BOUNDS FOR THE DISTRIBUTION FUNCTION OF A SUM OF INDEPENDENT, IDENTICALLY DISTRIBUTED RANDOM VARIABLES[1]

By Wassily Hoeffding and S. S. Shrikhande

University of North Carolina and College of Science, Nagpur, India

Summary. The problem is considered of obtaining bounds for the (cumulative) distribution function of the sum of n independent, identically distributed random variables with k prescribed moments and given range. For $n = 2$ it is shown that the best bounds are attained or arbitrarily closely approached with discrete random variables which take on at most $2k + 2$ values. For nonnegative random variables with given mean, explicit bounds are obtained when $n = 2$; for arbitrary values of n, bounds are given which are asymptotically best in the "tail" of the distribution. Some of the results contribute to the more general problem of obtaining bounds for the expected value of a given function of independent, identically distributed random variables when the expected values of certain functions of the individual variables are given. Although the results are modest in scope, the authors hope that this paper will draw attention to a problem of both mathematical and statistical interest.

1. Introduction. This paper considers part of the following general problem. Let $\mathfrak{D}$ be the class of all dfs (distribution functions) $F(x)$ on the real line which satisfy the conditions

$$\int g_i(x)\, dF(x) = c_i, \quad i = 1, \cdots, k; \qquad F(x) = \begin{cases} 0 & x < A, \\ 1 & x > B, \end{cases}$$

where the functions $g_1(x), \cdots, g_k(x)$ and the constants $c_1, \cdots, c_k$, A, and B are given. We allow that $A = -\infty$ and/or $B = \infty$. Here and in what follows, when the domain of integration is not indicated, the integral extends over the entire range of the variables involved.

Let $K(x_1, \cdots, x_n)$ be a function such that

$$\psi(F) = \int \cdots \int K(x_1, \cdots, x_n)\, dF(x_1) \cdots dF(x_n)$$

exists for all F in $\mathfrak{D}$ in the sense that the multiple integral is equal to the repeated integral taken in an arbitrary order. The problem is to determine upper and lower bounds for $\psi(F)$ when F is in $\mathfrak{D}$.

For $n = 1$, $g_i(x) = x^i$, and $K(x) = 1$ or 0 according as $x \leqq t$ or $> t$, as well

Received December 10, 1953.

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

as for other functions $K(x)$, an extensive literature on the subject exists; for some references see [4].

For n arbitrary, Robbins [6] showed that the Bienaymé-Tchebycheff bound for $\Pr(|X_1 + \cdots + X_n| \geqq t)$, where the X_i are independent and identically distributed with zero mean and given variance, can be improved when $n > 1$. Plackett [5], Gumbel [2], and Hartley and David [3] obtained the best possible bounds for the expected sample range and the expected value of the largest observation, in the case when the mean and the variance are given, assuming that the common df is continuous. In a problem analogous to the general problem stated above, but without the assumption that the n variables are identically distributed, one of the authors [4] showed that under general conditions the best bounds are attained or arbitrarily closely approached with step-functions in $\mathfrak{D}$ which have at most $k + 1$ steps.

The present paper concentrates attention on the case where $K = 1$ or 0 according as a given function $f(x_1, \cdots, x_n)$ is or is not contained in a given set. The method used permits one to obtain the closest bounds only for $n = 2$. If n is even, $f = x_1 + \cdots + x_n$, and $g_i(x) = x^i$, the bounds for $n = 2$ can be applied in an obvious way, but in general will not be the best ones. More general functions K are considered only insofar as they can be handled by the same method.

Theorem 2.1 states conditions under which we need consider only step-functions in $\mathfrak{D}$. Theorems 2.2 and 2.3 show that for functions $K(x, y)$ of a certain type we may restrict our attention to step-functions with a bounded number of steps. In Theorem 3.1 an explicit expression for the least upper bound of $\Pr(X + Y \geqq t)$ is obtained when X and Y are nonnegative, independent, and identically distributed with given mean. In Section 4 bounds for the analogous case with n summands are considered.

2. The least upper bound of $\int\int K(x, y)\, dF(x)\, dF(y)$. Let $K(x, y)$ be a function such that

$$\psi(F) = \iint K(x, y)\, dF(x)\, dF(y) \tag{2.1}$$

exists for all F in $\mathfrak{D}$, in the sense that the double integral equals the repeated integral. The problem is to determine the least upper bound of $\psi(F)$ for all F in $\mathfrak{D}$.

