ON MUTUALLY FAVORABLE EVENTS

By Kai-Lai Chung

Tsing Hua University, Kunming, China

Introduction. For a set of arbitrary events, E. J. Gumbel, M. Fréchet and the author[1] have recently obtained inequalities between sums of certain probability functions. One of the results of the author is the following:

Let $E_1, \cdots, E_n$ be n arbitrary events and let $p_m(\nu_1, \cdots, \nu_k)$ denote the probability of the occurrence of at least m events out of the k events $E_{\nu_1}, \cdots, E_{\nu_k}$. Then, for $k = 1, \cdots, n - 1$ and $1 \leqq m \leqq k$ we have

$$\binom{n-m}{k-m} \Sigma p_m(\nu_1, \cdots, \nu_{k+1}) \leqq \binom{n-m}{k-m+1} \Sigma p_m(\nu_1, \cdots, \nu_k),$$

where the summations extend respectively to all combinations of $k + 1$ and k indices out of the n indices $1, \cdots, n$.

In course of proof of the above inequalities it appears that similar inequalities between products instead of sums can be obtained under certain assumptions regarding the nature of interdependence of the events. We shall first study the nature of such assumptions, and then proceed to the proof of the said inequalities (Theorems 1 and 2). It may be noted that the inductive method used here serves equally well for the proof of the inequalities cited above, though somewhat longer, but apparently our former method is not applicable here.

That events satisfying our assumptions actually exist, is shown by an application to the elementary theory of numbers. The author feels incompetent to discuss other possible fields of application.

1. Let a set of events be given

$$E_1, E_2, \cdots, E_n, \cdots$$

and let E'_i denote the event non-E_i. Let $p(i)$ denote the probability of the occurrence of E_i, $p(i')$ that of the occurrence of E'_i. For convenience we assume that for any i $p_i(1 - p_i) \neq 0$; events with the exceptional probabilities 0 or 1 may evidently be left out of account.

Let $p(\nu_1 \cdots \nu_k)$ denote the probability of the occurrence of the conjunction $E_{\nu_1} \cdots E_{\nu_k}$ and let $p(\mu_1 \cdots \mu_h, \nu_1 \cdots \nu_k)$ denote the probability of the occurrence of $E_{\nu_1} \cdots E_{\nu_k}$, on the hypothesis that $E_{\mu_1} \cdots E_{\mu_h}$ have occurred. The μ's or ν's may be accented.

Definition 1: *If $p(\nu_1, \nu_2) > p(\nu_2)$, we say that the occurrence of the event E_{ν_1} is favorable to the occurrence of the event E_{ν_2}, or simply that E_{ν_1} is favorable to E_{ν_2}.*

---

[1] "On the probability of the occurrence of at least m events among n arbitrary events," *Annals of Math. Stat.* Vol. 12 (1941), pp. 328–338.

If $p(\nu_1, \nu_2) = p(\nu_2)$, *we say that* E_{ν_1} *is indifferent to* E_{ν_2}. *If* $p(\nu_1, \nu_2) < p(\nu_2)$, *we say that* E_{ν_1} *is unfavorable to* E_{ν_2}.

Thus the relations "favorableness," "indifference," and "unfavorableness" are mutually exclusive and together exhaustive. We state the following immediate consequences:

(i) Reflexity: An event is favorable to itself; in fact, $p(\nu, \nu) = 1 > p(\nu)$.

(ii) Symmetry: If E_1 is favorable (indifferent, unfavorable) to E_2, then E_2 is favorable (indifferent, unfavorable) to E_1. In fact, we have

$$p(1)p(1, 2) = p(12) = p(2)p(2, 1),$$

$$\frac{p(1, 2)}{p(2)} = \frac{p(2, 1)}{p(1)}.$$

Thus $p(1, 2) \gtreqless p(2)$ is equivalent to $p(2, 1) \gtreqless p(1)$.

In particular, if E_1 is indifferent to E_2, then so is E_2 to E_1. They are then usually said to be independent of each other.

(iii) If E_1 is favorable (indifferent, unfavorable) to E_2, then E_1' is unfavorable (indifferent, favorable) to E_2. For, we have

$$p(1)p(1, 2) + p(1')p(1', 2) = p(12) + p(1'2) = p(2),$$

whence

$$p(1')p(1', 2) = p(2) - p(1)p(1, 2).$$

On the other hand,

$$p(1')p(2) = [1 - p(1)]p(2) = p(2) - p(1)p(2).$$

Since by assumption $p(1')p(2) \neq 0$, we have

$$\frac{p(1', 2)}{p(2)} = \frac{p(2) - p(1)p(1, 2)}{p(2) - p(1)p(2)}.$$

Thus

$$p(1', 2) \gtreqless p(2) \text{ according as } p(1, 2) \gtreqless p(2).$$

For the sake of brevity we introduce the following symbolic notation:

$$E_1/E_2 = \begin{cases} 1, & \text{if } E_1 \text{ is favorable to } E_2 \\ 0, & \text{if } E_1 \text{ is indifferent to } E_2 \\ -1, & \text{if } E_1 \text{ is unfavorable to } E_2. \end{cases}$$

