum_{u=m}^{\ell-1} \binom{u}{r-1} \binom{u-r}{m-r}. \tag{14}$$

Proof. We prove Lemma 5 by induction on ℓ. First consider the base case that $\ell = 1$. Suppose $1 \le r \le m \le \ell$. Then we have $r = m = \ell = 1$, and hence the LHS of (14) is equal to 1. Thus, Lemma 5 holds for the base case that $\ell = 1$.

Now assume as the induction hypothesis that Lemma 5 holds for some $\ell - 1 \ge 1$. Suppose $1 \le r \le m \le \ell$. We consider the following three cases:

Case 1: $r = m = \ell$. In this case, the LHS of (14) is equal to 1. Thus, Lemma 5 holds in this case.

Case 2: $r < m = \ell$. In this case, we have

$$\begin{aligned}\sum_{k=r}^{m}(-1)^{k-r}\binom{k-1}{r-1}\binom{\ell}{k} &= \sum_{k=r}^{\ell}(-1)^{k-r}\binom{k-1}{r-1}\binom{\ell-1}{k-1} + \sum_{k=r}^{\ell-1}(-1)^{k-r}\binom{k-1}{r-1}\binom{\ell-1}{k} \\ &= \sum_{k'=r-1}^{\ell-1}(-1)^{k'-(r-1)}\binom{k'}{r-1}\binom{\ell-1}{k'} + 1 = 0 + 1 = 1,\end{aligned}$$

where the first equality follows from $m = \ell$, $\binom{\ell}{k} = \binom{\ell-1}{k-1} + \binom{\ell-1}{k}$ for $r \le k \le \ell - 1$, and $\binom{\ell}{\ell} = \binom{\ell-1}{\ell-1}$, the second equality follows from the induction hypothesis, and the third equality follows from (10). Thus, Lemma 5 holds in this case.

Case 3: $m \le \ell - 1$. In this case, we have

$$\begin{aligned}&\sum_{k=r}^{m}(-1)^{k-r}\binom{k-1}{r-1}\binom{\ell}{k} \\ &= \sum_{k=r}^{m}(-1)^{k-r}\binom{k-1}{r-1}\binom{\ell-1}{k-1} + \sum_{k=r}^{m}(-1)^{k-r}\binom{k-1}{r-1}\binom{\ell-1}{k} \\ &= \sum_{k'=r-1}^{m-1}(-1)^{k'-(r-1)}\binom{k'}{r-1}\binom{\ell-1}{k'} + \left(1 + (-1)^{m-r}\sum_{u=m}^{\ell-2}\binom{u}{r-1}\binom{u-r}{m-r}\right) \\ &= (-1)^{(m-1)-(r-1)}\binom{\ell-1}{r-1}\binom{(\ell-1)-(r-1)-1}{(m-1)-(r-1)} + \left(1 + (-1)^{m-r}\sum_{u=m}^{\ell-2}\binom{u}{r-1}\binom{u-r}{m-r}\right) \\ &= 1 + (-1)^{m-r}\sum_{u=m}^{\ell-1}\binom{u}{r-1}\binom{u-r}{m-r},\end{aligned}$$

where the second equality follows from the induction hypothesis and the third equality follows from (10). Thus, Lemma 5 holds in this case. ∎

(Proof of Theorem 3) Let B_J be given as in the proof of Theorem 1 for all $J \subseteq [n]$. Then we have

$$\begin{aligned}&\sum_{k=r}^{m}(-1)^{k-r}\binom{k-1}{r-1}S_k \\ &= \sum_{\ell=r}^{m}\sum_{\substack{J\subseteq[n]\\|J|=\ell}}\left(\sum_{k=r}^{\ell}(-1)^{k-r}\binom{k-1}{r-1}\binom{\ell}{k}\right)P(B_J) + \sum_{\ell=m+1}^{n}\sum_{\substack{J\subseteq[n]\\|J|=\ell}}\left(\sum_{k=r}^{m}(-1)^{k-r}\binom{k-1}{r-1}\binom{\ell}{k}\right)P(B_J) \\ &= \sum_{\ell=r}^{m}\sum_{\substack{J\subseteq[n]\\|J|=\ell}}1\cdot P(B_J) + \sum_{\ell=m+1}^{n}\sum_{\substack{J\subseteq[n]\\|J|=\ell}}\left(1 + (-1)^{m-r}\sum_{u=m}^{\ell-1}\binom{u}{r-1}\binom{u-r}{m-r}\right)P(B_J) \\ &= P(F_r) + (-1)^{m-r}\sum_{\ell=m+1}^{n}\sum_{u=m}^{\ell-1}\binom{u}{r-1}\binom{u-r}{m-r}P(E_\ell),\end{aligned}$$

where the first equality follows from the same arguments leading to (11) and the second equality follows from (14) in Lemma 5 and and $P(F_r) = \Sigma_{\ell=r}^{n}P(E_\ell)$. ∎

References

[1] J. Galambos and I. Simonelli, *Bonferroni-type Inequalities with Applications*, New York, NY: Springer-Verlag, 1996.

[2] C. Jordan, "De quelques formules de probabilité," *Comptes Rendus Academy Science, Paris*, 65 (1867), 993–994.

[3] W. Feller, *An Introduction to Probability Theory and Its Applications*, vol. I, 3rd ed., New York, NY: Wiley & Sons, 1968.

[4] F. S. Roberts and B. Tesman, *Applied Combinatorics*, 2nd ed., Boca Raton, FL: CRC Press, 2009.

[5] D. R. Mazur, *Combinatorics: A Guided Tour*, Washington, DC: The Mathematical Association of America, 2010.

[6] J. H. van Lint and R. M. Wilson, *A Course in Combinatorics*, 2nd ed., Cambridge, UK: Cambridge University Press, 2001.

[7] C. E. Bonferroni, "Teoria statistica delle classi e calcolo delle probabilitá," *Pubblicazioni del R. Istituto Superiore di Scienze Economiche e Commerciali di Firenze*, 8 (1936), 1–62.