The Annals of Probability
1975, Vol. 3, No. 1, 182–188

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

BY LEON JAY GLESER

Purdue University

Let S be the number of successes in n independent Bernoulli trials, where p_j is the probability of success on the jth trial. Let $\mathbf{p} = (p_1, p_2, \cdots, p_n)$, and for any integer c, $0 \leqq c \leqq n$, let $H(c \mid \mathbf{p}) = P\{S \leqq c\}$. Let $\mathbf{p}^{(1)}$ be one possible choice of $\mathbf{p}$ for which $E(S) = \lambda$. For any $n \times n$ doubly stochastic matrix Π, let $\mathbf{p}^{(2)} = \mathbf{p}^{(1)}\Pi$. Then in the present paper it is shown that $H(c \mid \mathbf{p}^{(1)}) \leqq H(c \mid \mathbf{p}^{(2)})$ for $0 \leqq c \leqq [\lambda - 2]$, and $H(c \mid \mathbf{p}^{(1)}) \geqq H(c \mid \mathbf{p}^{(2)})$ for $[\lambda + 2] \leqq c \leqq n$. These results provide a refinement of inequalities for $H(c \mid \mathbf{p})$ obtained by Hoeffding [3]. Their derivation is achieved by applying consequences of the partial ordering of majorization.

1. Introduction and summary. Let S be the number of successes in n independent Bernoulli trials, where p_j is the probability of success on the jth trial, $0 \leqq p_j \leqq 1$. Let

$$\mathbf{p} = (p_1, p_2, \cdots, p_n)\,, \tag{1.1}$$

and for any integer c, $0 \leqq c \leqq n$, let

$$H(c \mid \mathbf{p}) = P\{S \leqq c\}\,. \tag{1.2}$$

For fixed c, we are interested in the relationship between $H(c \mid \mathbf{p}^{(1)})$ and $H(c \mid \mathbf{p}^{(2)})$, where $\mathbf{p}^{(1)}$ and $\mathbf{p}^{(2)}$ each belong to the region

$$D_\lambda = \{\mathbf{p} : 0 \leqq p_i \leqq 1, i = 1, 2, \cdots, n; \textstyle\sum_{i=1}^n p_i = \lambda\}\,. \tag{1.3}$$

That is, $\mathbf{p}^{(1)}$ and $\mathbf{p}^{(2)}$ are sequences of probabilities for the independent Bernoulli trials each of which result in an expected number of successes, $E(S)$, equal to λ.

Hoeffding ([3] Theorem 4) has shown that for all $\mathbf{p} \in D_\lambda$,

$$0 \leqq H(c \mid \mathbf{p}) \leqq H(c \mid n^{-1}(\lambda, \lambda, \cdots, \lambda)) \qquad \text{if} \quad 0 \leqq c \leqq [\lambda - 2]\,, \tag{1.4}$$

$$H(c \mid n^{-1}(\lambda, \lambda, \cdots, \lambda)) \leqq H(c \mid \mathbf{p}) \leqq 1 \qquad \text{if} \quad [\lambda + 2] \leqq c \leqq n\,, \tag{1.5}$$

where

$$H(c \mid n^{-1}(\lambda, \lambda, \cdots, \lambda)) = \textstyle\sum_{k=0}^{c} \binom{n}{k} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k}\,, \tag{1.6}$$

and $[x]$ denotes the greatest integer $\leqq x$. Hoeffding ([3] Theorem 4) also

Received August 21, 1973; revised April 17, 1974.

[1] Research sponsored in part by Air Force Office of Scientific Research, Air Force Systems Command, USAF, under Grant No. 73-2432 at Purdue University. The United States Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation hereon.

AMS 1970 *subject classifications.* Primary 60C05, 60E05; Secondary 62E15, 26A86.

Key words and phrases. Independent Bernoulli trials, number of successes, inequalities on cumulative distribution function, inequalities on expected values, majorization, Schur condition.

obtained bounds on $H(c \mid \mathbf{p})$ for $c = [\lambda - 1]$, $[\lambda]$, and $[\lambda + 1]$. These will be discussed at the end of Section 3.