Then by (ii) and (iii) we have

$$E_1/E_2 = E_2/E_1,$$

$$E_1'/E_2 = E_2/E_1' = E_1/E_2' = E_2'/E_1 = -(E_1/E_2),$$

$$E_1'/E_2' = E_2'/E_1' = E_1/E_2,$$

analogous to the rules of signs in the multiplication of integers.

(iv) Non-transitivity: If E_1 is favorable to E_2, and E_2 is favorable to E_3, it does not necessarily follow that E_1 is favorable to E_3; in fact, it may happen that E_1 is unfavorable to E_3. For instance, imagine 11 identical balls in a bag marked respectively with the numbers

$$-11,\ -10,\ -3,\ -2,\ -1,\ 2,\ 4,\ 6,\ 11,\ 13,\ 16.$$

Let a ball be drawn at random. Let

$$E_1 = (\text{the event of the number on the ball being positive})$$

$$E_2 = (\text{the event of the number on the ball being even})$$

$$E_3 = (\text{the event of the number on the ball being of 1 digit})$$

We have

$$p(1, 2) = \tfrac{4}{6} > \tfrac{6}{11} = p(2),$$

$$p(2, 3) = \tfrac{4}{6} > \tfrac{6}{11} = p(3),$$

$$p(1, 3) = \tfrac{1}{2} < \tfrac{6}{11} = p(3).$$

(v) It may happen that $E_1/E_3 = 1$, $E_2/E_3 = 1$, but $E_1E_2/E_3 = -1$. In the example above,

$$p(2, 1) = \tfrac{4}{6} > \tfrac{6}{11} = p(1),$$

$$p(3', 1) = \tfrac{3}{5} > \tfrac{6}{11} = p(1),$$

$$p(23', 1) = \tfrac{1}{2} < \tfrac{6}{11} = p(1).$$

(vi) It may happen that $E_1/E_2 = 1$, $E_1/E_3 = 1$, but $E_1/E_2E_3 = -1$. Example:

$$p(1, 2) = \tfrac{4}{6} > \tfrac{6}{11} = p(2),$$

$$p(1, 3') = \tfrac{1}{2} > \tfrac{5}{11} = p(3'),$$

$$p(1, 23') = \tfrac{1}{6} < \tfrac{2}{11} = p(23').$$

(vii) It may happen that $E_1/E_3 = 1$, $E_2/E_3 = 1$, but the disjunction $(E_1 + E_2)/E_3 = -1$. For, by (v) we know that there exist events E_1', E_2', E_3' such that

$$E_1'/E_3' = 1, \qquad E_2'/E_3' = 1, \qquad E_1'E_2'/E_3' = -1.$$

Hence by (iii) there exist events E_1, E_2, E_3 such that

$$E_1/E_3 = 1, \qquad E_2/E_3 = 1, \qquad (E_1'E_2')'/E_3 = -1.$$

But $(E_1'E_2')' = E_1 + E_2$. Thus the last relation is $(E_1 + E_2)/E_3 = -1$.

(viii) It may happen that $E_1/E_2 = 1$, $E_1/E_3 = 1$, but $E_1/(E_2 + E_3) = -1$. This follows from (vi) as (vii) follows from (v).

After all these negative results in (iv)–(viii), we see that we cannot expect to go far without making stronger assumptions regarding the nature of interdependence between the events in the set. Firstly, in view of (iv), we shall restrict ourselves to consideration of a set of events in which each event is favorable to every other. Secondly, in view of (v), we shall only consider the case where the "favorableness," as defined above, shall be cumulative in its effect, that is to say, the more events favorable to a given event have been known to occur, the more probable this given event shall be esteemed. We formulate these two conditions in mathematical terms, as follows:

Definition 2: *A set of events* $E_1, \cdots, E_n, \cdots$ *is said to be strongly mutually favorable (in the first sense) if, for every integer* h *and every set of distinct indices (positive integers)* $\mu_1, \cdots, \mu_h$ *and* ν *we have*

$$p(\mu_1 \cdots \mu_h, \nu) > p(\mu_1 \cdots \mu_{h-1}, \nu).$$

This definition requires that there exist no implication relation between any event and any conjunction of events in the set; in particular, that the events are all distinct. It would be more convenient to consider the relation "favorable or indifferent to." This will be done later on. The present definitions have the advantage of being logically clear cut and also that of yielding unambiguous inequalities.

