

ON QUADRATIC INEQUALITIES IN PROBABILITY THEORY

by

J. GALAMBOS and A. RÉNYI

Summary

In this paper quadratic inequalities in the probabilities of Boolean functions of n variable events are considered. For a special class of such inequalities — called exact inequalities — a necessary and sufficient condition is given; this general theorem is applied to deduce certain special inequalities. Generalization to inequalities of degree higher than 2 is also considered.

§ 0. Notations

Let $S = (\Omega, \mathcal{A}, P)$ denote a probability space, i.e. let Ω be an arbitrary non-empty set, \mathcal{A} a σ -algebra¹ of subsets of Ω and P a measure on \mathcal{A} such that $P(\Omega) = 1$. We call the elements of \mathcal{A} events and denote them by capital letters. We denote by $A + B$ the union and by AB the intersection of the sets A and B , and by \bar{A} the complement of the set A with respect to Ω . As usual, \bar{A} is interpreted as the event consisting in the non-occurrence of the event A , while $A + B$ and AB respectively, are interpreted as the event that at least one of the events A, B occurs, resp. that both the events A, B occur.

Let p_1, p_2, \dots, p_r be any set of positive numbers such that

$$\sum_{j=1}^r p_j = 1$$

We shall denote by $S_r(p_1, \dots, p_r)$ that (finite) probability space in which the set Ω consists of r elements $\omega_1, \omega_2, \dots, \omega_r$. \mathcal{A} is the set of all 2^r subsets of Ω , and P is defined by

$$(0.1) \quad P(A) = \sum_{\omega_j \in A} p_j$$

Especially $S_1(1)$ is the trivial probability space which contains only two events: the "certain event" Ω and the "impossible event" \emptyset (the empty set). Further $S_2(\frac{1}{2}, \frac{1}{2})$ is the probability space (describing e.g. the throw of a fair coin) which contains only four events: $\Omega, \emptyset, \alpha = \{\omega_1\}$ and $\beta = \{\omega_2\}$ and $P(\alpha) = P(\beta) = \frac{1}{2}$.

A Boolean function $F = F(A_1, A_2, \dots, A_n)$ of n variable events A_1, \dots, A_n is a function of these events which can be expressed by means of the variables

¹ All results of this paper are valid also if \mathcal{A} is only an algebra of subsets of Ω and P a finitely additive nonnegative set function on \mathcal{A} for which $P(\Omega) = 1$.

A_1, \dots, A_n and a finite number of Boolean operations, i.e. the operations $A + B$, AB , \bar{A} . We introduce the notation

$$A^1 = A, \quad A^{-1} = \bar{A}.$$

Let us denote by $\delta_k(m)$ the k -th digit of the binary representation of the non-negative integer m , i.e. we put

$$(0.2) \quad m = \sum_{k \geq 0} \delta_k(m) 2^k$$

Let us put further

$$(0.3) \quad \varepsilon_k(m) = 2\delta_{k-1}(m) - 1 \quad (k=1, 2, \dots).$$

Clearly $\varepsilon_k(m) = \pm 1$, and if m runs over the integers $0, 1, \dots, 2^n - 1$, the n -tuple $\{\varepsilon_1(m), \dots, \varepsilon_n(m)\}$ runs over all 2^n possible n -tuples of the signs $+1$ and -1 .

Let us put

$$(0.4) \quad B_n(m) = A_1^{\varepsilon_1(m)} A_2^{\varepsilon_2(m)} \dots A_n^{\varepsilon_n(m)} \quad (0 \leq m \leq 2^n - 1)$$

We call the $B_n(m)$ the basic Boolean functions of the variables A_1, \dots, A_n . Clearly

$$(0.5) \quad B_n(m_1) B_n(m_2) = \emptyset \quad \text{if } m_1 \neq m_2$$

and

$$(0.6) \quad \sum_{m=0}^{2^n-1} B_n(m) = \Omega$$

It is well known that every Boolean function $F(A_1, \dots, A_n)$ can be uniquely represented in a „canonical form” as the sum of certain basic functions $B_n(m)$; thus there are only 2^{2^n} different Boolean functions of n variable events.

