On the Distribution of the Number of Successes in Independent Trials

by

Leon Jay Gleser
Purdue University

Department of Statistics
Division of Mathematical Sciences
Mimeograph Series #331
August 1973

---

*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.

SUMMARY

On the Distribution of the Number of Successes in Independent Trials

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

<u>AMS 1970 subject classification</u>. Primary 60C05, 60E05. Secondary 62E15, 26A86.

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

On the Distribution of the Number of Successes in Independent Trials

by

Leon Jay Gleser
Purdue University

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 j-th trial, $0 \le p_j \le 1$. Let

$$(1.1) \qquad \underset{\sim}{p} = (p_1, p_2, \ldots, p_n),$$

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

$$(1.2) \qquad H(c|\underset{\sim}{p}) = P\{S \le c\}.$$

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

$$(1.3) \qquad D_\lambda = \{\underset{\sim}{p}: \ 0 \le p_i \le 1, \ i = 1, 2, \ldots, n; \ \sum_{i=1}^{n} p_i = \lambda\}.$$

That is, $\underset{\sim}{p}^{(1)}$ and $\underset{\sim}{p}^{(2)}$ 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 [1; Theorem 4] has shown that for all $\underset{\sim}{p} \in D_\lambda$,

$$(1.4) \qquad 0 \le H(c|\underset{\sim}{p}) \le H(c|n^{-1}(\lambda, \lambda, \ldots, \lambda)), \text{ if } 0 \le c \le [\lambda - 2],$$

$$(1.5) \qquad H(c|n^{-1}(\lambda, \lambda, \ldots, \lambda)) \le H(c|\underset{\sim}{p}) \le 1, \text{ if } [\lambda + 2] \le c \le n,$$

where

$$(1.6) \qquad H(c|n^{-1}(\lambda,\lambda,\ldots,\lambda)) = \sum_{k=0}^{c} \binom{n}{k}\left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k},$$

and $[x]$ denotes the greatest integer $\leq x$. Hoeffding [1; Theorem 4] also obtained bounds on $H(c|\underset{\sim}{p})$ 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

$$(1.7) \qquad \underset{\sim}{p}^*(\lambda) = n^{-1}(\lambda,\lambda,\ldots,\lambda)$$

and

$$(1.8) \qquad \hat{\underset{\sim}{p}}(\lambda) = (1,1,\ldots,1,\lambda - [\lambda], 0, 0,\ldots,0),$$

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

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

$$(1.9) \qquad \underset{\sim}{p}^*(\lambda) = \underset{\sim}{p}\,\Pi^*,$$

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

$$(1.10) \qquad \underset{\sim}{p} = \hat{\underset{\sim}{p}}(\lambda) \quad \Pi(\underset{\sim}{p}),$$

where $\Pi(\underset{\sim}{p})$ is an nxn doubly stochastic matrix whose first $[\lambda + 1]$ rows are equal to $\underset{\sim}{p}$, and whose last $n - [\lambda + 1]$ rows are equal to $(n - [\lambda + 1])^{-1}(\underset{\sim}{1} - \lambda\underset{\sim}{p})$, where $\underset{\sim}{1} = (1,1,1\ldots,1)$ is the 1xn vector all of whose elements are ones. 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 $H(c|\underset{\sim}{p})$ as a function of $\underset{\sim}{p}$.

Theorem 1.1.

Let $p^{(1)} \in D_\lambda$ and suppose that there exists a doubly stochastic nxn matrix Π for which

$$(1.11) \qquad \underset{\sim}{p}^{(2)} = \underset{\sim}{p}^{(1)}\,\Pi.$$

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

$$(1.12) \qquad H(c|p^{(1)}) \leq H(c|p^{(2)}), \quad \text{if } 0 \leq c \leq [\lambda - 2],$$

and

$$(1.13) \qquad H(c|p^{(1)}) \geq H(c|p^{(2)}), \quad \text{if } [\lambda + 2] \leq c \leq n.$$