Let $\mathfrak{D}^*$ be the class of all F in $\mathfrak{D}$ which are step-functions with a finite number of steps. The following theorem shows that if $\mathfrak{D}$ is the class of dfs with k prescribed moments and given range, and $\psi(F)$ is the probability that two independent observations on a random variable with df F fall into a set of a rather general type, we may confine our attention to dfs in $\mathfrak{D}^*$.

THEOREM 2.1. *Let $g_i(x) = x^{m_i}$, where $m_1, \cdots, m_k$ are positive integers. Let $K(x, y) = 1$ or 0 according as (x, y) is or is not contained in a Borel set S such that the sets $\{x\colon (x, y) \in S, y \text{ fixed}\}$ and $\{y\colon (x, y) \in S, x \text{ fixed}\}$ are unions of a* *finite and bounded number of intervals (which may be infinite). Then*

$$\sup_{F \varepsilon \mathfrak{D}} \psi(F) = \sup_{F \varepsilon \mathfrak{D}^*} \psi(F).$$

The theorem follows immediately from an obvious analog of Lemma 2.1 in [4] and Lemma 3.1 and Theorem 4.1 of [4].

It can be seen from [4] that the reduction to distributions in $\mathfrak{D}^*$ is possible under more general conditions.

We shall now derive sufficient conditions under which, given a step-function F in $\mathfrak{D}$ with m steps, we can construct a step-function G in $\mathfrak{D}$ with less than m steps such that $\psi(G) \geqq \psi(F)$.

A step-function F in $\mathfrak{D}$ with exactly m steps is of the form

$$F(x) = P_j \quad \text{if} \quad a_j \leqq x < a_{j+1}, \qquad j = 0, 1, \cdots, m, \tag{2.2}$$

where

$$-\infty = a_0 < a_1 < a_2 < \cdots < a_m < a_{m+1} = \infty, \qquad A \leqq a_1, \qquad a_m \leqq B; \tag{2.3}$$

$$0 = P_0 < P_1 < \cdots < P_{m-1} < P_m = 1; \tag{2.4}$$

$$\sum_{j=1}^{m-1} h_{ij} P_j = c_i - g_i(a_m), \qquad i = 1, \cdots, k; \tag{2.5}$$

$$h_{ij} = g_i(a_j) - g_i(a_{j+1}), \qquad i = 1, \cdots, k; \; j = 1, \cdots, m-1. \tag{2.6}$$

Let

$$G(x) = P_j + t\,D_j \quad \text{if} \quad a_j \leqq x < a_{j+1}, \qquad j = 0, \cdots, m. \tag{2.7}$$

In order that $G(x)$ be a df in $\mathfrak{D}$ it is sufficient that the numbers t and D_j satisfy the conditions

$$D_0 = D_m = 0; \tag{2.8}$$

$$0 \leqq P_1 + t\,D_1 \leqq P_2 + t\,D_2 \leqq \cdots \leqq P_{m-1} + t\,D_{m-1} \leqq 1; \tag{2.9}$$

$$\sum_{j=1}^{m-1} h_{ij} D_j = 0, \qquad i = 1, \cdots, k. \tag{2.10}$$

If F and G are defined by (2.2) and (2.7), we have

$$\psi(G) - \psi(F) = t \sum_{j=1}^{m-1} L_j D_j + t^2 \sum_{i=1}^{m-1} \sum_{j=1}^{m-1} L_{ij} D_i D_j, \tag{2.11}$$

where, with $K_{ij} = K(a_i, a_j)$,

$$L_j = \sum_{i=1}^{m} (K_{ij} + K_{ji} - K_{i,j+1} - K_{j+1,i})(P_i - P_{i-1}), \tag{2.12}$$

$$L_{ij} = K_{ij} - K_{i+1,j} - K_{i,j+1} + K_{i+1,j+1}. \tag{2.13}$$

LEMMA 2.1. *Let F be a step-function in $\mathfrak{D}$ with exactly m steps, defined by* (2.2) *to* (2.6), *where* $m > k + 1$. *Suppose that the integers* $u_1, \cdots, u_{k+1}$ *can be so chosen that*

$$1 \leqq u_1 < u_2 < \cdots < u_{k+1} \leqq m - 1$$

and the equations

$$\sum_{r=1}^{k+1} h_{iu_r} x_r = 0, \qquad i = 1, \cdots, k, \tag{2.14}$$

imply

$$\sum_{r=1}^{k+1} \sum_{s=1}^{k+1} L_{u_r u_s} x_r x_s \geqq 0. \tag{2.15}$$

Then there exists a step-function G *in* $\mathfrak{D}$ *with less than* m *steps, for which* $\psi(G) \geqq \psi(F)$.