§ 1. Introduction

Some time ago, the second named author has proved ([1], see also [2]) the following

THEOREM 1. *Let $F_j = F_j(A_1, A_2, \dots, A_n)$ ($j=1, 2, \dots, N$) be arbitrary Boolean functions of the n variable events A_1, \dots, A_n . The linear inequality*

$$(1.1) \quad \sum_{j=1}^N c_j P(F_j) \geq 0$$

(where c_1, \dots, c_N are real constants) is valid in every probability space S if it is valid in the trivial probability space $S_1(1)$.

This simple theorem is useful because it makes it possible to reduce the proof of any linear inequality among probabilities of Boolean functions to a corresponding combinatorial inequality.

To make this paper self-contained we reproduce here the proof of Theorem 1, especially as the proof is very short.

PROOF OF THEOREM 1. Let the expression of the functions F_1, \dots, F_N in canonical form be

$$(1.2) \quad F_j = \sum_{m \in E_j} B_n(m) \quad (j=1, 2, \dots, N)$$

where E_j is some subset of the set $\{0, 1, \dots, 2^n - 1\}$. It follows from (0. 5) that

$$(1. 3) \quad P(F_j) = \sum_{m \in E_j} P(B_n(m))$$

and thus

$$(1. 4) \quad \sum_{j=1}^N c_j P(F_j) = \sum_{m=0}^{2^n-1} d_m P(B_n(m))$$

where

$$(1. 5) \quad d_m = \sum_{m \in E_j} c_j$$

Now evidently if $A_k = \Omega$ if $\varepsilon_k(m) = 1$ and $A_k = \emptyset$ if $\varepsilon_k(m) = -1$, then $B_n(m) = \Omega$ and $B_n(l) = \emptyset$ for $l \neq m, 0 \leq l \leq 2^n - 1$, thus for this special choice of the values of the variables A_1, \dots, A_n we have

$$(1. 6) \quad \sum_{j=1}^N c_j P(F_j) = d_m$$

Thus if (1.1) holds on $S_1(1)$ we have $d_m \geq 0$ for $m = 0, 1, \dots, 2^n - 1$ and thus it follows from (1. 4) that (1.1) holds for every choice of the values of the events A_1, \dots, A_n in every probability space S . Thus Theorem 1 is proved.

It is evident that Theorem 1 can be used also to prove identities. To prove that a relation

$$(1. 7) \quad \sum_{j=1}^N c_j P(F_j) = 0$$

is valid, according to Theorem 1 it is sufficient to verify that (1. 7) holds if all A_k are equal either to Ω or to \emptyset .

A typical example of an inequality which can be obtained as a special case of Theorem 1 is the following inequality, due to GUMBEL ([3]): Putting

$$(1. 8) \quad \sigma_k^{(n)} = \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} P(A_{i_1} A_{i_2} \dots A_{i_k}) \quad (k = 1, 2, \dots, n)$$

one has for $2 \leq k \leq n$

$$(1. 9) \quad (n - k + 1) \sigma_{k-1}^{(n)} \leq \binom{n}{k} + (k - 1) \sigma_k^{(n)}.$$

By means of Theorem 1 the proof of (1. 9) is reduced to a simple inequality between binomial coefficients (see [2], p. 30).

The aim of this paper is to prove a theorem similar to Theorem 1 for *quadratic* (instead of linear) inequalities. This will be done in § 2. In § 3 we give some applications of the general theorem of § 2. In § 4 we discuss the possibility of generalizing the result of § 2 to polynomial inequalities of the third and still higher degrees.

§ 2. A General Theorem on Quadratic Inequalities

In this § we consider quadratic inequalities of the form

$$(2. 1) \quad \sum_{i=1}^N \sum_{j=1}^N c_{i,j} P(F_i) P(F_j) \geq 0$$

where the $c_{i,j}$ are real constants, and F_1, F_2, \dots, F_N are Boolean functions of the variable events A_1, \dots, A_n .

