

ON THE DISTRIBUTION OF THE NUMBER OF  
SUCCESSSES IN INDEPENDENT TRIALS

by

Wassily Hoeffding  
University of North Carolina

This research was supported by the  
the United States Air Force, through  
the Office of Scientific Research of  
the Air Research and Development  
Command

ON THE DISTRIBUTION OF THE NUMBER OF  
SUCCESSSES IN INDEPENDENT TRIALS<sup>1</sup>

by

Wassily Hoeffding  
University of North Carolina

1. Introduction and Summary. Let  $S$  be the number of successes in  $n$  independent trials, and let  $p_j$  denote the probability of success in the  $j$ -th trial,  $j = 1, 2, \dots, n$  (Poisson trials). We consider the problem of finding the maximum and the minimum of  $E g(S)$ , the expected value of a given real-valued function of  $S$ , when  $ES = np$  is fixed. It is well known that the maximum of the variance of  $S$  is attained when  $p_1 = p_2 = \dots = p_n = p$ . This can be interpreted as showing that the variability in the number of successes is highest when the successes are equally probable (Bernoulli trials). This interpretation is further supported by the following two theorems proved in this paper. If  $b$  and  $c$  are two integers,  $0 \leq b \leq np \leq c \leq n$ , the probability  $P(b \leq S \leq c)$  attains its minimum if and only if  $p_1 = p_2 = \dots = p_n = p$  (Theorem 5, a corollary of Theorem 4, which gives the maximum and the minimum of  $P(S \leq c)$ ). If  $g$  is a strictly convex function,  $E g(S)$  attains its maximum if and only if  $p_1 = p_2 = \dots = p_n = p$  (Theorem 3). These results are obtained with the help of two theorems concerning the extrema of the expected value of an arbitrary function  $g(S)$  under the condition  $ES = np$ . Theorem 1 gives necessary conditions for the maximum and the minimum of  $E g(S)$ . Theorem 2 gives a partial characterization of the set of points at which an extremum is attained. Corollary 2.1 states that the maximum and the minimum are attained when  $p_1, p_2, \dots, p_n$  take on at most three different values, only one of which is distinct from 0 and 1. Applications of Theorems 3 and

---

1. This research was supported in part by the United States Air Force, through the Office of Scientific Research of the Air Research and Development Command.

5 to problems of estimation and testing are pointed out in section 5.

## 2. The extrema of the expected value of an arbitrary function of S.

The expected value of a function  $g(S)$  is

$$(1) \quad f(\underline{p}) = Eg(S) = \sum_{k=0}^n g(k) A_{nk}(\underline{p})$$

where  $\underline{p} = (p_1, p_2, \dots, p_n)$ , and  $A_{nk}(\underline{p})$ , the probability of  $S = k$ , is given by

$$A_{nk}(\underline{p}) = \sum_{\substack{i_j=0,1;j=1,\dots,n \\ i_1+\dots+i_n=k}} \prod_{j=1}^n p_j^{i_j} (1-p_j)^{1-i_j}, \quad k = 0, 1, \dots, n.$$

The function  $f(\underline{p})$  is symmetric in the components of  $\underline{p}$  and linear in each component. We observe in passing that conversely any function of  $\underline{p}$  with these two properties can be represented in the form (1). The problem to be considered is to find the maximum and the minimum of  $f(\underline{p})$  in the section D of the hyperplane

$$p_1 + p_2 + \dots + p_n = np \quad (0 < p < 1)$$

which is contained in the closed hypercube

$$0 \leq p_j \leq 1, \quad j = 1, 2, \dots, n.$$

We shall denote by  $\underline{p}^{i_1, i_2, \dots, i_m}$  the point in the  $(n - m)$  - dimensional space which is obtained from  $\underline{p}$  by omitting the coordinates  $p_{i_1}, p_{i_2}, \dots, p_{i_m}$ . Since

