ON THE DISTRIBUTION OF THE NUMBER OF SUCCESSES IN INDEPENDENT TRIALS[1]

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 jth trial, $j = 1, 2, \cdots, n$ (Poisson trials). We consider the problem of finding the maximum and the minimum of $Eg(S)$, 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 $p_1 = p_2 = \cdots = p_n = p$. 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, $0 \leqq b \leqq np \leqq c \leqq n$, the probability $P(b \leqq S \leqq c)$ attains its minimum if and only if $p_1 = p_2 = \cdots = p_n = p$, unless $b = 0$ and $c = n$ (Theorem 5, a corollary of Theorem 4, which gives the maximum and the minimum of $P(S \leqq c)$). If g is a strictly convex function, $Eg(S)$ attains its maximum if and only if $p_1 = p_2 = \cdots = p_n = p$ (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 $Eg(S)$. 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 $p_1, p_2, \cdots, p_n$ take on, at most, three different values, only one of which is distinct from 0 and 1. Applications of Theorems 3 and 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

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

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

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

The function $f(\mathbf{p})$ is symmetric in the components of $\mathbf{p}$ and linear in each component. We observe in passing that, conversely, any function of $\mathbf{p}$ with these two properties can be represented in the form (1). The problem to be con-

Received April 15, 1955.

[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.

sidered is to find the maximum and the minimum of $f(\mathbf{p})$ in the section D of the hyperplane

$$p_1 + p_2 + \cdots + p_n = np \qquad (0 < p < 1),$$

which is contained in the closed hypercube

$$0 \leqq p_j \leqq 1, \qquad j = 1, 2, \cdots, n.$$

We shall denote by $\mathbf{p}^{i_1, i_2, \cdots, i_m}$ the point in the $(n - m)$-dimensional space, which is obtained from $\mathbf{p}$ by omitting the coordinates $p_{i_1}, p_{i_2}, \cdots, p_{i_m}$. Since $f(\mathbf{p})$ is symmetric, and linear in each component, we can write

$$(2) \qquad f(\mathbf{p}) = f_{n-1,0}(\mathbf{p}^j) + p_j f_{n-1,1}(\mathbf{p}^j), \qquad j = 1, 2, \cdots, n,$$

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 $\mathbf{p}^j$. In general, we define the functions $f_{n-k,i}$ by $f_{n,0}(\mathbf{p}) = f(\mathbf{p})$, and

$$(3) \quad f_{n-k,i}(\mathbf{p}^{1,2,\cdots,k}) = f_{n-k-1,i}(\mathbf{p}^{1,2,\cdots,k+1}) + p_{k+1} f_{n-k-1,\, i+1}(\mathbf{p}^{1,2,\cdots,k+1}),$$
$$i = 0, 1, \cdots, k,\ k = 0, 1, \cdots, n-1.$$

Applying (3) repeatedly, we obtain

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

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

$$(5) \qquad C_{m0}(p_1, p_2, \cdots, p_m) = 1,$$

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

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 $(p_1, p_2, \cdots, p_m) = (0^{m-h} 1^h)$, $h = 0, 1, \cdots, m$, we obtain from (4) a system of linear equations for $f_{n-m,i}(\mathbf{p}^{1,2,\cdots,m})$ whose solution is

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

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

$$(7) \qquad f_{n-2,2}(\mathbf{a}^{ij}) \leqq 0 \qquad \textit{if } a_i \neq a_j,$$

$$(8) \qquad f_{n-2,2}(\mathbf{a}^{ij}) = 0 \qquad \textit{if } a_i \neq a_j,\ 0 < a_i < 1,\ 0 < a_j < 1,$$

$$(9) \qquad f_{n-2,2}(\mathbf{a}^{ij}) \geqq 0 \qquad \textit{if } 0 < a_i = a_j < 1.$$

The inequalities (7) *and* (9) *are strict if the maximum is not attained at the points in D which differ from* $\mathbf{a}$ *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 $\mathbf{a}'$ denote the point which is obtained from $\mathbf{a}$ if a_i and a_j are replaced by $a_i + x$ and $a_j - x$. The point $\mathbf{a}'$ is in D for all x in the interval I defined by $0 \leqq a_i + x \leqq 1$, $0 \leqq a_j - x \leqq 1$. By (4) we have