Note that it is no restriction that in (2.1) no linear terms occur, because one of the F_i may be equal to Ω (which is also a Boolean function, namely a constant function) and thus inequalities which contain both quadratic and linear terms can be also written in the form (2.1).

We shall call an inequality (2.1) *exact*, if in (2.1) the equality sign is valid every time when each A_k is equal either to Ω or to \emptyset . By other words (2.1) is exact if equality is valid in (2.1) when the variables A_1, \dots, A_n are restricted to events in the trivial probability space $S_1(1)$.

We shall prove now the following

THEOREM 2. *Let (2.1) be an exact inequality. In order that (2.1) should be valid on every probability space S it is sufficient (and of course also necessary) that it should be valid on the probability space $S_2(\frac{1}{2}, \frac{1}{2})$.*

PROOF OF THEOREM 2. Let again (1.2) be the expression of the function $F_j(1 \leq j \leq N)$ in canonical form. In view of (1.3) we get

$$(2.2) \quad \sum_{i=1}^N \sum_{j=1}^N c_{i,j} P(F_i)P(F_j) = \sum_{r=0}^{2^n-1} \sum_{s=0}^{2^n-1} d_{r,s} P(B_n(r))P(B_n(s)),$$

where

$$(2.3) \quad d_{r,s} = \sum_{\substack{r \in E_i \\ s \in E_j}} c_{i,j}$$

Now let us choose $A_k = \Omega$ if $\varepsilon_k(r) = 1$ and $A_k = \emptyset$ if $\varepsilon_k(r) = -1$ ($k = 1, 2, \dots, n$).

It follows that $P(B_n(r)) = 1$ and $P(B_n(s)) = 0$ if $s \neq r$; thus for this special choice of the values of the variables A_1, \dots, A_n we have

$$(2.4) \quad \sum_{i=1}^N \sum_{j=1}^N c_{i,j} P(F_i)P(F_j) = d_{r,r}$$

As we have supposed that the inequality (2.1) is exact, it follows that

$$(2.5) \quad d_{r,r} = 0 \quad \text{for } 0 \leq r \leq 2^n - 1.$$

Putting

$$(2.6) \quad D_{r,s} = d_{r,s} + d_{s,r} \quad \text{for } r \neq s$$

we obtain

$$(2.7) \quad \sum_{i=1}^N \sum_{j=1}^N c_{i,j} P(F_i)P(F_j) = \sum_{0 \leq r < s \leq 2^n - 1} D_{r,s} P(B_n(r))P(B_n(s))$$

Now let us choose an arbitrary pair (r, s) of integers, $0 \leq r < s \leq 2^n - 1$, and let us choose the values of the events A_k as follows:

$$(2.8) \quad \begin{aligned} A_k &= \Omega & \text{if } \varepsilon_k(r) = \varepsilon_k(s) = 1 \\ A_k &= \alpha & \text{if } \varepsilon_k(r) = 1 \quad \text{and} \quad \varepsilon_k(s) = -1 \\ A_k &= \beta & \text{if } \varepsilon_k(r) = -1 \quad \text{and} \quad \varepsilon_k(s) = +1 \\ A_k &= \emptyset & \text{if } \varepsilon_k(r) = \varepsilon_k(s) = -1 \end{aligned}$$

where α and β are the events $\alpha = \{\omega_1\}$, $\beta = \{\omega_2\}$ of the probability space $S_2(\frac{1}{2}, \frac{1}{2})$. For this special choice of the values of the variables A_k we have clearly

$$(2.9) \quad B_n(r) = \alpha, B_n(s) = \beta \quad \text{and} \quad B_n(t) = \emptyset \quad \text{for} \quad t \neq r, t \neq s.$$

Thus we obtain for this choice of the values of the A_k

$$(2.10) \quad P(B_n(r)) = P(B_n(s)) = \frac{1}{2}, \quad P(B_n(t)) = 0 \quad \text{for} \quad t \neq r, t \neq s,$$

and therefore

$$(2.11) \quad \sum_{i=1}^N \sum_{j=1}^N c_{i,j} P(F_i)P(F_j) = \frac{1}{4} D_{r,s}$$

Thus if (2.1) is valid on $S_2(\frac{1}{2}, \frac{1}{2})$, then we must have $D_{r,s} \geq 0$ for all pairs (r, s) and thus in view of (2.7) it follows that (2.1) is valid on every probability space S and for every choice of the value of the variables A_k .