$f(\underline{p})$  is symmetric and linear, we can write

$$(2) \quad f(\underline{p}) = f_{n-1,0}(\underline{p}^j) + p_j f_{n-1,1}(\underline{p}^j), \quad j = 1, 2, \dots, n,$$

where the functions  $f_{n-1,0}$  and  $f_{n-1,1}$  are independent of the index  $j$  and symmetric and linear in the components of  $\underline{p}^j$ . In general we define the functions  $f_{n-k,i}$  by  $f_{n,0}(\underline{p}) = f(\underline{p})$ ,

$$(3) \quad f_{n-k,i}(\underline{p}^{1,2,\dots,k}) = f_{n-k-1,i}(\underline{p}^{1,2,\dots,k+1}) + p_{k+1} f_{n-k-1,i+1}(\underline{p}^{1,2,\dots,k+1}),$$

$$i = 0, 1, \dots, k; \quad k = 0, 1, \dots, n-1.$$

Applying (3) repeatedly, we obtain

$$(4) \quad f(\underline{p}) = \sum_{i=0}^m C_{mi}(p_1, p_2, \dots, p_m) f_{n-m,i}(\underline{p}^{1,2,\dots,m}), \quad m = 1, 2, \dots, n,$$

where  $C_{m0}, C_{m1}, \dots, C_{m,m}$  are the symmetric sums

$$C_{m0}(p_1, p_2, \dots, p_m) = 1,$$

(5)

$$C_{mi}(p_1, p_2, \dots, p_m) = (p_1 p_2 \dots p_i) + (p_1 \dots p_{i-1} p_{i+1}) + \dots + (p_{m-i+1} p_{m-i+2} \dots p_m), \quad i > 0.$$

If we write  $(0^u 1^v)$  for the point whose first  $u$  coordinates are 0 and the remaining  $v$  coordinates are 1, and let  $(p_1, p_2, \dots, p_m) = (0^{m-h} 1^h)$ ,  $h = 0, 1, \dots, m$ , we obtain from (4) a system of linear equations for  $f_{n-m,i}(\underline{p}^{1,2,\dots,m})$  whose solution is

$$(6) \quad f_{n-m,i}(\underline{p}^{1,2,\dots,m}) = \sum_{h=0}^i (-1)^{i-h} \binom{i}{h} f(0^{m-h} 1^h, p_{m+1}, \dots, p_n), \quad i = 0, 1, \dots, m.$$

Theorem 1. Let  $\underline{a} = (a_1, a_2, \dots, a_n)$  be a point in  $D$  at which  $f(\underline{p})$  attains its maximum. Then for every two distinct indices  $i, j$  we have

$$(7) \quad f_{n-2,2}(\underline{a}^{ij}) \leq 0 \quad \text{if} \quad a_i \neq a_j,$$

$$(8) \quad f_{n-2,2}(\underline{a}^{ij}) = 0 \quad \underline{\text{if}} \quad a_i \neq a_j, \quad 0 < a_i < 1, \quad 0 < a_j < 1,$$

$$(9) \quad f_{n-2,2}(\underline{a}^{ij}) \geq 0 \quad \underline{\text{if}} \quad 0 < a_i = a_j < 1.$$

The inequalities (7) and (9) are strict if the maximum is not attained at the points in D which differ from  $\underline{a}$  only in that  $a_i$  and  $a_j$  are replaced by  $a_i + x$  and  $a_j - x$  with  $|x|$  positive and arbitrarily small.

Proof. Let  $\underline{a}'$  denote the point which is obtained from  $\underline{a}$  if  $a_i$  and  $a_j$  are replaced by  $a_i + x$  and  $a_j - x$ . The point  $\underline{a}'$  is in D for all  $x$  in the interval I defined by  $0 \leq a_i + x \leq 1$ ,  $0 \leq a_j - x \leq 1$ . By (4) we have

$$f(\underline{a}') = f_{n-2,0}(\underline{a}^{ij}) + (a_i + a_j) f_{n-2,1}(\underline{a}^{ij}) + (a_i + x)(a_j - x) f_{n-2,2}(\underline{a}^{ij}).$$