Proof. Let $G(x)$ be defined by (2.7), and let $D_j = 0$ for $j \neq u_1, \cdots, u_{k+1}$. Let $\lambda = 1$ or 0 according as the rank of the matrix

$$\begin{Vmatrix} h_{1u_1} & \cdots & h_{1u_{k+1}} \\ \cdots & \cdots & \cdots \\ h_{ku_1} & \cdots & k_{ku_{k+1}} \\ L_{u_1} & \cdots & L_{u_{k+1}} \end{Vmatrix}$$

is equal to or less than $k + 1$. Then the equations (2.14) and

$$\sum_{r=1}^{k+1} L_{u_r} x_r = \lambda$$

have a solution $(D_{u_1}, \cdots, D_{u_{k+1}}) \neq (0, \cdots, 0)$. Having thus fixed the D_j, let t be the largest number which satisfies the inequalities (2.9). This number exists and is positive. With this choice of the numbers t and D_j, G is a step-function in $\mathfrak{D}$ with less than m steps. Furthermore, by (2.11),

$$\psi(G) - \psi(F) = t\lambda + t^2 \sum_{r=1}^{k+1} \sum_{s=1}^{k+1} L_{u_r u_s} D_{u_r} D_{u_s} \geqq 0.$$

The proof is complete.

The next theorem shows that if $K(x, y)$ is of a certain form, and if we restrict ourselves to the class $\mathfrak{D}^*$ of step-functions in $\mathfrak{D}$ with a finite number of steps, we need consider only step-functions with a bounded number of steps.

Let $\mathfrak{D}_m$ be the class of all F in $\mathfrak{D}$ which are step-functions with at most m steps.

Theorem 2.2. *Suppose that* $K(x, y)$ *is of the form*

$$K(x, y) = \sum_{i=0}^{k} \sum_{j=0}^{k} a_{tij} g_i(x) g_j(y) \qquad \text{if } b_{t-1} \leqq f(x, y) < b_t, \qquad t = 1, \cdots, s,$$

where $g_0(x) = 1$, *the* a_{tij} *are arbitrary constants, the* b_t *satisfy*

$$-\infty = b_0 < b_1 < \cdots < b_{s-1} < b_s = \infty,$$

and $f(x, y)$ is a strictly increasing function in each of its arguments when the other argument is fixed. Then

$$\sup_{F \epsilon \mathfrak{D}^*} \psi(F) = \sup_{F \epsilon \mathfrak{D}_{sk+s}} \psi(F).$$

The theorem remains true if in the inequalities $b_{t-1} \leqq f(x, y) < b_t$ some signs $\leqq$ are replaced by $<$ or vice versa, provided that the s sets defined by the inequalities cover the entire plane.

PROOF. Let $F(x)$, as defined by (2.2) to (2.6), be an arbitrary step-function in $\mathfrak{D}$ with exactly m steps, where $m > sk + s$. It is sufficient to construct a step-function G in $\mathfrak{D}$ with less than m steps such that $\psi(G) \geqq \psi(F)$. Let m_t, for $t = 1, \cdots, s$, denote the number of indices u, with $1 \leqq u \leqq m$, for which

$$b_{t-1} \leqq f(a_u, a_u) < b_t .$$

Then $s \max (m_t) \geqq (m_1 + \cdots + m_s) = m > s(k + 1)$. Hence there exists a t for which $m_t \geqq k + 2$ and an integer n such that

$$b_{t-1} \leqq f(a_n, a_n) < f(a_{n+k+1}, a_{n+k+1}) < b_t .$$

The assumption about $f(x, y)$ implies that

$$K_{vw} = \sum_{i=0}^{k} \sum_{j=0}^{k} a_{tij}\, g_i(a_v) g_j(a_w) \qquad n \leqq v, w \leqq n + k + 1.$$

By (2.13) and (2.6) this implies

$$L_{vw} = \sum_{i=1}^{k} \sum_{j=1}^{k} a_{tij}\, h_{iu}\, h_{jw} \qquad n \leqq v, w \leqq n + k.$$

Hence if we let $u_r = n + r - 1$ for $r = 1, 2, \cdots, k + 1$, the conditions of Lemma 2.1 are satisfied. The proof is complete.