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

Hence,

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

Since $f(\mathbf{a})$ is a maximum, the right side of (10) must be negative or zero for all x in I. We may assume that $a_i \leqq a_j$. 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 $\mathbf{a}'$ 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 $f_{n-2,2}(\mathbf{a}^{ij})$ will be useful in the applications of Theorem 1. It is easily seen (for instance, from probability considerations) that

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

Hence, from (6) and (1),

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

Alternatively, this can be written in the forms

$$f_{n-2,2}(\mathbf{a}^{ij}) = \sum_{k=0}^{n-1} \{g(k+1) - g(k)\}\{A_{n-2,k-1}(\mathbf{a}^{ij}) - A_{n-2,k}(\mathbf{a}^{ij})\} \tag{12}$$

and

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

In general, the maximum or the minimum of $f(\mathbf{p})$ can be attained at more than one point in D. Thus, if $np < n - 1$, the function $p_1p_2 \cdots p_n$ 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* $\mathbf{a}$ *be a point in D at which $f(\mathbf{p})$ attains its maximum or its minimum. Suppose that* $\mathbf{a}$ *has at least two unequal coordinates which are distinct from* 0 *and* 1. *Then,*

(i) *$f(\mathbf{p})$ 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* $\mathbf{a}$ *has;*

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

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

and we have

$$g(s + k) = kg(s + 1) - (k - 1)g(s), \qquad k = 2, \cdots, n - r - s. \tag{15}$$

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

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

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

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, $2 \leqq k < m$. Let

$$\mathbf{b}_k = (b_1, b_2, \cdots, b_k, a_{k+1}, \cdots, a_n), \tag{17}$$

where

$$b_1 + \cdots + b_k = a_1 + \cdots + a_k, \qquad 0 \leqq b_i \leqq 1, i = 1, \cdots, k. \tag{18}$$

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

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

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

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

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

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

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

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

To prove part (ii), we observe that

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

and we can put $(a_{m+1}, \cdots, a_n) = (0^r1^s)$. Hence, by (19) and (16), with $k = m$,

$$f(\mathbf{a}) = f_{n-m,0}(0^r1^s) + (np - s)f_{n-m,1}(0^r1^s) \tag{20}$$

and

$$f_{n-m,i}(0^r1^s) = 0, \qquad i = 2, \cdots, m. \tag{21}$$

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

$$\begin{aligned} f_{n-m,i}(0^r1^s) &= \sum_{h=0}^{i} (-1)^{i-h} \binom{i}{h} f(0^{m+r-h}1^{s+h}) \\ &= \sum_{h=0}^{i} (-1)^{i-h} \binom{i}{h} g(s+h), \qquad i = 0, 1, \cdots, m. \end{aligned}$$

Hence, (14) follows from (20). Equations (21) state that the second differences of g are zero in the indicated range. Therefore, the first differences, $g(s + k) - g(s + k - 1)$, are constant for $k = 1, 2, \cdots, m$, which is equivalent to (15). 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 $f(\mathbf{p})$ 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 (except perhaps 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*

$$g(k + 2) - 2g(k + 1) + g(k) > 0, \qquad k = 0, 1, \cdots, n - 2, \tag{22}$$

then

$$Eg(S) \leqq \sum_{k=0}^{n} g(k) \binom{n}{k} p^k(1 - p)^{n-k}, \tag{23}$$

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

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 of the p_j are equal.

PROOF OF THEOREM 3. Let $\mathbf{a} = (a_1, a_2, \cdots, a_n)$ be a point in D at which $f(\mathbf{p}) = Eg(S)$ attains its maximum. Suppose that $a_i \neq a_j$ for some i, j. By Theorem 1, $f_{n-2,2}(\mathbf{a}^{ij}) \leqq 0$. By (13) and (22), this implies $A_{n-2,k}(\mathbf{a}^{ij}) = 0$, $k = 0, 1, \cdots, n - 2$. But this is impossible, since the sum of the probabilities $A_{n-2,0}$,

$A_{n-2,1}, \cdots, A_{n-2,n-2}$ is 1. Hence, the maximum is attained if and only if all the a_j are equal, i.e., if $a_1 = a_2 = \cdots = a_n = p$. 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 $P(S \leqq c)$ and $P(b \leqq S \leqq c)$ when $ES = np$.