Hence

$$(10) \quad f(\underline{a}') - f(\underline{a}) = x(a_j - a_i - x) f_{n-2,2}(\underline{a}^{ij}).$$

Since  $f(\underline{a})$  is a maximum, the right side of (10) must be negative or zero for all  $x$  in I. We may assume that  $a_i \leq a_j$ . If  $a_i < a_j$ , and  $x$  is positive and sufficiently small, then  $x$  is in I. Hence (7) must hold. If  $0 < a_i < 1$  and  $0 < a_j < 1$ , then the point  $x = 0$  is in the interior of I. Hence (8) and (9) must hold.

If the maximum is not attained at  $\underline{a}'$  when  $x$  is in I and is different from and sufficiently close to zero, the inequalities (7) and (9) must be strict. The proof is complete.

The following explicit expressions for  $f_{n-2,2}(\underline{a}^{ij})$  will be useful in the applications of Theorem 1. It is easily seen from probability considerations that

$$A_{nk}(0^{2-h}, 1^h, p_3, \dots, p_n) = A_{n-2,k-h}(p_3, \dots, p_n), \quad h = 0, 1, 2.$$

Hence, from (6) and (1),

$$(11) \quad f_{n-2,2}(\underline{a}^{ij}) = \sum_{k=0}^n g(k) \left\{ A_{n-2,k-2}(\underline{a}^{ij}) - 2A_{n-2,k-1}(\underline{a}^{ij}) + A_{n-2,k}(\underline{a}^{ij}) \right\}.$$

Alternatively this can be written in the forms

$$(12) \quad f_{n-2,2}(\underline{a}^{ij}) = \sum_{k=0}^{n-1} \left\{ g(k+1) - g(k) \right\} \left\{ A_{n-2,k-1}(\underline{a}^{ij}) - A_{n-2,k}(\underline{a}^{ij}) \right\}$$

and

$$(13) \quad f_{n-2,2}(\underline{a}^{ij}) = \sum_{k=0}^{n-2} \left\{ g(k+2) - 2g(k+1) + g(k) \right\} A_{n-2,k}(\underline{a}^{ij}).$$

In general the maximum or the minimum of  $f(\underline{p})$  can be attained at more than one point in  $D$ . Thus if  $np < n-1$ , the function  $p_1 p_2 \dots p_n$  attains its minimum 0 at every point in  $D$  with at least one zero coordinate, and there are infinitely many points with this property. The following theorem gives some information about the set of points at which an extremum is attained.

Theorem 2. Let  $\underline{a}$  be a point in  $D$  at which  $f(\underline{p})$  attains its maximum or its minimum. Suppose that  $\underline{a}$  has at least two unequal coordinates which are distinct from 0 and 1. Then

(i)  $f(\underline{p})$  attains its maximum (or minimum) at any point in  $D$  which has the same number of zero coordinates and the same number of unit coordinates as  $\underline{a}$  has;

(ii) if  $\underline{a}$  has exactly  $r$  zero coordinates and  $s$  unit coordinates, the maximum (or minimum) of  $f(\underline{p})$  is equal to

$$(14) \quad f(\underline{a}) = (1-np+s) g(s) + (np-s) g(s+1),$$

and we have

$$(15) \quad g(s+k) = kg(s+1) - (k-1) g(s), \quad k = 2, \dots, n-r-s.$$

Proof. Let  $m = n-r-s$  be the number of coordinates of  $\underline{a} = (a_1, a_2, \dots, a_n)$  which are distinct from 0 and 1. We may take  $a_1, a_2, \dots, a_m$  to be these coordinates, and we may assume that  $a_1 \neq a_2$ . We first show that

$$(16) \quad f_{n-k,i}(a_{k+1}, \dots, a_n) = 0, \quad i = 2, \dots, k,$$

for  $k = 2, \dots, m$ .