To motivate the major result of the present paper, let

$$\mathbf{p}^*(\lambda) = n^{-1}(\lambda, \lambda, \cdots, \lambda) \tag{1.7}$$

and

$$\hat{\mathbf{p}}(\lambda) = (1, 1, \cdots, 1, \lambda - [\lambda], 0, 0, \cdots, 0)\,, \tag{1.8}$$

where in $\hat{\mathbf{p}}(\lambda)$ there are $[\lambda]$ ones and $n - [\lambda] - 1$ zeroes. Note that both $\mathbf{p}^*(\lambda)$ and $\hat{\mathbf{p}}(\lambda)$ are elements of D_λ. We have already noted the role $\mathbf{p}^*(\lambda)$ plays in Hoeffding's bounds (1.4) and (1.5), giving the upper bound to $H(c \mid \mathbf{p})$ for $0 \leqq c \leqq [\lambda - 2]$ in (1.4) and the lower bound to $H(c \mid \mathbf{p})$ for $[\lambda + 2] \leqq c \leqq n$ in (1.5). On the other hand, we have $H(c \mid \hat{\mathbf{p}}(\lambda)) = 0$ for $0 \leqq c \leqq [\lambda - 2]$ and $H(c \mid \hat{\mathbf{p}}(\lambda)) = 1$ for $[\lambda + 2] \leqq c \leqq n$; these, of course, are the lower and upper bounds to $H(c \mid \mathbf{p})$ in (1.4) and (1.5) respectively.

Now note that for any $\mathbf{p} \in D_\lambda$, we can write

$$\mathbf{p}^*(\lambda) = \mathbf{p}\Pi^*\,, \tag{1.9}$$

where Π^* is an $n \times n$ doubly stochastic matrix, all of whose elements are n^{-1}. Also, for any $\mathbf{p} \in D_\lambda$, we can write

$$\mathbf{p} = \hat{\mathbf{p}}(\lambda)\Pi(\mathbf{p})\,, \tag{1.10}$$

where $\Pi(\mathbf{p})$ is an $n \times n$ doubly stochastic matrix. [This fact follows directly from Lemma 2.1 of Section 2.] It is thus apparent that proof of the following theorem would yield Hoeffding's inequalities (1.4) and (1.5) as corollaries, and would provide a more detailed picture of the behavior of $H(c \mid \mathbf{p})$ as a function of $\mathbf{p}$.

THEOREM 1.1. *Let $\mathbf{p}^{(1)} \in D_\lambda$ and suppose that there exists a doubly stochastic $n \times n$ matrix Π for which*

$$\mathbf{p}^{(2)} = \mathbf{p}^{(1)}\Pi\,. \tag{1.11}$$

Then $\mathbf{p}^{(2)} \in D_\lambda$ and

$$H(c \mid \mathbf{p}^{(1)}) \leqq H(c \mid \mathbf{p}^{(2)}) \qquad \text{if} \quad 0 \leqq c \leqq [\lambda - 2]\,, \tag{1.12}$$

and

$$H(c \mid \mathbf{p}^{(1)}) \geqq H(c \mid \mathbf{p}^{(2)}) \qquad \text{if} \quad [\lambda + 2] \leqq c \leqq n\,. \tag{1.13}$$

In Section 2, we apply Ostrowski's ([1] Theorem 15) fundamental theorem on majorization to the problem of ordering, over various choices of $\mathbf{p} \in D_\lambda$, the expected values $Eg(S)$ of any function $g(k)$ on $0, 1, 2 \cdots, n$. The results obtained in Section 2 are then used in Section 3 to prove Theorem 1.1.