If $g_i(x) = x^i$, that is, if $\mathfrak{D}$ is the class of distributions with given moments up to order k and given range, the assumption of Theorem 2.2 means that $K(x, y)$ is piecewise polynomial, of bounded degrees, in sections of the plane separated by curves of *negative* slope. If $K(x, y)$ is piecewise polynomial in sections separated by curves of *positive* slope, a similar reduction of the problem to the case of step-functions with a bounded number of steps is in general impossible. For example, let $K(x, y) = \max (x, y)$, and let $\mathfrak{D}$ be the class of dfs F with

$$\int x\, dF(x) = 0, \qquad \int x^2\, dF(x) = 1.$$

Under the restriction to continuous functions $F(x)$, this is a special case of a problem considered by Hartley and David [3] and Gumbel [2]. For an arbitrary df $F(x)$ we can write

$$\psi(F) = 2 \int x\bar{F}(x)\, dF(x), \qquad \bar{F}(x) = \tfrac{1}{2}[F(x - 0) + F(x + 0)].$$

Using Schwarz's inequality, we have for any constant c and any F in $\mathfrak{D}$

$$\psi(F) + c = 2\int (x + c)\bar{F}(x)\, dF(x) \leqq 2\left((1 + c^2)\int \bar{F}(x)^2\, dF(x)\right)^{1/2}$$

If $F(x)$ is continuous, $\int\bar{F}(x)^2\, dF(x) = \frac{1}{3}$, and the bound

$$\psi(F) \leqq \min_c \{2\cdot 3^{-\frac{1}{2}}(1 + c^2)^{\frac{1}{2}} - c\}$$

is attained with a continuous df in $\mathfrak{D}$, as shown by Hartley and David.

Now let $F(x)$ be a step-function with at most m steps which takes on the values $0 = P_0 \leqq P_1 \leqq \cdots \leqq P_{m-1} \leqq P_m = 1$. Then

$$4\int \bar{F}(x)^2\, dF(x) = \sum_{j=1}^{m} (P_{j-1} + P_j)^2(P_j - P_{j-1}).$$

This can be written

$$12\int \bar{F}(x)^2\, dF(x) = 4 - \sum_{j=1}^{m} p_j^3, \qquad p_j = P_j - P_{j-1}.$$

The conditions $\sum p_j = 1$ and $p_j \geqq 0$ imply $\sum p_j^3 \geqq m^{-2}$. Hence

$$\int \bar{F}(x)^2\, dF(x) \leqq \tfrac{1}{3} - 1/12m^2,$$

and the Hartley-David bound cannot be approached arbitrarily closely with a step-function in $\mathfrak{D}$ having a bounded number of steps.

Combining Theorems 2.1 and 2.2 we can state that if the conditions of both theorems are satisfied, then

$$\sup_{F\varepsilon\mathfrak{D}} \psi(F) = \sup_{F\varepsilon\mathfrak{D}_{sk+s}} \psi(F).$$

In particular, the conditions of Theorem 2.2 are fulfilled if $\psi(F) = P_F\{f(X, Y) \geqq c\}$, or $= P_F\{\,|f(X, Y)| \geqq c\}$, etc., where $P_F\{\cdots\}$ is the probability of the event in braces when X and Y are independent with common df F, and $f(x, y)$ has the property stated in the theorem. Using Theorem 2.1, we obtain:

THEOREM 2.3. *Let $\mathfrak{D}$ be the class of* dfs *$F(x)$ which satisfy the conditions*

$$\int x^{m_i}\, dF(x) = c_i, \quad i = 1, \cdots, k, \qquad F(x) = \begin{cases} 0 & x < A, \\ 1 & x > B, \end{cases}$$

with given integers $m_1, \cdots, m_k$ and given numbers $c_1, \cdots, c_k, A, B$, where we may have $A = -\infty$ and/or $B = \infty$. Let $f(x, y)$ be a strictly increasing function in each of its arguments when the other argument is fixed. Then

$$\sup_{F\varepsilon\mathfrak{D}} P_F\{f(X, Y) \geqq c\} = \sup_{F\varepsilon\mathfrak{D}_{2k+2}} P_F\{f(X, Y) \geqq c\}.$$

3. The least upper bound of $P(X + Y \geqq t)$ when X and Y are nonnegative, independent, and identically distributed with given mean. As an application of the results of Section 2 we shall prove the following theorem.