From Definition 2 we deduce the following consequences:

(1) If the set $(\mu_1^*, \cdots, \mu_j^*)$ is a sub-set of $(\mu_1, \cdots, \mu_h)$, we have

$$p(\mu_1 \cdots \mu_h, \nu) > p(\mu_1^* \cdots \mu_j^*, \nu).$$

(2) For any positive integer k and any two sets $(\nu_1, \cdots, \nu_k)$ and $(\mu_1, \cdots, \mu_h)$ where all the indices are distinct, we have

$$p(\mu_1 \cdots \mu_h, \nu_1 \cdots \nu_k) > p(\mu_1 \cdots \mu_{h-1}, \nu_1 \cdots \nu_k).$$

More generally, we have as in (1),

$$p(\mu_1 \cdots \mu_h, \nu_1 \cdots \nu_k) > p(\mu_1^* \cdots \mu_j^*, \nu_1 \cdots \nu_k).$$

Proof: We have only to prove the first inequality. For $k = 1$ this is the assumption in Definition 2. Suppose that the inequality holds for $k - 1$, we shall prove that it holds for k, too:

$$\begin{aligned}
\frac{p(\mu_1 \cdots \mu_h, \nu_1 \cdots \nu_k)}{p(\mu_1 \cdots \mu_{h-1}, \nu_1 \cdots \nu_k)} &= \frac{p(\mu_1 \cdots \mu_{h-1})p(\mu_1 \cdots \mu_h)p(\mu_1 \cdots \mu_h, \nu_1 \cdots \nu_k)}{p(\mu_1 \cdots \mu_h)p(\mu_1 \cdots \mu_{h-1})p(\mu_1 \cdots \mu_{h-1}, \nu_1 \cdots \nu_k)} \\
&= \frac{p(\mu_1 \cdots \mu_{h-1})p(\mu_1 \cdots \mu_h \nu_1 \cdots \nu_k)}{p(\mu_1 \cdots \mu_h)p(\mu_1 \cdots \mu_{h-1} \nu_1 \cdots \nu_k)} \\
&= \frac{p(\mu_1 \cdots \mu_{h-1})p(\mu_1 \cdots \mu_h)p(\mu_1 \cdots \mu_h, \nu_1)p(\mu_1 \cdots \mu_h \nu_1, \nu_2 \cdots \nu_k)}{p(\mu_1 \cdots \mu_h)p(\mu_1 \cdots \mu_{h-1})p(\mu_1 \cdots \mu_{h-1}, \nu_1)p(\mu_1 \cdots \mu_{h-1} \nu_1, \nu_2 \cdots \nu_k)} \\
&= \frac{p(\mu_1 \cdots \mu_h, \nu_1)}{p(\mu_1 \cdots \mu_{h-1}, \nu_1)} \frac{p(\mu_1 \cdots \mu_h \nu_1, \nu_2 \cdots \nu_k)}{p(\mu_1 \cdots \mu_{h-1} \nu_1, \nu_2 \cdots \nu_k)} \\
&> \frac{p(\mu_1 \cdots \mu_h \nu_1, \nu_2 \cdots \nu_k)}{p(\mu_1 \cdots \mu_{h-1} \nu_1, \nu_2 \cdots \nu_k)} > 1.
\end{aligned}$$

Observe that none of the denominators vanish by our original assumption and by Definition 2.

Therefore we see that when the failure in (v) is remedied by our definition, the failure in (vi) is automatically remedied too.

2. Theorem 1: *Let $n > 1$ and let $E_1, \cdots, E_n, \cdots$ be a set of strongly mutually favorable events (in the first sense). Then we have, for $k = 1, \cdots, n-1$,*

$$\prod_{\nu_1,\cdots,\nu_{k+1}} [p(\nu_1 \cdots \nu_{k+1})]^{\binom{n-1}{k}^{-1}} > \prod_{\nu_1,\cdots,\nu_k} [p(\nu_1 \cdots \nu_k)]^{\binom{n-1}{k-1}^{-1}}$$

where the products extend respectively to all combinations of $k+1$ and k distinct indices out of the indices $1, \cdots, n$.