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

2. Majorization.

A 1xn vector $\underset{\sim}{x}$ is said to majorize a 1xn vector $\underset{\sim}{y}$ if $x_{[1]} \geq y_{[1]}$, $x_{[1]} + x_{[2]} \geq y_{[1]} + y_{[2]}, \ldots, \sum_{i=1}^{n-1} x_{[i]} \geq \sum_{i=1}^{n-1} y_{[i]}$, 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 $\underset{\sim}{x}$ and $\underset{\sim}{y}$, respectively, arranged in descending order ($x_{[1]} \geq x_{[2]} \geq \ldots \geq x_{[n]}$, and similarly for the $y_{[i]}$'s). The relation of majorization to doubly stochastic matrices is given by the following result of Karamata [1; Theorem 14].

Lemma 2.1.

The vector $\underset{\sim}{x}$ majorizes the vector $\underset{\sim}{y}$ if and only if there exists an n x n doubly stochastic matrix Π such that $\underset{\sim}{y} = \underset{\sim}{x}\,\Pi$.

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

Lemma 2.2.

Let $F(\underset{\sim}{z})$ be a function defined on n-dimensional vectors $\underset{\sim}{z} = (z_1, z_2, \ldots, z_n)$. For any i, j, $i \neq j$, and all $\underset{\sim}{z}$ in a region D, suppose that

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

whenever $z_i \geq z_j$. If $\underset{\sim}{x}, \underset{\sim}{y} \in D$, and if $\underset{\sim}{x}$ majorizes $\underset{\sim}{y}$, then

$$(2.2) \qquad F(\underset{\sim}{x}) \geq F(\underset{\sim}{y}).$$

A function satisfying (2.1) over a region D is said to satisfy a Schur condition on D.

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

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

for $\underset{\sim}{p} = (p_1, p_2, \ldots, p_n)$. Then $h(\underset{\sim}{p})$ is a function defined over the region

$$D = \{\underset{\sim}{z}: \quad 0 \leq z_i \leq 1, \ i = 1, 2, \ldots, n\}.$$

Lemma 2.3.

For any two components p_i and p_j, $i < j$, of $\underset{\sim}{p}$,

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

where for any function s(k) defined on the non-negative integers

$$\Delta s(k) \equiv s(k+2) - 2s(k+1) + s(k) \tag{2.5}$$

is the second difference of s(k), where $\underset{\sim}{p}^{ij}$ is the $1 \times (n-2)$ vector formed by deteting the i-th and j-th components of $\underset{\sim}{p}$, and where

$$f(k|\underset{\sim}{p}^{ij}) = \text{probability of } k \text{ successes in the } n-2 \text{ trials other than trials } i \text{ and } j; \tag{2.6}$$

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

Proof.

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

$$\begin{aligned} P\{S = k\} = (1-p_i)(1-p_j)\, f(k|\underset{\sim}{p}^{ij}) + (p_i + p_j - 2p_ip_j)\, f(k-1|\underset{\sim}{p}^{ij}) \\ + p_i\, p_j\, f(k-2|\underset{\sim}{p}^{ij}). \end{aligned} \tag{2.7}$$

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

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

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

As a corollary of Lemma 2.3, we can prove a result earlier obtained by Karlin and Novikoff [4].

Corollary 2.1.

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

(2.9) $\quad h(\underset{\sim}{p}^{(1)}) \leq h(\underset{\sim}{p}^{(2)})$.

Proof.

Since $\Delta g(k) \geq 0$, $k = 0, 1, \ldots, n - 2$, $-h(\underset{\sim}{p})$ satisfies a Schur condition, as can be seen from (2.4). Hence, Lemmas 2.1 and 2.2 imply that $-h(p^{(1)}) \geq -h(\underset{\sim}{p}^{(2)})$, from which (2.9) immediately follows. Q.E.D.