THEOREM 3.1. *Let X and Y be two independent random variables with common* cdf $F(x)$. *Let $\mathfrak{D}$ be the class of* dfs F *with $F(x) = 0$ for $x < 0$ and $\int x \, dF(x) = \mu$, where $\mu > 0$. Then*

$$\sup_{F \varepsilon \mathfrak{D}} P_F(X + Y \geqq c\mu) = \begin{cases} 1, & c \leqq 2; \\ 4/c^2, & 2 \leqq c \leqq \frac{5}{2}; \\ 2/c - 1/c^2, & \frac{5}{2} \leqq c. \end{cases} \tag{3.1}$$

The three bounds are attained with the respective distributions

$$P(X = \mu) = 1;$$

$$P(X = 0) = 1 - 2/c, \qquad P(X = \tfrac{1}{2}c\mu) = 2/c;$$

$$P(X = 0) = 1 - 1/c, \qquad P(X = c\mu) = 1/c.$$

Theorem 3.1 should be compared with the solution by Birnbaum, Raymond, and Zuckerman [1] of the analogous problem without the restriction that X and Y be identically distributed. If $M(c)$ denotes the least upper bound of $P(X + Y \geqq c\mu)$ when X and Y are nonnegative, independent, and have the common mean μ, we have by [1]

$$M(c) = \begin{cases} 1, & c \leqq 2; \\ 1/(c - 1), & 2 \leqq c \leqq \frac{1}{2}(3 + \sqrt{5}); \\ 2/c - 1/c^2, & \frac{1}{2}(3 + \sqrt{5}) \leqq c. \end{cases} \tag{3.2}$$

Hence the bound (3.1) is smaller than the Birnbaum-Raymond-Zuckerman bound if and only if $2 < c < \frac{1}{2}(3 + \sqrt{5})$.

PROOF OF THEOREM 3.1. We may and shall assume that $\mu = 1$. By Theorem 2.3 we need consider only dfs F in $\mathfrak{D}$ which are step-functions with $m \leqq 4$ steps. Then F is of the form (2.2) to (2.4), where $A = 0$ and $B = \infty$, and $\sum_1^m a_j(P_j - P_{j-1}) = 1$. We have

$$K(x, y) = \begin{cases} 1, & x + y \geqq c; \\ 0, & x + y < c. \end{cases}$$

Hence the numbers $K_{ij} = K(a_i, a_j)$ satisfy the conditions

$$K_{ij} = 0 \text{ or } 1; \qquad K_{ij} = K_{ji}; \qquad K_{ij} \leqq K_{i'j} \quad \text{if} \quad i < i'.$$

The sequence $(K_{11}, K_{22}, \cdots, K_{mm})$ consists of a sequence of zeros followed by a sequency of ones. The reasoning used in the proof of Theorem 2.2 shows that any distribution for which there are more than two consecutive zeros or more than two consecutive ones in this sequence can be replaced by a distribution with less than m steps which does not decrease the value of $\psi(F)$.

Hence for $m = 4$ we need consider only matrices $\|K_{ij}\|$ of the four types

$$\begin{Vmatrix} 0 & 0 & 0 & \cdot \\ 0 & 0 & 0 & \cdot \\ 0 & 0 & 1 & 1 \\ & & 1 & 1 \end{Vmatrix}, \quad \begin{Vmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{Vmatrix}, \quad \begin{Vmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{Vmatrix}, \quad \begin{Vmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{Vmatrix},$$

where the numbers represented by dots need not be specified. The corresponding matrices $\|L_{ij}\|$ are

$$\text{I} \quad \begin{Vmatrix} 0 & 0 & \cdot \\ 0 & 1 & \cdot \\ \cdot & \cdot & 0 \end{Vmatrix}, \quad \text{II} \quad \begin{Vmatrix} 0 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 0 \end{Vmatrix}, \quad \text{III} \quad \begin{Vmatrix} 0 & 1 & -1 \\ 1 & -1 & 0 \\ -1 & 0 & 0 \end{Vmatrix}, \quad \text{IV} \quad \begin{Vmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{Vmatrix}$$