2. Majorization. A $1 \times n$ vector $\mathbf{x}$ is said to *majorize* a $1 \times n$ vector $\mathbf{y}$ if $x_{[1]} \geqq y_{[1]}$, $x_{[1]} + x_{[2]} \geqq y_{[1]} + y_{[2]}$, $\cdots$, $\sum_{i=1}^{n-1} x_{[i]} \geqq \sum_{i=1}^{n-1} y_{[i]}$, and $\sum_{i=1}^{n} x_{[i]} = \sum_{i=1}^{n} y_{[i]}$, where the $x_{[i]}$'s and $y_{[i]}$'s are the components of $\mathbf{x}$ and $\mathbf{y}$, respectively, arranged in descending order ($x_{[1]} \geqq x_{[2]} \geqq \cdots \geqq x_{[n]}$, and similarly for the $y_{[i]}$'s). The relation of majorization to doubly stochastic matrices is given by the following result of Hardy, Littlewood, and Pólya ([2] page 49). (Note: In [1] this result is incorrectly attributed to Karamata.)

LEMMA 2.1. *The vector* $\mathbf{x}$ *majorizes the vector* $\mathbf{y}$ *if and only if there exists an* $n \times n$ *doubly stochastic matrix* Π *such that* $\mathbf{y} = \mathbf{x}\Pi$.

The following result, originally due to Ostrowski (see [1] pages 30–33), relates majorization to the ordering of the values of functions $F(\mathbf{z})$ over regions of n-dimensional Euclidean space.

LEMMA 2.2. *Let* $F(\mathbf{z})$ *be a permutation-symmetric function defined on n-dimensional vectors* $\mathbf{z} = (z_1, z_2, \cdots, z_n)$. *For any* i, j, $i \neq j$, *and all* $\mathbf{z}$ *in a permutation-symmetric region* D, *suppose that*

$$(2.1) \qquad (z_i - z_j)\left(\frac{\partial F}{\partial z_i} - \frac{\partial F}{\partial z_j}\right) \geqq 0 \, .$$

If $\mathbf{x}, \mathbf{y} \in D$, *and if* $\mathbf{x}$ *majorizes* $\mathbf{y}$, *then*

$$(2.2) \qquad F(\mathbf{x}) \geqq F(\mathbf{y}) \, .$$

A permutation-symmetric function satisfying (2.1) over a region D is said to satisfy a *Schur condition on* D. It should be remarked that the condition that F be permutation-symmetric (*i.e.* $F(z_1, z_2, \cdots, z_n) = F(z_{\sigma(1)}, z_{\sigma(2)}, \cdots, z_{\sigma(n)})$ for all permutations σ, all $z_1, \cdots, z_n$) is incorrectly omitted from the statement of Ostrowski's lemma in [1].

Let $g(k)$ be any function on $0, 1, \cdots, n$, and let S be the number of successes in n independent Bernoulli trials, where p_j is the probability of success on the jth trial. Let

$$(2.3) \qquad h(\mathbf{p}) = Eg(S) \, ,$$

for $\mathbf{p} = (p_1, p_2, \cdots, p_n)$. Then $h(\mathbf{p})$ is a permutation-symmetric function defined over the region

$$D = \{\mathbf{z} : 0 \leqq z_i \leqq 1, i = 1, 2, \cdots, n\} \, .$$

LEMMA 2.3. *For any two components* p_i *and* p_j, $i < j$, *of* $\mathbf{p}$,

$$(2.4) \qquad (p_i - p_j)\left(\frac{\partial h(\mathbf{p})}{\partial p_i} - \frac{\partial h(\mathbf{p})}{\partial p_j}\right) = -(p_i - p_j)^2 \textstyle\sum_{k=0}^{n-2} f(k \mid \mathbf{p}^{ij}) \Delta g(k) \, ,$$

where for any function $s(k)$ *defined on the nonnegative integers*

$$(2.5) \qquad \Delta s(k) \equiv s(k + 2) - 2s(k + 1) + s(k)$$

is the second difference of $s(k)$, *where* $\mathbf{p}^{ij}$ *is the* $1 \times (n - 2)$ *vector formed by deleting the ith and jth components of* $\mathbf{p}$, *and where*

(2.6) $f(k \mid \mathbf{p}^{ij})$ = *probability of* k *successes in the* $n - 2$ *trials other than trials* i *and* j;

for $k = 0, 1, \cdots, n - 2$.