Equations (16) will be proved by induction on  $k$ . That (16) is true for  $k = 2$  follows from Theorem 1, (8). Assume that (16) is true for a fixed  $k$ ,  $2 \leq k < m$ . Let

$$(17) \quad \underline{b}_k = (b_1, b_2, \dots, b_k, a_{k+1}, \dots, a_n),$$

where

$$(18) \quad b_1 + \dots + b_k = a_1 + \dots + a_k, \quad 0 \leq b_i \leq 1, \quad i = 1, \dots, k.$$

The point  $\underline{b}_k$  is in  $D$ . By (4) and the induction hypothesis,

$$(19) \quad f(\underline{b}_k) = f_{n-k,0}(a_{k+1}, \dots, a_n) + (a_1 + \dots + a_k) f_{n-k,1}(a_{k+1}, \dots, a_n) = f(\underline{a}).$$

Thus the maximum is attained at every point  $\underline{b}_k$  which satisfies (17) and (18). In particular, (18) can be satisfied with  $b_1 \neq b_2$ ,  $b_1 \neq a_{k+1}$ ,  $b_2 \neq a_{k+1}$ ,  $0 < b_i < 1$ ,  $i = 1, 2, \dots, k$ . Under these assumptions we can apply the induction hypothesis (16) with  $\underline{a}$  replaced by the point  $\underline{b}_k$  whose first  $k+1$  coordinates can be suitably rearranged. Hence

$$f_{n-k,i}(b_1, a_{k+2}, \dots, a_n) = 0, \quad f_{n-k,i}(b_2, a_{k+2}, \dots, a_n) = 0, \quad i = 2, \dots, k.$$

Applying (3) to the left sides of these equations we obtain

$$f_{n-k-1,i}(a_{k+2}, \dots, a_n) + b_h f_{n-k-1,i+1}(a_{k+2}, \dots, a_n) = 0, \quad i = 2, \dots, k; \quad h = 1, 2.$$

Since  $b_1 \neq b_2$  we find that (16) is satisfied with  $k$  replaced by  $k + 1$ . Thus (16) holds for  $k = 2, \dots, m$ .

By (16) with  $k = m$ , equations (19) hold with  $k = m$  for every  $\underline{b}_m$  which satisfies (17) and (18). Since  $f$  is symmetric, this implies part (i) of the theorem.

To prove part (ii) we observe that

$$a_1 + a_2 + \dots + a_m = np - s,$$

and we can put  $(a_{m+1}, \dots, a_n) = (0^r 1^s)$ . Hence, by (19) and (16) with  $k = m$ ,

$$(20) \quad f(\underline{a}) = f_{n-m,0}(0^r 1^s) + (np-s) f_{n-m,1}(0^r 1^s)$$

and

$$(21) \quad f_{n-m,i}(0^r 1^s) = 0, \quad i = 2, \dots, m.$$

Applying (6) and then (1), we obtain

$$\begin{aligned} f_{n-m,i}(0^r 1^s) &= \sum_{h=0}^i (-1)^{i-h} \binom{i}{h} f(0^{m+r-h} 1^{s+h}) \\ &= \sum_{h=0}^i (-1)^{i-h} \binom{i}{h} g(s+h), \quad i = 0, 1, \dots, m. \end{aligned}$$

Hence (14) follows from (20), and (15) is obtained by solving the equations (21) for  $g(s+2), \dots, g(s+m)$ . The proof is complete.

The following immediate corollary of Theorem 2(i) is often convenient for finding an extremum.

Corollary 2.1. The maximum and the minimum of  $f(p)$  in  $D$  are attained at points whose coordinates take on at most three different values, only one of which is distinct from 0 and 1.

Thus to find an extremum, it is sufficient to determine the numbers  $r$  and  $s$  of



the zero and unit coordinates of an extremal point whose remaining coordinates are all equal. We shall see that  $r$  and  $s$  can sometimes be determined with the help of Theorem 1. If an extremum is attained at only one point (perhaps except for permutations of the coordinates), part (ii) of Theorem 2 will prove useful to establish the uniqueness.