We shall apply Lemma 2.1 to show that in every case there exists a df in $\mathfrak{D}$ with at most three steps which does not decrease the value of $\psi(F)$. It is sufficient to find integers u and v with $1 \leqq u < v \leqq 3$ such that the equation

$$(a_u - a_{u+1})x + (a_v - a_{v+1})y = 0 \tag{3.3}$$

implies

$$L_{uu}x^2 + 2L_{uv}xy + L_{vv}y^2 \geqq 0. \tag{3.4}$$

Inequality (3.4) is satisfied in Case I with $u = 1$ and $v = 2$, and in Cases II and IV with $u = 1$ and $v = 3$. In Case III, when $u = 1$ and $v = 3$, the left side of (3.4) is $-2xy$, which is nonnegative by (3.3), since $a_j - a_{j+1} < 0$.

Hence we may confine our attention to step-functions in $\mathfrak{D}$ with $m \leqq 3$ steps.

If $m = 3$, we have to consider the matrices $\|K_{ij}\|$ of the forms

$$\begin{Vmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{Vmatrix}, \quad \begin{Vmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \end{Vmatrix}, \quad \begin{Vmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{Vmatrix},$$

$$\begin{Vmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{Vmatrix}, \quad \begin{Vmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{Vmatrix}, \quad \begin{Vmatrix} 0 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{Vmatrix}$$

The corresponding matrices $\|L_{ij}\|$ are

$$A \begin{Vmatrix} 0 & 0 \\ 0 & 1 \end{Vmatrix}, \quad B \begin{Vmatrix} 0 & 1 \\ 1 & -1 \end{Vmatrix}, \quad C \begin{Vmatrix} 0 & 0 \\ 0 & -1 \end{Vmatrix}, \quad D \begin{Vmatrix} 1 & 0 \\ 0 & 0 \end{Vmatrix}, \quad E \begin{Vmatrix} 1 & -1 \\ -1 & 0 \end{Vmatrix}, \quad F \begin{Vmatrix} -1 & 0 \\ 0 & 0 \end{Vmatrix}$$

In applying Lemma 2.1 we have to take $u = 1$ and $v = 2$ and show that (3.3) implies (3.4). This is true for the matrices A, D, and E. In Cases B, C, and F, Lemma 2.1 is not applicable.

In Case C, $\psi(F) = 1 - P_2^2$. If $G(x)$ is defined by (2.7) with $m = 3$, we have

$$\psi(G) - \psi(F) = -t\,D_2[2(P_2 + t\,D_2) - t\,D_2],$$

where t and D_2 satisfy the conditions

$$(a_1 - a_2)\,D_1 + (a_2 - a_3)\,D_2 = 0, \tag{3.5}$$

$$0 \leqq P_1 + t\,D_1 \leqq P_2 + t\,D_2 \leqq 1. \tag{3.6}$$

Let $D_2 = -1$. Then D_1 is given by (3.5). Let t be the largest number which satisfies (3.6). Then $t > 0$, G is in $\mathfrak{D}_2$, and $\psi(G) - \psi(F) \geqq 0$.

In case F, $\psi(F) = 1 - P_1^2$, and similar reasoning shows that this case also can be reduced to a step-function with at most two steps.

The only remaining case with $m = 3$ is Case B. Here we can write $\psi(F) = 2p_2p_3 + p_3^2$, where (admitting the possibility that F has less than three steps)

$$(3.7) \qquad p_1 + p_2 + p_3 = 1, \qquad a_1p_1 + a_2p_2 + a_3p_3 = 1;$$

$$(3.8) \qquad p_1 \geqq 0, \qquad p_2 \geqq 0, \qquad p_3 \geqq 0;$$

$$(3.9) \qquad 0 \leqq a_1 \leqq a_2 \leqq a_3 ;$$

$$(3.10) \qquad a_1 + a_3 < c, \qquad 2a_2 < c, \qquad c \leqq a_2 + a_3 .$$

Expressing $\psi(F)$ in terms of a_1, a_2, a_3, p_1, we get

$$\psi(F) = (1 - p_1)^2 - p_2^2, \qquad p_2 = \frac{a_3 - 1 - (a_3 - a_1)p_1}{a_3 - a_2}.$$