PROOF. We adopt the convention that $f(k \mid \mathbf{p}^{ij}) = 0$ for $k < 0$ or $k > n - 2$. Under this convention,

$$P\{S = k\} = (1 - p_i)(1 - p_j)f(k \mid \mathbf{p}^{ij}) + (p_i + p_j - 2p_ip_j)f(k - 1 \mid \mathbf{p}^{ij}) + p_ip_jf(k - 2 \mid \mathbf{p}^{ij}) . \tag{2.7}$$

Using (2.7), we find that the left-hand side of (2.4) is

$$(p_i - p_j)\left(\frac{\partial h(\mathbf{p})}{\partial p_i} - \frac{\partial h(\mathbf{p})}{\partial p_j}\right) = \sum_{k=0}^{n} g(k)\Delta f(k - 2 \mid \mathbf{p}^{ij}) . \tag{2.8}$$

It is now easily shown that the right-hand sides of (2.4) and (2.8) are equal. □

As a corollary of Lemma 2.3, we can prove a result earlier obtained by Karlin and Novikoff [5].

COROLLARY 2.1. *Suppose that $g(k)$ is convex on $0, 1, \cdots, n - 2$, in the sense that $\Delta g(k) \geqq 0$, $k = 0, 1, \cdots, n - 2$. If $\mathbf{p}^{(1)} \in D_\lambda$ and if $\mathbf{p}^{(2)} = \mathbf{p}^{(1)}\Pi$, where Π is any $n \times n$ doubly stochastic matrix, then $\mathbf{p}^{(2)} \in D_\lambda$ and*

$$h(\mathbf{p}^{(1)}) \leqq h(\mathbf{p}^{(2)}) . \tag{2.9}$$

PROOF. Since $\Delta g(k) \geqq 0$, $k = 0, 1, \cdots, n - 2$, $-h(\mathbf{p})$ satisfies a Schur condition, as can be seen from (2.4). Hence, Lemmas 2.1 and 2.2 imply that $-h(\mathbf{p}^{(1)}) \geqq -h(\mathbf{p}^{(2)})$, from which (2.9) immediately follows. □

Karlin and Novikoff [5] proved Corollary 2.1 in a somewhat different way. Their proof, however, embodies the ideas underlying the usual proof of Lemma 2.2.

From Corollary 2.1 and the arguments in Section 1 relating any $\mathbf{p} \in D_\lambda$ by doubly stochastic matrices to $\mathbf{p}^*(\lambda)$ and $\hat{\mathbf{p}}(\lambda)$, it follows that for any $g(k)$ convex on $0, 1, \cdots, n - 2$, and any $\mathbf{p} \in D_\lambda$,

$$\begin{aligned}(1 - \delta)g([\lambda]) + \delta g([\lambda + 1]) &\leqq h(\mathbf{p}) = Eg(S) \\ &\leqq \sum_{k=0}^{n} g(k)\binom{n}{k}\left(\frac{\lambda}{n}\right)^k\left(1 - \frac{\lambda}{n}\right)^{n-k} ,\end{aligned} \tag{2.10}$$

where $\delta \equiv \lambda - [\lambda]$.

The result (2.10) implies that $E|S - b|^a$, for any $a \geqq 1$ and any real number b, is highest over D_λ when S has a binomial distribution with parameters n and $n^{-1}\lambda$ (i.e., $\mathbf{p} = \mathbf{p}^*(\lambda)$), and lowest when

$$\begin{aligned} S &= [\lambda + 1] , && \text{with probability } \delta , \\ &= [\lambda] , && \text{with probability } 1 - \delta \end{aligned}$$

(i.e., $\mathbf{p} = \hat{\mathbf{p}}(\lambda)$). The upper bound in (2.10) was first obtained (using a different method) by Hoeffding [3]. The lower bound in (2.10) can also be obtained by the methods in Hoeffding's [3] paper.