3. The maximum of the expected value of a convex function of  $S$ .

Theorem 3. If  $ES = np$  and

$$(22) \quad g(k+2) - 2g(k+1) + g(k) > 0, \quad k = 0, 1, \dots, n-2,$$

then

$$(23) \quad E g(S) \leq \sum_{k=0}^n g(k) \binom{n}{k} p^k (1-p)^{n-k},$$

where the sign of equality holds if and only if  $p_1 = p_2 = \dots = p_n = p$ .

Thus in particular every absolute moment of  $S$ ,  $E(|S-b|^c)$ , about an arbitrary point  $b$ , which is of order  $c > 1$ , attains its maximum if and only if all the  $p_j$  are equal.

Proof of Theorem 3. Let  $\underline{a} = (a_1, a_2, \dots, a_n)$  be a point in  $D$  at which  $f(\underline{p}) = Eg(S)$  attains its maximum. Suppose that  $a_i \neq a_j$  for some  $i, j$ . By Theorem 1,  $f_{n-2,2}(\underline{a}^{ij}) \leq 0$ . By (13) and (22) this implies  $A_{n-2,k}(\underline{a}^{ij}) = 0$ ,  $k = 0, 1, \dots, n-2$ . But this is impossible since the sum of the probabilities  $A_{n-2,0}, A_{n-2,1}, \dots, A_{n-2,n-2}$  is 1. Hence the maximum is attained if and only if all the  $a_j$  are equal, that is,  $a_1 = a_2 = \dots = a_n = p$ . This implies (23) and completes the proof.

Observe that in the proof of Theorem 3 no use was made of Theorem 2. Only inequality (7) of Theorem 1 was needed.

#### 4. The extrema of certain probabilities.

In this section we consider the determination of the maxima and the minima of the probabilities  $P(S \leq c)$  and  $P(b \leq S \leq c)$  when  $ES = np$ .

Theorem 4. If  $ES = np$ , and  $c$  is an integer,

$$(24) \quad 0 \leq P(S \leq c) \leq \sum_{k=0}^c \binom{n}{k} p^k (1-p)^{n-k} \quad \text{if } 0 \leq c \leq np - 1,$$

$$(25) \quad 0 < 1 - Q(n-c-1, 1-p) \leq P(S \leq c) \leq Q(c, p) < 1 \quad \text{if } np-1 < c < np,$$

$$(26) \quad \sum_{k=0}^c \binom{n}{k} p^k (1-p)^{n-k} \leq P(S \leq c) \leq 1 \quad \text{if } np \leq c \leq n,$$

where

$$(27) \quad Q(c, p) = \max_{0 \leq s \leq c} \sum_{k=0}^{c-s} \binom{n-s}{k} a^k (1-a)^{n-s-k},$$

$$(28) \quad a = \frac{np-s}{n-s}.$$

The maximizing value of  $s$  satisfies the inequality

$$(29) \quad (c+1-np)(n-s) < n-np$$

unless  $c = n-1$ , in which case  $s = n-1$ .

All bounds are attained. The upper bound for  $0 \leq c \leq np-1$  and the lower bound for  $np \leq c < n$  are attained only if  $p_1 = p_2 = \dots = p_n = p$ .

Theorem 5. If  $ES = np$ , and  $b$  and  $c$  are two integers such that

$0 \leq b \leq np \leq c \leq n$ , then

$$(30) \quad \sum_{k=b}^c \binom{n}{k} p^k (1-p)^{n-k} \leq P(b \leq S \leq c) \leq 1.$$

Both bounds are attained. The lower bound is attained only if  $p_1 = p_2 = \dots = p_n = p$  unless  $b = 0$  and  $c = n$ .