Proof. We may assume that the indices are written so that

$$1 \leqq \nu_1 < \nu_2 < \cdots < \nu_{k+1} \leqq n.$$

Taking logarithms, we have

$$\binom{n-1}{k-1} \sum_{\nu_1,\cdots,\nu_{k+1}} \log p(\nu_1 \cdots \nu_{k+1}) > \binom{n-1}{k} \sum_{\nu_1,\cdots,\nu_k} \log p(\nu_1 \cdots \nu_k).$$

Substituting from the obvious formula

$$p(\nu_1 \cdots \nu_k) = p(\nu_1)p(\nu_1, \nu_2)p(\nu_1\nu_2, \nu_3) \cdots p(\nu_1 \cdots \nu_{k-1}, \nu_k),$$

and writing $\log p(\cdots) = q(\cdots)$, the inequality becomes

$$\begin{aligned} &\binom{n-1}{k-1} \Sigma[q(\nu_1) + q(\nu_1, \nu_2) + \cdots + q(\nu_1 \cdots \nu_k, \nu_{k+1})] \\ (1) \qquad &\qquad > \binom{n-1}{k} \Sigma[q(\nu_1) + q(\nu_1, \nu_2) + \cdots + q(\nu_1 \cdots \nu_{k-1}, \nu_k)]. \end{aligned}$$

Immediately we observe that the number of terms of the form $q(\nu_1 \cdots \nu_s, \mu)(0 \leqq s \leqq \mu - 1)$ with a fixed μ after the comma in the bracket is the same on both sides, since

$$(2) \qquad \binom{n-1}{k-1}\binom{n-1}{k} = \binom{n-1}{k}\binom{n-1}{k-1}.$$

Let the sums of such q's on the left and right of (1) be $\sigma^{(1)} = \sigma^{(1)}(\mu)$ and $\sigma^{(2)} = \sigma^{(2)}(\mu)$ respectively. To prove our theorem it is sufficient to prove that $\sigma^{(1)}(\mu) \geqq \sigma^{(2)}(\mu)$ for every μ and $\sigma^{(1)}(\mu) > \sigma^{(2)}(\mu)$ for at least one μ.

Now the terms in $\sigma^{(1)}$ (or $\sigma^{(2)}$) fall into classes according to the number s of the μ_i's before the comma in the bracket. Let those terms having s μ_i's before the comma belong to the s-th class. It is evident that the number of terms of the s-th class in $\sigma^{(1)}(\mu)$ is equal to

$$\binom{n-1}{k-1}\binom{\mu-1}{s}\binom{n-\mu}{k-s}$$

for $s = 0, 1, \cdots, \mu - 1$; where we make the convention that

$$\binom{0}{0} = 1, \qquad \binom{a}{b} = 0 \quad \text{if } a < b \text{ or if } b < 0.$$

Thus for a fixed μ, when the terms in $\sigma^{(1)}(\mu)$ are classified in the above manner, its total number of terms may be written as the following sum, in which vanishing terms may occur:

$$\begin{aligned}\binom{n-1}{k-1}\binom{n-1}{k} = \binom{n-1}{k-1}\biggl\{&\binom{\mu-1}{\mu-1}\binom{n-\mu}{k-\mu+1} \\ &+ \binom{\mu-1}{\mu-2}\binom{n-\mu}{k-\mu+2} + \cdots + \binom{\mu-1}{s}\binom{n-\mu}{k-s} \\ &+ \cdots + \binom{\mu-1}{0}\binom{n-\mu}{k}\biggr\}.\end{aligned}$$

Similarly the total number of terms in $\sigma^{(2)}(\mu)$ may be written as the following sum:

$$\begin{aligned}\binom{n-1}{k}\binom{n-1}{k-1} = \binom{n-1}{k}\biggl\{&\binom{\mu-1}{\mu-1}\binom{n-\mu}{k-\mu} + \binom{\mu-1}{\mu-2}\binom{n-\mu}{k-\mu+1} \\ &+ \cdots + \binom{\mu-1}{s}\binom{n-\mu}{k-s-1} + \cdots + \binom{\mu-1}{0}\binom{n-\mu}{k-1}\biggr\}.\end{aligned}$$

LEMMA 1: *For $0 \leqq s \leqq k$, we have, taking account of our conventions about the binomial coefficients,*

$$(3) \quad \binom{n-1}{k-1}\binom{n-\mu}{k-s} > \binom{n-1}{k}\binom{n-\mu}{k-s-1} \quad \textit{for} \quad s > (\mu - 1)k/n;$$

$$(4) \quad \binom{n-1}{k-1}\binom{n-\mu}{k-s} \leqq \binom{n-1}{k}\binom{n-\mu}{k-s-1} \quad \textit{for} \quad s \leqq (\mu - 1)k/n.$$

PROOF: Suppose $s \geqq k - n + \mu$, then

$$\binom{n-1}{k-1}\binom{n-\mu}{k-s} \gtreqless \binom{n-1}{k}\binom{n-\mu}{k-s-1}$$

according as

$$\frac{k}{n-k} \gtreqless \frac{k-s}{n-\mu-k+s+1},$$

i.e. according as

$$s \gtreqless (\mu - 1)k/n.$$

But, since $k < n$ and $\mu \leqq n$, we have

$$n - k - k/n + 1 > (n - k)\mu/n$$

$$(\mu - 1)k/n > k - n + \mu - 1$$

so that

$$(\mu - 1)k/n + 1 \geqq k - n + \mu.$$

Therefore if $s > (\mu - 1)k/n$, then $s \geqq (\mu - 1)k/n + 1 \geqq k - n + \mu$, and (3) holds.