Thus Theorem 2 is proved.

Similarly as Theorem 1, Theorem 2 can be used also to prove identities. As a matter of fact we obtain from Theorem 2 the following

COROLLARY. *If*

$$(2.12) \quad \sum_{i=1}^N \sum_{j=1}^N c_{i,j} P(F_i)P(F_j) = 0$$

holds on $S_1(1)$ and on $S_2(\frac{1}{2}, \frac{1}{2})$, then it holds identically on every probability space.

§ 3. Some Applications of the General Theorem of § 2

In this § we consider some examples of quadratic inequalities which can be easily proved by means of Theorem 2.

EXAMPLE 1. Let us put $\sigma_0^{(n)} = 1$ and

$$(3.1) \quad \sigma_k^{(n)} = \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} P(A_{i_1} A_{i_2} \dots A_{i_k})$$

We shall prove that the inequality

$$(3.2) \quad k\sigma_k^{(n)} \geq \sigma_{k-1}^{(n)}(\sigma_1^{(n)} - k + 1) \quad (k = 1, 2, \dots, n)$$

is valid.

To prove (3.2) we first remark that it is a quadratic inequality of type (2.1). Further it is easy to see that (3.2) is an exact inequality. As a matter of fact if l among the events A_1, \dots, A_n are equal to Ω and the other $n-l$ to \emptyset , then three cases are possible:

a) either $l \leq k-2$, in which case $\sigma_k^{(n)} = \sigma_{k-1}^{(n)} = 0$ and thus both sides of (3.2) are equal to 0,

b) or $l = k-1$ in which case $\sigma_k^{(n)} = 0$ and $\sigma_1^{(n)} - k + 1 = 0$ and thus again both sides of (3.2) are equal to 0,

c) or $l \geq k$, in which case $\sigma_k^{(n)} = \binom{l}{k}$, $\sigma_{k-1}^{(n)} = \binom{l}{k-1}$ and $\sigma_1^{(n)} = l$. As however

$$k \binom{l}{k} = \binom{l}{k-1} (l - k + 1)$$

we have equality in (3.2) in this case too. Thus (3.2) is exact. Now let us check that (3.2) holds for $S_2(\frac{1}{2}, \frac{1}{2})$. Suppose that among the events A_1, \dots, A_n l_1 are equal to Ω , l_2 to α , l_3 to β ($l_1 + l_2 + l_3 \leq n$) and the remaining $n - l_1 - l_2 - l_3$ to \emptyset . In this case

$$\sigma_j^{(n)} = \frac{1}{2} \left[\binom{l_1 + l_2}{j} + \binom{l_1 + l_3}{j} \right] \quad \text{for } 1 \leq j \leq n$$

and thus

$$(3.3) \quad k\sigma_k^{(n)} - \sigma_{k-1}^{(n)}(\sigma_1^{(n)} - k + 1) = \frac{1}{4}(l_2 - l_3) \left[\binom{l_1 + l_2}{k-1} - \binom{l_1 + l_3}{k-1} \right] \geq 0$$

Thus by Theorem 2 (3.2) holds on every probability space S for any choice of the events A_1, \dots, A_n .

It is interesting to compare (3.2) with GUMBEL's inequality (1.9). The fact that (3.2) is exact, while in GUMBEL's inequality we have equality (as seen from the proof) on $S_1(1)$ only if $l = n$ or $l = n - 1$, shows that, (3.2) gives sometimes a better estimate than (1.9). Another such instance is when the events all have probability $\frac{1}{2}$, and $k = 2$. In this case (1.9) gives for $\sigma_2^{(n)}$ only the trivial lower estimate 0, while (3.2) gives the non-trivial (in fact, asymptotically best possible) lower estimate $\sigma_2^{(n)} \geq \frac{n(n-2)}{8}$.