3. Proof of Theorem 1.1. For fixed integer c, $0 \leqq c \leqq n$, let

$$\begin{aligned} g_c(k) &= 1 && \text{if } 0 \leqq k \leqq c \\ &= 0 && \text{if } c + 1 \leqq k \leqq n . \end{aligned} \tag{3.1}$$

Then

$$H(c \mid \mathbf{p}) = E(g_c(S)) . \tag{3.2}$$

Note that $g_c(k)$ is not convex on $0, 1, \cdots, n-2$ when $c \leqq n-1$, so that we cannot directly use Corollary 2.1 to prove Theorem 2.1. Instead, we make use of Lemmas 2.1 to 2.3.

First note that for $c \leqq n-1$

$$\begin{aligned} \Delta g_c(k) &= 1 , & k &= c , \\ &= -1 , & k &= c-1 , \\ &= 0 , & &\text{otherwise.} \end{aligned} \tag{3.3}$$

Thus, from Lemma 2.3,

$$\begin{aligned} (p_i - p_j)\left(\frac{\partial H(c \mid \mathbf{p})}{\partial p_i} - \frac{\partial H(c \mid \mathbf{p})}{\partial p_j}\right) \\ = -(p_i - p_j)^2 (f(c \mid \mathbf{p}^{ij}) - f(c-1 \mid \mathbf{p}^{ij})) . \end{aligned} \tag{3.4}$$

Now, Samuels [7] has shown (using a well-known inequality attributed to Newton) that if $f(k)$ is the probability of k successes in m independent Bernoulli trials, and if $\sum_{k=0}^{m} kf(k) = \tau$, then $f(k)$ is increasing in k for $k \leqq [\tau]$ and decreasing in k for $k \geqq [\tau + 1]$. Hence, using the characterization of $f(k \mid \mathbf{p}^{ij})$ given in (2.6), and noting that $\sum_{k=0}^{n-2} kf(k \mid \mathbf{p}^{ij}) = \lambda - p_i - p_j$, we have that (3.4) is nonnegative for $c \geqq [\lambda - p_i - p_j + 2]$ and nonpositive for $c \leqq [\lambda - p_i - p_j]$. Since $0 \leqq p_i + p_j \leqq 2$, all $i \neq j$, this result means that for all $\mathbf{p} \in D_\lambda$, (3.4) is $\leqq 0$ for $c \leqq [\lambda - 2]$ and $\geqq 0$ for $c \geqq [\lambda + 2]$. Thus, the bounds (1.12) and (1.13) in Theorem 1.1 follow by a direct application of Lemmas 2.1 and 2.2. □

REMARK 1. Hoeffding ([3] Theorem 4) also showed that for all $\mathbf{p} \in D_\lambda$,

$$0 \leqq H([\lambda - 1] \mid \mathbf{p}) \leqq H([\lambda - 1] \mid n^{-1}(\lambda, \lambda, \cdots, \lambda)) , \tag{3.5}$$

and

$$H([\lambda + 1] \mid n^{-1}(\lambda, \lambda, \cdots, \lambda)) \leqq H([\lambda + 1] \mid \mathbf{p}) \leqq 1 . \tag{3.6}$$

It might be thought that more detailed results for the cases $c = [\lambda - 1]$, $c = [\lambda + 1]$, similar to the results in Theorem 1.1, can be obtained. That is, we might suspect that $\mathbf{p}^{(1)} \in D_\lambda$, $\mathbf{p}^{(2)} = \mathbf{p}^{(1)}\Pi$ for doubly stochastic Π, implies that

$$H([\lambda - 1] \mid \mathbf{p}^{(1)}) \leqq H([\lambda - 1] \mid \mathbf{p}^{(2)}) , \tag{3.7}$$

$$H([\lambda + 1] \mid \mathbf{p}^{(1)}) \geqq H([\lambda + 1] \mid \mathbf{p}^{(2)}) . \tag{3.8}$$