Again, if $k - n + \mu \leqq s \leqq (\mu - 1)k/n$, then (4) holds; while if $s < k - n + \mu$, then the left-hand side of (4) vanishes while the right-hand side is non-negative, thus (4) holds for $s \leqq (\mu - 1)k/n$. The lemma is proved.

If we put $(s = 0, 1, \cdots, k)$

$$\binom{n-1}{k-1}\binom{n-\mu}{k-s} - \binom{n-1}{k}\binom{n-\mu}{k-s-1} = d_s,$$

then by Lemma 1,

$$d_s \gtreqqless 0 \quad \text{according as} \quad s \gtreqqless (\mu - 1)k/n.$$

This means that although the total number of terms of the form $p(\mu_1 \cdots \mu_s, \mu)$ is the same on both sides of (1), the left-hand side is more abundant in terms with larger s while the right-hand side is more abundant in terms with smaller s. Now we have

$$q(\mu_1 \cdots \mu_i, \mu) > q(\mu_1^* \cdots \mu_j^*, \mu)$$

if $i > j$ and if $(\mu_1^* \cdots \mu_j^*)$ is a subset of $(\mu_1 \cdots \mu_i)$. Hence it is natural to suppose that the left-hand side must be greater because it is more abundant in terms of larger values. Unfortunately even if $i > j$, the last inequality is in general not true if the set $(\mu_1^* \cdots \mu_j^*)$ is not a sub-set of $(\mu_1 \cdots \mu_i)$. Therefore we cannot as yet conclude that $\sigma^{(1)} \geqq \sigma^{(2)}$.

To prove that actually we have $\sigma^{(1)} \geqq \sigma^{(2)}$, we make the following "process of compensation":

We have, by (2) and the definition of d_s, the following equality:

$$\binom{\mu-1}{0}d_0 + \binom{\mu-1}{1}d_1 + \cdots + \binom{\mu-1}{\mu-1}d_{\mu-1} = 0.$$

where $d_j = 0$ if $j > k$. Thus

$$d_s \leqq 0 \quad \text{for} \quad s \leqq k(\mu - 1)/n,$$

$$d_s \geqq 0 \quad \text{for} \quad s > k(\mu - 1)/n.$$

Hence

$$\binom{\mu-1}{0}d_0 + \binom{\mu-1}{1}d_1 + \cdots + \binom{\mu-1}{\mu-1}d_{\mu-1} \leqq 0 \tag{5}$$

$$\text{for} \quad l = 0, 1, \cdots, \mu - 1.$$

For the fixed μ, let

$$\rho_l^{(1)} = \binom{n-1}{k-1}\left\{\binom{n-\mu}{k}q_\mu + \binom{n-\mu}{k-1}\sum_{\mu_1<\mu} q(\mu_1, \mu) + \cdots \right.$$
$$\left. + \binom{n-\mu}{k-l}\sum_{\mu_1<\cdots<\mu_l<\mu} q(\mu_1 \cdots \mu_l, \mu)\right\}$$

$$\rho_l^{(2)} = \binom{n-1}{k}\left\{\binom{n-\mu}{k-1}q_\mu + \binom{n-\mu}{k-2}\sum_{\mu_1<\mu} q(\mu_1, \mu) + \cdots \right.$$
$$\left. + \binom{n-\mu}{k-l-1}\sum_{\mu_1<\cdots<\mu_l<\mu} q(\mu_1 \cdots \mu_l, \mu)\right\}$$

so that

$$\rho_{\mu-1}^{(1)} = \sigma^{(1)}(\mu), \qquad \rho_{\mu-1}^{(2)} = \sigma^{(2)}(\mu).$$

For $\mu = 1$, $l = 0$, we have

$$\sigma^{(1)}(1) = \rho_0^{(1)} = \rho_0^{(2)} = \sigma^{(2)}(1).$$

LEMMA 2: *For $\mu > 1$ and $0 \leqq l < \mu - 1$, we have*

$$\sum_{1\leqq\mu_1<\cdots<\mu_l<\mu} q(\mu_1 \cdots \mu_l, \mu) < \frac{l+1}{\mu - l - 1}\sum_{1\leqq\mu_1<\cdots<\mu_{l+1}<\mu} q(\mu_1 \cdots \mu_{l+1}, \mu).$$