For $k = 2$ we obtain as a special case of (3.2) the well known inequality

$$(3.4) \quad \sigma_2^{(n)} \geq \binom{\sigma_1^{(n)}}{2}.$$

It follows from (3.2) by induction that

$$(3.5) \quad \sigma_k^{(n)} \geq \binom{\sigma_1^{(n)}}{k}.$$

It should be noted that one can deduce from (3.4) the following inequality:

$$\text{If } \sigma_2^{(n)} \leq \binom{n}{2} p^2 \quad \text{then} \quad \sigma_1^{(n)} \leq np + \frac{1}{2}(1-p) + \frac{1-p^2}{4p(2n-1)}$$

As a matter of fact, it follows from (3.4) and the inequality $\sqrt{1+x} \leq 1 + \frac{x}{2}$ that

$$\sigma_1^{(n)} \leq \frac{1 + \sqrt{1 + 8\sigma_2^{(n)}}}{2} \leq \frac{1}{2} + \frac{1}{2}(2np - p) \sqrt{1 + \frac{1-p^2}{(2np-p)^2}}$$

and thus that

$$\sigma_1^{(n)} \leq np + \frac{1-p}{2} + \frac{1-p^2}{4p(2n-1)}$$

REMARK. The exact maximum of $\sigma_1^{(n)}$ under condition $\sigma_2^{(n)} \leq \binom{n}{p} p^2$ was determined in [4].

EXAMPLE 2. Let us consider the quadratic relation

$$(3.6) \quad P^2(A+B) + P^2(AB) = P^2(A) + P^2(B) + 2P(A\bar{B})P(\bar{A}B)$$

It is evidently valid on $S_1(1)$ and also on $S_2(\frac{1}{2}, \frac{1}{2})$, thus it holds identically.

§ 4. Cubic Inequalities

Theorem 2 can be generalized for cubic inequalities

$$(4.1) \quad \sum_{i_1=1}^N \sum_{i_2=1}^N \sum_{i_3=1}^N c_{i_1, i_2, i_3} P(F_{i_1})P(F_{i_2})P(F_{i_3}) \cong 0$$

where F_1, \dots, F_N are Boolean functions of the variable events A_1, \dots, A_n . The inequality (4.1) is called *exact of order 2* if for every p ($0 \leq p \leq 1$) equality stands in (4.1), if A_1, \dots, A_n are all events of $S_2(p, 1-p)$. (Clearly an inequality which is exact of order 2 is exact.)

We prove the following

THEOREM 3. *Let (4.1) be an inequality which is exact of order 2. If (4.1) holds on $S_3(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$, it holds on every probability space.*

PROOF. If (1.2) is the canonical form of F_j we have

$$(4.2) \quad \begin{aligned} & \sum_{i_1=1}^N \sum_{i_2=1}^N \sum_{i_3=1}^N c_{i_1, i_2, i_3} P(F_{i_1})P(F_{i_2})P(F_{i_3}) = \\ & = \sum_{r_1=0}^{2^n-1} \sum_{r_2=0}^{2^n-1} \sum_{r_3=0}^{2^n-1} d(r_1, r_2, r_3) P(B_n(r_1))P(B_n(r_2))P(B_n(r_3)) \end{aligned}$$

where

$$(4.3) \quad d(r_1, r_2, r_3) = \sum_{r_h \in E_{i_h} (h=1, 2, 3)} c_{i_1, i_2, i_3} (h=1, 2, 3)$$