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

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

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

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

where

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

$$a = \frac{np - s}{n - s}. \tag{28}$$

The maximizing value of s satisfies the inequality

$$(c + 1 - np)(n - s) < n - np \tag{29}$$

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

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

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

$$0 \leqq b \leqq np \leqq c \leqq n,$$

then

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

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

PROOF OF THEOREM 4. We first consider the maximum of $f(\mathbf{p}) = P(S \leqq c \mid \mathbf{p})$ in D. By Corollary 2.1, the maximum is attained at a point $\mathbf{a} = (0^r a^{n-r-s} 1^s)$ (using a notation similar to that employed in Section 2), where $r \geqq 0$, $s \geqq 0$, $n - r - s \geqq 0$, and

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

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

Now let $0 \leqq c < np$. If $s > c$, $P(S \leqq c \mid \mathbf{a}) = 0$. But

$$P(S \leqq c \mid p, p, \cdots, p) > 0$$

for all $c \geqq 0$. Hence, we must have $s \leqq c$. Since $a \leqq 1$, we have $n - r \geqq np$. If $n - r = np$, then $\mathbf{a} = (0^r 1^{n-r})$ and $n - r > c$, hence $P(S \leqq c \mid \mathbf{a}) = 0$. Thus, we must have $n - r > np$. Consequently, we have the inequalities

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

and this implies $0 < a < 1$.

We have $P(S \leqq c) = Eg(S)$, where $g(k) = 1$ or 0, according as $k \leqq c$ or $k > c$. Hence, by (12),

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

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

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

Then,

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

and

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

Since $0 < a < 1$, we see that if $v \leqq c \leqq n - u - 1$, $f_{n-2,2}(\mathbf{a}^{ij})$ 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 $f_{n-2,2}(0^{r-1} a^{n-r-s-1} 1^s) \leqq 0$. Hence, $(n - r - s)a - c + s = np - c \leqq 0$. But this contradicts the assumption. Thus, $r = 0$, $\mathbf{a} = (a^{r-s} 1^s)$, $(n - s)a = np - s$, $0 \leqq s \leqq c$.

Suppose that $s > 0$. By Theorem 1, with $a_i = a$, $a_j = 1$, we must have $f_{n-2,2}(a^{n-s-1} 1^s) \leqq 0$, i.e.,

$$(n-s)a - c + s - 1 = np - c - 1 \leqq 0. \tag{31}$$

Hence, if $c < np - 1$, we must have $r = s = 0$, $a = p$. Thus, the second inequality (24) holds for $0 \leqq c < np - 1$, 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 $f_{n-2,2}s(a^{n-s-2}1^s) \geqq 0$, i.e., $(n - s - 1)a - c + s \geqq 0$. This is equivalent to

$$(c + 1 - np)(n - s) \leqq n - np. \tag{32}$$

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 $P(S \leqq c)$ 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 $0 \leqq c < np$, 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 **a** which has at least two unequal coordinates which are distinct from 0 and 1. Let s be the number of unit coordinates in **a**. By Theorem 2, equation (14), we must have $f(\mathbf{a}) = 1$ if $s < c$, $f(\mathbf{a}) = 1 - np + s$ if $s = c$, and $f(\mathbf{a}) = 0$ if $s > c$. Since for $0 < c < np$ 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 **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

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

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

Proof of Theorem 5. Since $P(b \leqq S \leqq c) = P(S \leqq c) - P(S \leqq b - 1)$, 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 $(c\text{-}b)a = np - b$.

5. Statistical applications. The lower bound for $P(b \leqq S \leqq c)$, 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 $\frac{1}{2}$ and the power greater than $\frac{1}{2}$.) 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 if $W(p, t(S))$ 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.