PROOF: We have, for any $\nu < \mu$, $\nu \neq \mu_i$ $(i = 1, \cdots, l)$

$$q(\mu_1 \cdots \mu_l\nu, \mu) > q(\mu_1 \cdots \mu_l, \mu).$$

Summing with respect to all such ν's,

$$\sum_\nu q(\mu_1 \cdots \mu_l\nu, \mu) > (\mu - l - 1)q(\mu_1 \cdots \mu_l, \nu).$$

Summing with respect to all $1 \leqq \mu_1 < \cdots < \mu_l < \mu$,

$$\sum_{1\leqq\mu_1<\cdots<\mu_l<\mu}\sum_\nu q(\mu_1 \cdots \mu_l\nu, \mu) = (l+1)\sum_{1\leqq\mu_1<\cdots<\mu_{l+1}<\mu} q(\mu_1 \cdots \mu_{l+1}, \mu)$$
$$> (\mu - l - 1)\sum_{1\leqq\mu_1<\cdots<\mu_l<\mu} q(\mu_1 \cdots \mu_l, \mu).$$

The lemma is proved.

Now we use induction to prove that for $\mu > 1$ and $l = 1, \cdots, \mu - 1$

$$\rho_l^{(1)} - \rho_l^{(2)} > \frac{d_0 + \binom{\mu-1}{1}d_1 + \cdots + \binom{\mu-1}{l}d_l}{\binom{\mu-1}{l}}$$
$$\times \sum_{1\leqq\mu_1<\cdots<\mu_l<\mu} q(\mu_1 \cdots \mu_l, \mu) \geqq 0.$$

This inequality holds for $l = 1$ by Lemma 2. Assume that it holds for l, $(l < \mu - 1)$. Then we have, by (5) and the fact that each $q < 0$,

$$\begin{aligned}
\rho_{l+1}^{(1)} - \rho_{l+1}^{(2)} &= \rho_l^{(1)} - \rho_l^{(2)} + d_{l+1} \sum_{1 \leqq \mu_1 < \cdots < \mu_{l+1} < \mu} q(\mu_1 \cdots \mu_{l+1}, \mu) \\
&> \frac{d_0 + \binom{\mu-1}{1} d_1 + \cdots + \binom{\mu-1}{l} d_l}{\binom{\mu-1}{l}} \sum q(\mu_1 \cdots \mu_l, \mu) \\
&\qquad\qquad + d_{l+1} \sum q(\mu_1 \cdots \mu_{l+1}, \mu) \\
&\geqq \left(\frac{d_0 + \binom{\mu-1}{1} d_1 + \cdots + \binom{\mu-1}{l} d_l}{\binom{\mu-1}{l}} \, \frac{l+1}{\mu - l - 1} + d_{l+1} \right) \sum q(\mu_1 \cdots \mu_{l+1}, \mu) \\
&= \frac{d_0 + \binom{\mu-1}{1} d_1 + \cdots + \binom{\mu-1}{l} d_l + \binom{\mu-1}{l+1} d_{l+1}}{\binom{\mu-1}{l+1}} \sum q(\mu_1 \cdots \mu_{l+1}, \mu) \geqq 0.
\end{aligned}$$

Therefore, for $\mu > 1$, we have

$$\sigma^{(1)}(\mu) - \sigma^{(2)}(\mu) = \rho_{\mu-1}^{(1)} - \rho_{\mu-1}^{(2)} > 0.$$

Since $n > 1$ and $1 \leqq \mu \leqq n$, there exists a $\mu > 1$. Hence

$$\sum_{\mu=1}^{n} \sigma^{(1)}(\mu) > \sum_{\mu=1}^{n} \sigma^{(2)}(\mu)$$

which is equivalent to the inequality (1).

3. Our next step will be to obtain a generalization of Theorem 1. Consider a derived event defined by a disjunction of a (finite) number of events in the set, as follows:

$$E_{\nu_1} + E_{\nu_2} + \cdots + E_{\nu_m}.$$

We call such a disjunction a disjunction of the m-th order.

Definition 3: *A set of events is said to be strongly mutually favorable in the second sense if for every positive integer m, the derived set of events consisting of all the disjunctions of the m-th order forms a strongly mutually favorable set of events* (*in the first sense*).

Let $D = D(m)$ denote in general a disjunction of the m-th order; let $p(D_1 \cdots D_h, D)$ denote the probability of the occurrence of the disjunction D, on the hypothesis that the conjunction of the h disjunctions $D_1 \cdots D_h$ has occurred. Then Definition 3 says that for any positive integer h and any set of distinct D's we have

$$p(D_1 \cdots D_h, D) > p(D_1 \cdots D_{h-1}, D).$$

Since a disjunction of the 1st order is an event E, we see that Definition 3 includes Definition 2.