If a_2, a_3, and p_1 are held fixed, $\psi(F)$ is a decreasing function of a_1. Hence we maximize $\psi(F)$ by choosing the least possible value for a_1. This is the greatest of the bounds given by the inequalities $p_j \geqq 0$ and $a_1 \geqq 0$. If this bound is given by one of the equations $p_j = 0$, we get a distribution in $\mathfrak{D}_2$. Hence we may assume that the least value is $a_1 = 0$, so

$$p_2 = \frac{a_3 - 1 - a_3p_1}{a_3 - a_2}$$

and $\psi(F)$ is a decreasing function of a_2 when a_3 and p_1 are fixed. The only lower bound for a_2 which does not necessarily correspond to a distribution in $\mathfrak{D}_2$ is $a_2 = c - a_3$. In this case

$$p_2 = \frac{(1 - p_1)a_3 - 1}{2a_3 - c} = \frac{1 - p_1}{2} + \frac{c(1 - p_1) - 2}{2(2a_3 - c)}.$$

This is a monotonic function of a_3 (possibly a constant) when p_1 is held fixed. Hence the maximum is attained at one of the endpoints of the range of a_3. This range is given by the inequalities (3.8) to (3.10) with $a_1 = 0$ and $a_2 = c - a_3$. Its endpoints correspond either to distributions in $\mathfrak{D}_2$ or (if given by $a_1 + a_3 = c$ or $2a_2 = c$) to cases where the value of $\psi(F)$ exceeds $2p_2p_3 + p_3^2$ and which already have been disposed of.

Thus we need consider only dfs in $\mathfrak{D}_2$.

If $c \leqq 2$, we have $\psi(F) = 1$ for the df in $\mathfrak{D}_1$ which has a single step at $x = 1$. Thus

$$(3.11) \qquad \sup_{F \in \mathfrak{D}} \psi(F) = 1 \qquad c \leqq 2.$$

Henceforth we assume that $c > 2$.

A distribution F in $\mathfrak{D}_2$ assigns to the points a_1 and a_2 the respective probabilities

$$p_1 = \frac{a_2 - 1}{a_2 - a_1}, \qquad p_2 = \frac{1 - a_1}{a_2 - a_1}, \qquad 0 \leqq a_1 \leqq 1 \leqq a_2.$$

If $c \leqq 2a_1$, we have $c \leqq 2$, a case already considered. If $c > 2a_2$, then $\psi(F) = 0$, a case which may be disregarded. We are left with the two cases

$$\text{(i)} \quad 2a_1 < c \leqq a_1 + a_2, \qquad \text{(ii)} \quad a_1 + a_2 < c \leqq 2a_2.$$

In Case (i), $\psi(F) = 1 - p_1^2$, which is a decreasing function of a_1. The lower bound for a_1 is max $(0, \quad c - a_2)$.

If $a_1 = 0 \geqq c - a_2$, then $p_1 = 1 - 1/a_2$, so that $\psi(F)$ is a decreasing function of a_2. The lower bound for a_2 is max $(1, c) = c$, and we obtain

$$\psi(F) = 1 - (1 - 1/c)^2 = 2/c - 1/c^2.$$

If $a_1 = c - a_2 \geqq 0$, then

$$p_1 = \frac{a_2 - 1}{2a_2 - c} = \frac{1}{2} + \frac{c - 2}{2(2a_2 - c)},$$

so that $\psi(F)$ is an increasing function of a_2. Since $a_2 \leqq c$, we obtain the same maximum of $\psi(F)$ as in the previous case.

In Case (ii), $\psi(F) = p_2^2$, which is a decreasing function of a_2. Hence we let $a_2 = \frac{1}{2}c$. Then $\psi(F)$ is a decreasing function of a_1, and hence is maximized for $a_1 = 0$. We get $\psi(F) = 4/c^2$. Hence

$$\sup_{F \in \mathfrak{D}} \psi(F) = \max \{2/c - 1/c^2, \ 4/c^2\}, \qquad c > 2. \tag{3.12}$$

Theorem 3.1 now follows from (3.11), (3.12) and the stated conditions under which the bounds are attained.