Karlin and Novikoff [4] 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 $\underset{\sim}{p} \in D_\lambda$ by doubly stochastic matrices to $\underset{\sim}{p}^*(\lambda)$ and $\hat{\underset{\sim}{p}}(\lambda)$, it follows that for any $g(k)$ convex on $0, 1, \ldots, n - 2$, and any $\underset{\sim}{p} \in D_\lambda$,

$$(1 - \delta)g([\lambda]) + \delta g([\lambda + 1])$$

$$\leq h(\underset{\sim}{p}) = Eg(S) \tag{2.10}$$

$$\leq \sum_{k=0}^{n} g(k) \binom{n}{k} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n - k},$$

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

The result (2.10) implies that $E|S - b|^a$, for any $a > o$ 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., $\underset{\sim}{p} = \underset{\sim}{p}^*(\lambda)$), and lowest when

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

(i.e., $\underset{\sim}{p} = \hat{\underset{\sim}{p}}(\lambda)$). The upper bound in (2.10) was first obtained (using a different method) by Hoeffding [2]. The lower bound in (2.10) can also be obtained by the methods in Hoeffding's [2] paper.

3. Proof of Theorem 1.1.

For fixed integer c, $0 \leq c \leq n$, let

$$g_c(k) = \begin{cases} 1, \text{ if } 0 \leq k \leq c \\ 0, \text{ if } c + 1 \leq k \leq n. \end{cases} \tag{3.1}$$

Then

$$(3.2) \qquad H(c|\underset{\sim}{p}) = E(g_c(S)).$$

Note that $g_c(k)$ is not convex on $0, 1,\ldots,n - 2$ when $c \le n - 1$, 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 $c \le n - 1$

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

Thus, from Lemma 2.3,

$$(3.4) \quad (p_i - p_j)\left(\frac{\partial H(c|\underset{\sim}{p})}{\partial p_i} - \frac{\partial H(c|\underset{\sim}{p})}{\partial p_j} \right) = -(p_i - p_j)^2 \left(f(c|\underset{\sim}{p}^{ij}) - f(c-1|\underset{\sim}{p}^{ij})\right).$$

Now, Samuels [6] 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 $\sum_{k=0}^{m} k\, f(k) = \tau$, then $f(k)$ is increasing in k for $k \le [\tau]$ and decreasing in k for $k \ge [\tau + 1]$. Hence, using the characterization of $f(k|\underset{\sim}{p}^{ij})$ given in (2.6), and noting that $\sum_{k=0}^{n-2} kf(k|\underset{\sim}{p}^{ij}) = \lambda - p_i - p_j$, we have that (3.4) is non-negative for $c \ge [\lambda - p_i - p_j + 2]$ and non-positive for $c \le [\lambda - p_i - p_j]$. Since $0 \le p_i + p_j \le 2$, all $i \ne j$, this result means that for all $\underset{\sim}{p} \in D_\lambda$, (3.4) is ≤ 0 for $c \le [\lambda - 2]$ and ≥ 0 for $c \ge [\lambda + 2]$. Thus, the bounds (1.12) and (1.13) in Theorem 1.1 follow by a direct application of Lemmas 2.1 and 2.2. Q.E.D.

<u>Remark I</u>.

Hoeffding [2; Theorem 4] also showed that for all $\underset{\sim}{p} \in D_\lambda$,

$$(3.5) \qquad 0 \le H([\lambda - 1]|\underset{\sim}{p}) \le H([\lambda - 1]|n^{-1}(\lambda, \lambda,\ldots,\lambda)),$$

and

$$(3.6) \qquad H([\lambda + 1]|n^{-1}(\lambda, \lambda,\ldots,\lambda)) \le H([\lambda + 1]|\underset{\sim}{p}) \le 1.$$

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 $\underset{\sim}{p}^{(1)} \in D_\lambda$, $\underset{\sim}{p}^{(2)} = \underset{\sim}{p}^{(1)}\Pi$ for doubly stochastic Π, implies that

$$(3.7) \qquad H([\lambda - 1]|\underset{\sim}{p}^{(1)}) \leq H([\lambda - 1]|\underset{\sim}{p}^{(2)}),$$

$$(3.8) \qquad H([\lambda + 1]|\underset{\sim}{p}^{(1)}) \geq H([\lambda + 1]|\underset{\sim}{p}^{(2)}).$$

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

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

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

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

of D_λ.