Let $D_m(\nu_1, \cdots, \nu_k)$, $\nu_1 < \cdots < \nu_k$ denote the derived event

$$\prod_{\mu_1, \cdots, \mu_m} (E_{\mu_1} + \cdots + E_{\mu_m})$$

where the product (conjunction) extends to all combinations of m indices out of the indices $\nu_1, \cdots, \nu_k$. Let $p_m^*(\nu_1, \cdots, \nu_k)$ denote the probability of the occurrence of $D_m(\nu_1, \cdots, \nu_k)$. It is seen that $p_1^*(\nu_1, \cdots, \nu_k) = p(\nu_1 \cdots \nu_k)$ in our previous notation.

We merely state Theorem 2, whose proof is analogous to that of Theorem 1 but requires more cumbersome expressions.

THEOREM 2: *Let $n > k \geqq m \geqq 1$ and let $E_1, \cdots, E_n$ be a set of mutually strongly favorable events in the second sense. Then we have*

$$\prod_{1 \leqq \nu_1 < \cdots < \nu_{k+1} \leqq n} [p_m^*(\nu_1, \cdots, \nu_{k+1})]^{\binom{n-m}{k-m+1}^{-1}} > \prod_{1 \leqq \nu_1 < \cdots < \nu_k \leqq n} [p_m^*(\nu_1, \cdots, \nu_k)]^{\binom{n-m}{k-m}^{-1}}.$$

To give an interpretation of $p_m^*(\nu_1, \cdots, \nu_k)$, we prove the symbolic equation between events:

$$D_m = \prod_{\nu_1 \leqq \mu_1 < \cdots < \mu_m \leqq \nu_k} (E_{\mu_1} + \cdots + E_{\mu_m}) = \sum_{\nu_1 \leqq \mu_1 < \cdots < \mu_{k-m+1} \leqq \nu_k} (E_{\mu_1} \cdots E_{\mu_{k-m+1}}) = C_{k-m+1},$$

where product means conjunction and sum means disjunction.

To prove this, we write for a general event E, $E = 1$ when E occurs, $E = 0$ when E does not occur. Now if $C_{k-m+1} = 0$, then at most $k - m$ events among the k given events occur, so that there exist m events such that $E_{\lambda_1} = 0$, $E_{\lambda_2} = 0$, $E_{\lambda_m} = 0$, thus

$$E_{\lambda_1} + E_{\lambda_2} + \cdots + E_{\lambda_m} = 0.$$

Now the last disjunction is contained in D_m as a factor, therefore $D_m = 0$.

Conversely, if $D_m = 0$, at least one of its factors $= 0$, so that there exist m events, such that $E_{\lambda_1} = 0$, $E_{\lambda_2} = 0, \cdots, E_{\lambda_m} = 0$. Thus at most $k - m$ events out of the k given events occur and so by definition $C_{k-m+1} = 0$. Q.e.d.

From the above it immediately follows that

$$p_m^*(\nu_1, \cdots, \nu_k) = p_{k-m+1}(\nu_1, \cdots, \nu_k)$$

where $p_{k-m+1}(\nu_1, \cdots, \nu_k)$ is defined in the Introduction. Then Theorem 2 may be written as

$$\Pi[p_{k-m+2}(\nu_1, \cdots, \nu_{k+1})]^{\binom{n-m}{k-m+1}^{-1}} > \Pi[p_{k-m+1}(\nu_1, \cdots, \nu_k)]^{\binom{n-m}{k-m}^{-1}}$$

or again as

$$\Pi[w_{m-1}(\nu_1', \cdots, \nu_{k+1}')]^{\binom{n-m}{k-m+1}^{-1}} > \Pi[w_{m-1}(\nu_1', \cdots, \nu_k')]^{\binom{n-m}{k-m}^{-1}}$$

where $w_{m-1}(\nu_1', \cdots, \nu_k')$ denotes the probability of the occurrence of at most $m - 1$ events out of the k events $E_{\nu_1}', \cdots, E_{\nu_k}'$.

REMARK. If in our Definitions 2 and 3 we replace the sign "$>$" by the sign "$\geqq$", then we obtain the inequalities in Theorems 1 and 2 with the sign "$>$" replaced by "$\geqq$". The corresponding set of events thus newly defined will be said to be strongly mutually favorable or indifferent (in the first or second sense).

After this modification, we can include events with the probability 1 in our considerations. Also, the events need no longer be distinct and there may now exist implication relations between events or their conjunctions. This modification is useful for the following application.