4. Bounds for $P(X_1 + \cdots + X_n \geqq c)$.

Let $\bar{X}_n = n^{-1}(X_1 + \cdots + X_n)$, and let $\omega_n(t)$ denote the least upper bound of $P(\bar{X}_n \geqq t\mu)$ when $X_1, \cdots, X_n$ are nonnegative, independent, and identically distributed with mean μ. It is easily seen that for every n

$$\omega_n(t) = 1, \quad \text{if } t \leqq 1; \qquad \omega_{sn}(t) \leqq \omega_s(t), \quad s = 1, 2, \cdots.$$

By Markov's inequality, $\omega_1(t) = 1/t$ if $1 \leqq t$. By Theorem 3.1,

$$\omega_2(t) = \begin{cases} 1/t^2 & 1 \leqq t \leqq 5/4, \\ 1/t - 1/4t^2 & 5/4 \leqq t. \end{cases}$$

Let $\omega_n^*(t)$ be the least upper bound of $P(\bar{X}_n \geqq t\mu)$ when $X_1, \cdots, X_n$ are independent and nonnegative with common mean μ. Clearly, $\omega_n(t) \leqq \omega_n^*(t)$. From [1] (in particular, Corollary 2.2) we have

$$\omega_n^*(t) \leqq \begin{cases} \dfrac{1}{t} - \dfrac{1}{4t^2} & \dfrac{3 + \sqrt{5}}{4} \leqq t; \quad n \text{ even}; \\[2ex] \dfrac{1}{t} - \dfrac{n^2 - 1}{n^2}\dfrac{1}{4t^2} & \dfrac{3n + 1 + (5n^2 + 6n + 5)^{1/2}}{4n} \leqq t; \quad n \text{ arbitrary}. \end{cases}$$

On the other hand, for any random variables X_i which satisfy our assumptions, $P(\bar{X}_n \geqq t\mu)$ is a lower bound for $\omega_n(t)$. In particular, if $nt \geqq 1$ and $X_i = 0$ or nt with respective probabilities $1 - 1/nt$ and $1/nt$, we get

$$\omega_n(t) \geqq 1 - \left(1 - \frac{1}{nt}\right)^n > \frac{1}{t} - \frac{n-t}{2n}\frac{1}{t^2}.$$

Hence we have for all positive integers n

$$(4.1) \quad \omega_n(t) = \frac{1}{t} - \frac{1+\theta}{4}\frac{n-1}{n}\frac{1}{t^2} \quad \text{if} \quad \frac{3n+1+(5n^2+6n+5)^{1/2}}{4n} \leqq t, \quad \frac{1}{n} \leqq \theta < 1;$$

$$(4.2) \quad \omega_{2n}(t) = \frac{1}{t} - \frac{1+\theta'}{4}\frac{1}{t^2} \quad \text{if} \quad \frac{5}{4} \leqq t, \qquad 0 \leqq \theta' < 1 - \frac{2}{n}.$$

Equation (4.1) is also true for $\omega_n^*(t)$, and (4.2) holds for $\omega_{2n}^*(t)$ if $\frac{1}{4}(3 + \sqrt{5}) \leqq t$. Thus for large values of t the known bounds for $\omega_n(t)$ and $\omega_n^*(t)$ cannot be improved substantially.

REFERENCES

[1] Z. W. Birnbaum, J. Raymond, and H. S. Zuckerman, "A generalization of Tchebychev's inequality to two dimensions", *Ann. Math. Stat.*, Vol. 18 (1947), pp. 70–79.

[2] E. J. Gumbel, "The maxima of the mean largest value and of the range", *Ann. Math. Stat.*, Vol. 25 (1954), pp. 76–84.

[3] H. O. Hartley and H. A. David, "Universal bounds for mean range and extreme observation", *Ann. Math. Stat.*, Vol. 25 (1954), pp. 85–99.

[4] W. Hoeffding, "The extrema of the expected value of a function of independent random variables", *Ann. Math. Stat.*, Vol. 26 (1955), pp. 268–275.

[5] R. L. Plackett, "Limits of the ratio of mean range to standard deviation", *Biometrika*, Vol. 34 (1947), pp. 120–122.

[6] H. Robbins, "Some remarks on the inequality of Tchebycheff", *Courant Anniversary Volume*, 1948, pp. 345–350.