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

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

Here, $\lambda = 2$, $\underset{\sim}{p}^{(2)} = (3/4, 3/4, 1/4, 1/4)$, $[\lambda - 1] = 1$,

and

$$H([\lambda - 1]|\underset{\sim}{p}^{(1)}) = \frac{9}{32} > \frac{69}{256} = H([\lambda - 1]|\underset{\sim}{p}^{(2)}).$$

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

$$H([\lambda + 1]|\underset{\sim}{p}^{(1)}) = \frac{55}{64} < \frac{883}{1024} = H([\lambda + 1]|\underset{\sim}{p}^{(2)}).$$

Since Theorem 4 of [2] does not even show that $H([\lambda]|\underset{\sim}{p})$ is bounded by the values of $H([\lambda]|\underset{\sim}{p})$ for $\underset{\sim}{p} = \underset{\sim}{p}^*(\lambda)$ and $\underset{\sim}{p} = \hat{\underset{\sim}{p}}(\lambda)$, it is unlikely that an ordering between $H([\lambda]|\underset{\sim}{p}^{(1)})$ and $H([\lambda]|\underset{\sim}{p}^{(2)})$, for $\underset{\sim}{p}^{(2)} = \underset{\sim}{p}^{(1)}\Pi$, that always goes in the same direction for all $\underset{\sim}{p}^{(1)} \in D_\lambda$, all doubly stochastic Π, can be demonstrated. Indeed, it is easy to find examples in which $H([\lambda]|\underset{\sim}{p}^{(1)}) < H([\lambda]|\underset{\sim}{p}^{(2)})$, and examples in which $H([\lambda]|\underset{\sim}{p}^{(1)}) > H([\lambda]|\underset{\sim}{p}^{(2)})$.

Remark 2.

Hoeffding [2; Theorem 5] also showed that if $0 \le b \le \lambda \le c \le n$, then for all $\underset{\sim}{p} \in D_\lambda$,

$$\begin{aligned} &H(c|\underset{\sim}{p}^*(\lambda)) - H(b - 1|\underset{\sim}{p}^*(\lambda)) \\ (3.10)\qquad &\le P\{b \le S \le c\} = H(c|\underset{\sim}{p}) - H(b - 1|\underset{\sim}{p}) \\ &\le 1. \end{aligned}$$

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

Theorem 3.1.

Suppose $0 \le b \le [\lambda - 1]$ and $[\lambda + 2] \le c \le n$. Let $\underset{\sim}{p}^{(1)} \in D_\lambda$ and let $\underset{\sim}{p}^{(2)} = \underset{\sim}{p}^{(1)}\Pi$, where Π is an n x n doubly stochastic matrix. Then

$$(3.11)\qquad H(c|\underset{\sim}{p}^{(2)}) - H(b - 1|\underset{\sim}{p}^{(2)}) \le H(c|\underset{\sim}{p}^{(1)}) - H(b - 1|\underset{\sim}{p}^{(1)}).$$

Remark 3.

For the possible statistical applications of the results obtained in this paper, the reader is urged to read Section 5 of [2], and also the comments in [3] and [6].

Acknowledgements.

I would like to thank S. M. Samuels and Prem S. Puri for some helpful discussions.

BIBLIOGRAPHY

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

[2] Hoeffding, W. (1956). On the distribution of the number of successes in independent trials. Ann. Math. Statist. 27 713-721.

[3] 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.

[4] Karlin, S.J. and Novikoff, A. (1963). Generalized convex inequalities. Pacific J. Math. 13 1251-1279.

[5] Karlin, S.J. and Studden, W.J. (1966). Tchebycheff Systems: With Applications in Analysis and Statistics. Wiley, New York.

[6] Samuels, S.M. (1965). On the number of successes in independent trials. Ann. Math. Statist. 36 1272-1278.