Clearly (4.1) being exact implies that

$$d(r, r, r) = 0 \quad (0 \leq r \leq 2^n - 1).$$

Let us put for $r \neq s$

$$D(r, s) = d(r, r, s) + d(r, s, r) + d(s, r, r).$$

Now from the supposition that in (4.1) equality holds on $S_2(p, q)$ ($q = 1-p$) it follows that for any pair of numbers r, s ($r \neq s$)

$$(4.4) \quad D(r, s)p + D(s, r)q = 0.$$

By supposition (4.4) holds for $p = \frac{1}{2}$ and also for some p for which $0 < p < \frac{1}{2}$; it follows that

$$(4.5) \quad D(r, s) = 0 \quad \text{if } s \neq r.$$

Thus we obtain, putting

$$\begin{aligned} D(r_1, r_2, r_3) &= d(r_1, r_2, r_3) + d(r_1, r_3, r_2) + d(r_2, r_1, r_3) + \\ &+ d(r_2, r_3, r_1) + d(r_3, r_1, r_2) + d(r_3, r_2, r_1) \end{aligned}$$

that

$$(4.6) \quad \begin{aligned} & \sum_{i_1=1}^N \sum_{i_2=1}^N \sum_{i_3=1}^N c_{i_1, i_2, i_3} P(F_{i_1})P(F_{i_2})P(F_{i_3}) = \\ & = \sum_{0 \leq r_1 < r_2 < r_3 \leq 2^n - 1} D(r_1, r_2, r_3) P(B_n(r_1))P(B_n(r_2))P(B_n(r_3)) \end{aligned}$$

Now let r_1, r_2, r_3 be any three different numbers, $0 \leq r_1 < r_2 < r_3 \leq 2^n - 1$. Let us denote the atoms of $S_3(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ by α_1, α_2 and α_3 . Let us put

$$A_k = \sum_{\varepsilon_k(r_i)=1} \alpha_i$$

It is easy to show that for this choice of the values of the variable events A_k we have

$$(4.7) \quad B_n(r_i) = \alpha_i \quad (i=1, 2, 3).$$

As a matter of fact $A_k^{\varepsilon_k(r_i)} \supseteq \alpha_i$ ($k=1, 2, \dots, n$) thus

$$(4.8) \quad B_n(r_i) = \prod_{k=1}^n A_k^{\varepsilon_k(r_i)} \supseteq \alpha_i$$

As however the events $B_n(r_1), B_n(r_2), B_n(r_3)$ are disjoint, (4.8) implies (4.7).

Clearly (4.7) implies that for any s , different from each of r_1, r_2, r_3 , one has $B_n(s) = \emptyset$. Thus for the above choice of the values of the variables A_1, \dots, A_n we have

$$(4.9) \quad \sum_{i_1=1}^N \sum_{i_2=1}^N \sum_{i_3=1}^N c_{i_1, i_2, i_3} P(F_{i_1})P(F_{i_2})P(F_{i_3}) = \frac{1}{27} D(r_1, r_2, r_3)$$

As by supposition (4.1) holds on $S_3(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$, we obtain from (4.9)

$$(4.10) \quad D(r_1, r_2, r_3) \geq 0 \quad \text{for} \quad 0 \leq r_1 < r_2 < r_3 \leq 2^n - 1.$$

In view of (4.6) it follows that (4.1) holds for every probability space S .

As an example consider the following cubic inequality

$$(4.11) \quad P(AB)P(BC)P(AC) \geq P^2(ABC)[P(AB) + P(AC) + P(BC) - 2P(ABC)]$$

Clearly (4.11) is exact of order two. Thus we have to check only that (4.11) holds on $S_3(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$, which is easily done.

Theorem 3 could be generalized also for polynomial inequalities of degree greater than 3.

REFERENCES

- [1] RÉNYI, A.: Quelques remarques sur les probabilités d'événements dépendantes, *Journal de Math.* 37 (1958) 393—398.
- [2] RÉNYI, A.: *Wahrscheinlichkeitsrechnung mit einer Anhang über Informationstheorie*, Deutscher Verlag der Wissenschaften, Berlin, 1962.
- [3] FRÉCHET, M.: *Les probabilités associées a un système d'événements compatibles et dépendants*, I, II, Hermann, Paris, 1940, 1943.
- [4] RÉNYI, A., ERDŐS, P., NEVEU, J.: An elementary inequality between the probability of events, *Math. Scandinavica* 13 (1963) 99—104.

University of Ghana, Legon (Accra) and Eötvös Loránd University, Budapest

(Received August 10, 1967.)