ON THE DEVIATIONS OF THE EMPIRIC DISTRIBUTION FUNCTION OF VECTOR CHANCE VARIABLES

BY

J. KIEFER[1] AND J. WOLFOWITZ[2]

1. **Introduction.** Let F be a distribution function (d.f.) on Euclicean m-space, and let $X_1, X_2, \cdots, X_n$, be independent chance variables with the common d.f. F. The empiric d.f. S_n is a chance d.f. defined for any $x = (x_1, \cdots, x_m)$ as follows: $nS_n(x)$ is the number of X_i's, $i = 1, \cdots, n$, such that, for $j = 1, \cdots, m$, the jth component $X_i^{(j)}$ of X_i is less than x_j. Define

$$D_n = \sup_x |S_n(x) - F(x)|,$$

$$D_n^+ = \sup_x (S_n(x) - F(x)),$$

$$D_n^- = \sup_x (F(x) - S_n(x)),$$

and

$$G_n(r) = P\{D_n < r/n^{1/2}\},$$
$$G_n^+(r) = P\{D_n^+ < r/n^{1/2}\},$$
$$G_n^-(r) = P\{D_n^- < r/n^{1/2}\}.$$

In this paper we prove two theorems, of which the first is the following:

THEOREM 1. *For each m there exist positive constants c_0 and c such that, for all n, all F, and all positive r,*

(1.1) $$1 - G_n(r) < c_0 e^{-cr^2},$$

(1.2) $$1 - G_n^+(r) < c_0 e^{-cr^2},$$

(1.3) $$1 - G_n^-(r) < c_0 e^{-cr^2}.$$

The nub of the theorem is, of course, that it sets a minimum rate at which $G_n(r)$, $G_n^+(r)$, and $G_n^-(r)$ go to one as $r \to \infty$, independent of n and F. It is rather curious that a bound independent of F can be given, since the limits of G_n, G_n^+, and G_n^- (as $n \to \infty$) depend on F for $m > 1$. The limits of G_n, G_n^+, and G_n^- for $m = 1$ are of course known [1] and independent of F when F is continuous. The limits for $m > 1$ are at present writing unknown.

Theorem 1 for $m=1$ was proved in [2] by a method which took as its point of departure an exact expression for $G_n^+(r)$ due to Smirnov [3]. No such formula is known for the case $m>1$. The method of the present paper is entirely different and does not use the result of [2][3]. The extension of the result from $m=1$ to $m=2$ presents difficulties; the extension of the result for $m=2$ to larger values of m by our method of proof is obvious, and proceeds by induction on m. Theorem 1 is used in proving Theorem 2.

The constants c_0 and c in general depend upon m. We make no attempt in this paper to obtain the best possible constants or even to perform some tedious calculations which would improve them. At the end of the proof of Theorem 1 we calculate possible values of c and give some suggestions for improving the constants (it is shown in [2] that 2 is the best value of c for $m=1$; we also show at the end of the proof of Theorem 1 that $c<2$ for $m>1$). We also point out that the supremum operation can be performed over a larger class of sets without affecting the result.

Before stating Theorem 2 we introduce some additional notation. For fixed F and positive integral k, write A_k^m for the subset consisting of every point in Euclidean m-space for which, for $1\leq j\leq m$, the jth coordinate w_j satifies $F_j(w_j)\leq h_j/k\leq F_j(w_j+0)$ for any integers h_j, where F_j is the (marginal) d.f. of $X_i^{(j)}$. Write

$$D_{n,k} = \max_{x\in A_k^m} \left| S_n(x) - F(x) \right|.$$

$$D_{n,k}^+ = \max_{x\in A_k^m} \left[S_n(x) - F(x) \right].$$

$$D_{n,k}^- = \max_{x\in A_k^m} \left[F(x) - S_n(x) \right].$$

as well as

$$G_{n,k}(r) = P\{ D_{n,k} < r/n^{1/2} \},$$

$$G_{n,k}^+(r) = P\{ D_{n,k}^+ < r/n^{1/2} \},$$

$$G_{n,k}^-(r) = P\{ D_{n,k}^- < r/n^{1/2} \},$$

and

$$H_n(r, r') = P\{ D_n^+ < r/n^{1/2},\ D_n^- < r'/n^{1/2} \},$$

$$H_{n,k}(r, r') = P\{ D_{n,k}^+ < r/n^{1/2},\ D_{n,k}^- < r'/n^{1/2} \}.$$

We shall also denote by

$$G_{\infty,k}, \qquad G_{\infty,k}^+, \qquad G_{\infty,k}^-, \qquad H_{\infty,k}$$