The inequalities (3.7) and (3.8) do not, however, always hold. Inequality (3.7) holds if $\mathbf{p}^{(1)}$ is restricted to belong to the subset

$$D_\lambda^0 = \{\mathbf{p} : \mathbf{p} \in D_\lambda; [\lambda - 1] \leqq [\lambda - p_i - p_j] \leqq [\lambda + 1], \text{ all } i \neq j\}$$

of D_λ, as can be seen from the proof of Theorem 1.1. (Note: if $\mathbf{p}^{(1)} \in D_\lambda^0$ and $\mathbf{p}^{(2)} = \mathbf{p}^{(1)}\Pi$, Π doubly stochastic, then $\mathbf{p}^{(2)} \in D_\lambda^0$.) Similarly, inequality (3.8) holds if $\mathbf{p}^{(1)}$ is restricted to the subset

$$D_\lambda^1 = \{\mathbf{p} : \mathbf{p} \in D_\lambda;\ [\lambda] \leqq [\lambda - p_i - p_j] \leqq [\lambda + 1],\ \text{all}\ \ i \neq j\}$$

of D_λ.

That (3.7) does not hold in general can be seen by letting $n = 4$, $\mathbf{p}^{(1)} = (1, \frac{1}{2}, \frac{1}{4}, \frac{1}{4})$, and

$$\Pi = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{3.9}$$

Here, $\lambda = 2$, $\mathbf{p}^{(2)} = (\frac{3}{4}, \frac{3}{4}, \frac{1}{4}, \frac{1}{4})$, $[\lambda - 1] = 1$, and

$$H([\lambda - 1] \mid \mathbf{p}^{(1)}) = \tfrac{9}{32} > \tfrac{69}{256} = H([\lambda - 1] \mid \mathbf{p}^{(2)}) .$$

That (3.8) does not hold in general can be seen by letting $n = 4$, $\mathbf{p}^{(1)} = (\frac{1}{4}, 0, \frac{3}{4}, \frac{3}{4})$, and Π be as in (3.9). Here $\lambda = \frac{7}{4}$, $\mathbf{p}^{(2)} = (\frac{1}{8}, \frac{1}{8}, \frac{3}{4}, \frac{3}{4})$, $[\lambda + 1] = 2$ and

$$H([\lambda + 1] \mid \mathbf{p}^{(1)}) = \tfrac{55}{64} < \tfrac{883}{1024} = H([\lambda + 1] \mid \mathbf{p}^{(2)}) .$$

Since when λ is not an integer, Theorem 4 of [3] does not even show that $H([\lambda] \mid \mathbf{p})$ is bounded by the values of $H([\lambda] \mid \mathbf{p})$ for $\mathbf{p} = \mathbf{p}^*(\lambda)$ and $\mathbf{p} = \hat{\mathbf{p}}(\lambda)$, it is unlikely that an ordering between $H([\lambda] \mid \mathbf{p}^{(1)})$ and $H([\lambda] \mid \mathbf{p}^{(2)})$, for $\mathbf{p}^{(2)} = \mathbf{p}^{(1)}\Pi$ that always goes in the same direction for all $\mathbf{p}^{(1)} \in D_\lambda$, all doubly stochastic Π, can be demonstrated. Indeed, it is easy to find examples in which $H([\lambda] \mid \mathbf{p}^{(1)}) < H([\lambda] \mid \mathbf{p}^{(2)})$ for one choice of $\mathbf{p}^{(1)} \in D_\lambda$ and a doubly-stochastic matrix Π, and in which $H([\lambda] \mid \mathbf{p}^{(1)}) > H([\lambda] \mid \mathbf{p}^{(2)})$ for another choice of Π and $\mathbf{p}^{(1)} \in D_\lambda$. The case when λ is an integer is more interesting, since in this case Theorem 4 of [3] states that

$$H(\lambda \mid n^{-1}(\lambda, \lambda, \cdots, \lambda)) \leqq H(\lambda \mid \mathbf{p}) \leqq 1$$

for all $\mathbf{p} \in D_\lambda$. Thus, for any $\mathbf{p}^{(1)} \in D_\lambda$ it is possible to find a doubly stochastic matrix Π mapping $\mathbf{p}^{(1)}$ into $\mathbf{p}^{(2)} = n^{-1}(\lambda, \lambda, \cdots, \lambda)$ for which