Proof of Theorem 4. We first consider the maximum of  $f(\underline{p}) = P(S \leq c \mid \underline{p})$  in  $D$ . By Corollary 2.1 the maximum is attained at a point  $\underline{a} = (0^r a^{n-r-s} 1^s)$  (using a notation similar to that employed in section 2), where  $r \geq 0$ ,  $s \geq 0$ ,  $n-r-s \geq 0$ , and

$$(n - r - s) a = np - s.$$

If  $c \geq np$ , let  $s$  be the greatest integer contained in  $np$ , and  $r = n - s - 1$ . Then  $a = np - s$  and  $P(S \leq c \mid \underline{a}) = 1$ . Hence the (obvious) upper bound in (26) is attained.

Now let  $0 \leq c < np$ . If  $s > c$ ,  $P(S \leq c \mid \underline{a}) = 0$ . But  $P(S \leq c \mid p, p, \dots, p) > 0$  for all  $c \geq 0$ . Hence we must have  $s \leq c$ . Since  $a \leq 1$ , we have  $n - r \geq np$ . If  $n - r = np$ , then  $\underline{a} = (0^r 1^{n-r})$  and  $n - r > c$ , hence  $P(S \leq c \mid \underline{a}) = 0$ . Thus we must have  $n - r > np$ . Hence we have the inequalities

$$0 \leq s \leq c < np < n - r \leq n,$$

and this implies  $0 < a < 1$ .

We have  $P(S \leq c) = \mathbb{E}g(S)$ , where  $g(k) = 1$  or  $0$  according as  $k \leq c$  or  $k > c$ . Hence by (12),

$$f_{n-2,2}(\underline{a}^{ij}) = A_{n-2,c}(\underline{a}^{ij}) - A_{n-2,c-1}(\underline{a}^{ij}).$$

If  $\underline{a} = (0^r a^{n-r-s} 1^s)$ ,  $\underline{a}^{ij}$  is of the form

$$\underline{a}^{ij} = (0^u a^{n-u-v-2} 1^v).$$

Then

$$A_{n-2,k}(\underline{a}^{ij}) = \binom{n-u-v-2}{k-v} a^{k-v} (1-a)^{n-k-u-2}$$

and

$$f_{n-2,2}(\underline{a}^{ij}) = \frac{(n-u-v-2)!}{(c-v)!(n-c-u-1)!} a^{c-v-1} (1-a)^{n-c-u-2} \left\{ (n-u-v-1)a - c + v \right\}.$$

Since  $0 < a < 1$ , we see that if  $v \leq c \leq n - u - 1$ ,  $f_{n-2,2}(\underline{a}^{ij})$  has the same sign as

$$(n - u - v - 1) a - c + v.$$

Suppose that  $r > 0$ . By Theorem 1 with  $a_i = 0$ ,  $a_j = a$  we must have

$f_{n-2,2}(0^{r-1} a^{n-r-s-1} 1^s) \leq 0$ . Hence  $(n - r - s)a - c + s = np - c \leq 0$ . But this contradicts the assumption. Thus  $r = 0$ ,  $\underline{a} = (a^{r-s} 1^s)$ ,  $(n-s)a = np - s$ ,  $0 \leq s \leq c$ .

Suppose that  $s > 0$ . By Theorem 1 with  $a_i = a$ ,  $a_j = 1$  we must have

$f_{n-2,2}(a^{n-s-1} 1^s) \leq 0$ , that is,

$$(31) \quad (n - s)a - c + s - 1 = np - c - 1 \leq 0.$$

Hence if  $c < np - 1$ , we must have  $r = s = 0$ ,  $a = p$ . Thus the second inequality (24) holds for  $0 \leq c < np - 1$ , and the bound is attained. (We postpone the proof for  $c = np - 1$ .)

Now suppose that  $n - s > 1$ . By Theorem 1 with  $a_i = a_j = a$  we must have

$f_{n-2,2}(a^{n-s-2} 1^s) \geq 0$ , that is,  $(n - s - 1)a - c + s \geq 0$ . This is equivalent to

$$(32) \quad (c + 1 - np)(n - s) \leq n - np.$$

If  $c = n - 1$ , this contradicts the assumption  $n - s > 1$ , and we must have  $s = n - 1$ .