4. Consider the divisibility of a random positive integer by the set of positive integers. To each positive integer there corresponds an event, namely the event that the random positive integer is divisible by it. The enumerable set of events

$$E_1, E_2, E_3, E_4, \cdots, E_n, \cdots$$

where E_n = the event of divisibility by n, with the probabilities

$$1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \cdots, \frac{1}{n}, \cdots$$

evidently forms a set of strongly mutually favorable or indifferent events in the second sense.

Again, the enumerable set of events

$$E_2', E_3', E_4', \cdots, E_n', \cdots$$

where E_n' = the event of non-divisibility by n, with the probabilities

$$\frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \cdots, \frac{n-1}{n}, \cdots$$

evidently also forms a set of strongly mutually favorable or indifferent events in the second sense.

Hence our Theorem 2 can be applied to both sets and in this way we obtain results which belong properly to the elementary theory of numbers.

We shall content ourselves with indicating a few examples.

Let $\{a_1, \cdots, a_n\}$ denote the least common multiple of the natural numbers $a_1, \cdots, a_n$. Then Theorem 1, when applied to the two sets above, gives respectively

THEOREM 1.1: *Let $a_1, \cdots, a_n$ be any positive integers, then we have, for $k = 1, \cdots, n - 1$*

$$\left(\prod_{1 \leqq \nu_1 < \cdots < \nu_{k+1} \leqq n} \frac{1}{\{a_{\nu_1}, \cdots, a_{\nu_{k+1}}\}}\right)^{\binom{n-1}{k}^{-1}} \geqq \left(\prod_{1 \leqq \nu_1 < \cdots < \nu_k \leqq n} \frac{1}{\{a_{\nu_1}, \cdots, a_{\nu_k}\}}\right)^{\binom{n-1}{k-1}^{-1}}.$$

THEOREM 1.2: *Also we have,*

$$\prod_{1\leqq\nu_1<\cdots<\nu_{k+1}\leqq n}\left(1-\sum_{\nu_1\leqq\mu_1\leqq\nu_{k+1}}\frac{1}{a_{\mu_1}}+\sum_{\nu_1\leqq\mu_1<\mu_2\leqq\nu_{k+1}}\frac{1}{\{a_{\mu_1},a_{\mu_2}\}}-+\cdots+(-1)^{k+1}\frac{1}{\{a_{\nu_1},\cdots,a_{\nu_{k+1}}\}}\right)^{\binom{n-1}{k}^{-1}}$$

$$\geqq\prod_{1\leqq\nu_1<\cdots<\nu_k\leqq n}\left(1-\sum_{\nu_1\leqq\mu_1\leqq\nu_k}\frac{1}{a_{\mu_1}}+\sum_{\nu_1\leqq\mu_1<\mu_2\leqq\nu_k}\frac{1}{\{a_{\mu_1},a_{\mu_2}\}}-+\cdots+(-1)^{k}\frac{1}{\{a_{\nu_1},\cdots,a_{\nu_k}\}}\right)^{\binom{n-1}{k-1}^{-1}}$$

A trivial corollary of Theorem 1 is

$$p(12\cdots n)\geqq p_1p_2\cdots p_n.$$

Correspondingly we have

$$1-\sum_{1\leqq\mu_1\leqq n}\frac{1}{a_{\mu_1}}+\sum_{1\leqq\mu_1<\mu_2\leqq n}\frac{1}{\{a_{\mu_1},a_{\mu_2}\}}-+\cdots+(-1)^n\frac{1}{\{a_1,\cdots,a_n\}}$$

$$\geqq\left(1-\frac{1}{a_1}\right)\left(1-\frac{1}{a_2}\right)\cdots\left(1-\frac{1}{a_n}\right).$$

If we multiply by $a_1a_2\cdots a_n$, we get

$$A(a_1,a_2,\cdots,a_n)\geqq(a_1-1)(a_2-1)\cdots(a_n-1),$$

where $A(a_1,\cdots,a_n)$ denotes the number of positive integers $\leqq a_1a_2\cdots a_n$ that are not divisible by any of the a_i $(i=1,\cdots,n)$.

This last result, which is almost obvious here, was proved by H. Rohrbach and H. Heilbronn independently.[2] See also my generalization[3] (also obvious from the present point of view) of this result to higher dimensional sets of positive integers and to sets of ideals in any algebraic number field.

---

[2] "Beweis einer zahlentheoretische Ungleichung," *Jour. für Math.*, Vol. 177 (1937), pp. 193–196. "On an inequality in the elementary theory of numbers," *Proc. Camb. Phil. Soc.*, Vol. 33, (1937), pp. 207–209.

[3] "A generalization of an inequality in the elementary theory of numbers," *Jour. für Math.*, Vol. 183 (1941), p. 103.