$$H(\lambda \mid \mathbf{p}^{(2)}) \leqq H(\lambda \mid \mathbf{p}^{(1)}) ,$$

and unless $\mathbf{p}^{(1)} = n^{-1}(\lambda, \lambda, \cdots, \lambda)$, this inequality will be strict (see [3]). On the other hand, even in this special case it is unfortunately true that we can find a $\mathbf{p}^{(1)} \in D_\lambda$ and a doubly stochastic matrix Π such that

$$H(\lambda \mid \mathbf{p}^{(1)}\Pi) > H(\lambda \mid \mathbf{p}^{(1)}) .$$

For example, let $\mathbf{p}^{(1)} = (\frac{1}{2}, 0, \frac{3}{4}, \frac{3}{4})$, $n = 4$, and Π be given by (3.9). Then $\lambda = 2$, $\mathbf{p}^{(2)} = (\frac{1}{4}, \frac{1}{4}, \frac{3}{4}, \frac{3}{4})$, and

$$H(2 \mid \mathbf{p}^{(2)}) = \tfrac{187}{256} > \tfrac{184}{256} = H(2 \mid \mathbf{p}^{(1)}) .$$

REMARK 2. Hoeffding ([3] Theorem 5) also showed that if $0 \leqq b \leqq \lambda \leqq c \leqq n$, then for all $\mathbf{p} \in D_\lambda$,

$$\begin{aligned} &H(c \mid \mathbf{p}^*(\lambda)) - H(b - 1 \mid \mathbf{p}^*(\lambda)) \\ &\qquad \leqq P\{b \leqq S \leqq c\} = H(c \mid \mathbf{p}) - H(b - 1 \mid \mathbf{p}) \\ &\qquad \leqq 1 . \end{aligned} \tag{3.10}$$

Correspondingly, as a corollary to Theorem 1.1, we can establish the following result.

THEOREM 3.1. *Suppose* $0 \leqq b \leqq [\lambda - 1]$ *and* $[\lambda + 2] \leqq c \leqq n$. *Let* $\mathbf{p}^{(1)} \in D_\lambda$ *and let* $\mathbf{p}^{(2)} = \mathbf{p}^{(1)}\Pi$, *where* Π *is an* $n \times n$ *doubly stochastic matrix. Then*

$$H(c \mid \mathbf{p}^{(2)}) - H(b - 1 \mid \mathbf{p}^{(2)}) \leqq H(c \mid \mathbf{p}^{(1)}) - H(b - 1 \mid \mathbf{p}^{(1)}) . \tag{3.11}$$

REMARK 3. For the possible statistical applications of the results obtained in this paper, the reader is urged to read Section 5 of [3], and also the comments in [4] and [7].

Acknowledgments. I would like to thank S. M. Samuels, A. W. Marshall, and Prem S. Puri for some helpful discussions and comments. I would also like to thank the referees for their careful attention to my paper.

REFERENCES

[1] BECKENBACH, E. F. and BELLMAN, R. (1965). *Inequalities* (2nd rev. printing). Springer-Verlag, New York.

[2] HARDY, G. H., LITTLEWOOD, J. E., and PÓLYA, G. (1959). *Inequalities*, 2nd. ed. Cambridge Univ. Press.

[3] HOEFFDING, W. (1956). On the distribution of the number of successes in independent trials. *Ann. Math. Statist.* **27** 713–721.

[4] JOGDEO, K. and SAMUELS, S. M. (1968). Monotone convergence of binomial probabilities and a generalization of Ramanujan's equation. *Ann. Math. Statist.* **39** 1191–1195.

[5] KARLIN, S. J. and NOVIKOFF, A. (1963). Generalized convex inequalities. *Pacific J. Math.* **13** 1251–1279.

[6] KARLIN, S. J. and STUDDEN, W. J. (1966). *Tchebycheff Systems: With Applications in Analysis and Statistics.* Wiley, New York.

[7] SAMUELS, S. M. (1965). On the number of successes in independent trials. *Ann. Math. Statist.* **36** 1272–1278.

DEPARTMENT OF STATISTICS
MATHEMATICAL SCIENCES BUILDING
PURDUE UNIVERSITY
WEST LAFAYETTE, INDIANA 47907