If  $c \neq n - 1$ , we have  $c < n - 1$  and  $n - s > 1$ , so that (32) must be satisfied. Hence if  $c < np$ , the maximum of  $P(S \leq c)$  is  $Q(c, p)$  as defined in (27) and (28), and the maximizing value of  $s$  satisfies (32) and is equal to  $n - 1$  if  $c = n - 1$ . (We postpone the proof of strict inequality in (32) for  $c \neq n - 1$ .) Since  $a > 0$  and  $c - s < n - s$ , we have  $Q(c, p) < 1$ .

We next show that if  $0 \leq c < np$ , the maximum can be attained only at a point whose coordinates which are distinct from 0 and 1 are all equal. Suppose the maximum is attained at a point  $\underline{a}$  which has at least two unequal coordinates which are distinct from 0 and 1. Let  $s$  be <sup>the</sup> number of unit coordinates in  $\underline{a}$ . By Theorem 2, equation (14), we must have  $f(\underline{a}) = 1$  if  $s < c$ ,  $f(\underline{a}) = 1 - rp + s$  if  $s = c$ , and  $f(\underline{a}) = 0$  if  $s > c$ . Since for  $0 \leq c < np$  the maximum is positive and less than 1, we must have  $s = c$ . By (15) with  $s = c$ ,  $k = 2$  we must then have  $g(c + 2) = -1$ , which is not true. Hence the coordinates of  $\underline{a}$  which are not 0 or 1 must be all equal.

By Theorem 1 this implies that the inequalities in (31) and (32) are strict. All statements of Theorem 4 concerning the upper bounds are now easily seen to be true.

The statements concerning the lower bounds follow from the equation

$$P(S \leq c \mid \underline{p}) = 1 - P(S \leq n - c - 1 \mid \underline{q}),$$

where  $\underline{q} = (1 - p_1, 1 - p_2, \dots, 1 - p_n)$ . The proof is complete.

Proof of Theorem 5. Since  $P(b \leq S \leq c) = P(S \leq c) - P(S \leq b - 1)$ , the lower bound in (30) and the condition for its attainment follow from Theorem 4. The upper bound 1 is attained at  $(0^{n-c} a^{c-b} 1^b)$ , where  $(c-b)a = np - b$ .

5. Statistical applications. The lower bound for  $P(b \leq S \leq c)$  which is given in Theorem 5 shows that the usual (one-sided and two-sided) tests for the constant probability of "success" in  $n$  independent (Bernoulli) trials can be used as tests for the average probability  $p$  of success when the probability of success varies from trial to trial. That is to say, the significance level of these tests (which is understood as the upper bound for the probability of an error of the first kind) remains unchanged. Moreover, we can obtain lower bounds for the power of these tests when the alternative is not too close to the hypothesis which is being tested. (Very

roughly, the significance level has to be less than  $1/2$  and the power greater than  $1/2$ .) We can also obtain a confidence interval for  $p$  with a prescribed (sufficiently high) confidence coefficient, and an upper bound for the probability that the confidence interval covers a wrong value of  $p$  when the latter is not too close to the true value. Details are left to the reader.

Theorem 3 can be applied in certain point estimation problems. Suppose we want to estimate a function  $\theta(p)$ , and the loss due to saying  $\theta(p) = t$  is  $W(p, t)$ . If the estimator  $t(S)$  is a function of  $S$  only, and  $W(p, t(S))$  is a strictly convex function of  $S$  for every  $p$ , then Theorem 3 implies that the risk,  $EW(p, t(S))$ , is maximized when all the  $p_j$  are equal. It follows in particular that if  $t(S)$  is a minimax estimator under the assumption that the  $p_j$  are all equal, it retains this property when the assumption is not satisfied (with no restriction on the class of estimators).

One may doubt whether these problems are statistically meaningful since the average probability of success depends on the sample size. The main interest of these results to the practicing statistician seems to be in cases where he assumes that the probability of success is constant, but there is the possibility that this assumption is violated.