_____

[3] In a first, unpublished, version of [2], a weaker result than that mentioned below as appearing in [2] for the case $m=1$, was proved by a method which has points in common with the present proof of Theorem 1; one idea used in this method is due to P. Erdös.

the respective limits as $n \to \infty$ of the d.f.'s $G_{n,k}$, $G_{n,k}^+$, $G_{n,k}^-$, and $H_{n,k}$; the existence of these limits is a consequence of the multivariate central limit theorem.

Our second result is

THEOREM 2. *For every m and F, there exists a d.f. G (resp., G^+, G^-, H) such that the sequence of d.f.'s G_n (resp., G_n^+, G_n^-, H_n) converges to G (resp., G^+, G^-, H) at every continuity point of the latter as $n \to \infty$ and such that the sequence of d.f.'s $G_{\infty,k}$ (resp., $G_{\infty,k}^+$, $G_{\infty,k}^-$, $H_{\infty,k}$) converges to this d.f. G (resp., G^+, G^-, H) at every continuity point of the latter as $k \to \infty$.*

It is obvious that G, G^+, G^-, and H cannot be degenerate unless F is. Of course, as noted above, these d.f.'s depend on F.

Theorem 2 generalizes the result of Donsker [4] for the case $m = 1$; we remark that our proof starts ab initio and does not make use of Donsker's result or method. Donsker's result is needed to justify Doob's [5] computation of G, G^+, G^-, and H in the case $m = 1$, and Theorem 2 could perhaps prove of similar use in the more difficult problem of computing these limiting d.f.'s when $m > 1$ if this is to be done by consideration of a Gaussian process (depending on F) with m-dimensional time. Donsker's result was also used in [2] in the case $m = 1$ for proving certain asymptotic optimum properties of S_n in estimating F.

(*Added in proof*: In another paper we shall prove that analogous optimality results hold for S_n when $m > 1$, *even though D_n is no longer distribution free and the distribution theory of D_n is unknown*. These results follow from those of the present paper, arguments like those of [1], and the fact that the integral of a continuous bounded function with respect to G_n converges to that with respect to G uniformly in continuous F; the latter result will also appear in another paper.)

Some generalizations of Theorem 2 are mentioned at the end of §3.

2. **Proof of Theorem 1.** We shall give a detailed proof for $m = 2$. As we have remarked earlier, the proof for general m is by induction on m and is obvious to carry out. We shall indicate below the point where induction would be used. The result for the case $m = 1$ can be obtained by an argument similar to but simpler than that used below. Alternatively, it can be obtained from Lemma 2 of [2].

Throughout this section c_0 and c will be a generic notation for positive constants which do not depend on n, r, and F. Hence these symbols in different places will not, in general, stand for the same numbers. No confusion will be caused by this.

We have

$$(1 - G_n(r)) \leq (1 - G_n^+(r)) + (1 - G_n^-(r)).$$

We will content ourselves with proving (1.2). The proof of (1.3) follows in the same fashion, and (1.1) then follows from (1.2) and (1.3).

We shall assume that F is continuous. If this is not so (1.1), (1.2), and (1.3) hold a fortiori; the proof of this is the same as in the one-dimensional case and therefore obvious. Since F is continuous we may transform $X_i^{(1)}$ and $X_i^{(2)}$ separately so that the marginal d.f.'s F_j are uniform on $[0, 1]$ without changing $G_n(r)$, $G_n^+(r)$, or $G_n^-(r)$; we hereafter assume that the F_j are uniform on $[0, 1]$.

In the discussion which follows we shall always assume, to simplify the discussion, that, for any given number x_1, there is at most one i such that $X_i^{(1)} = x_1$. The probability that this be not so is zero.

In the course of the proof we shall always assume that $r < n^{1/2}$. The theorem is trivially true when this is not so.

If the theorem is true for all $r > R > 0$, it is true for all $r \geq 0$. One has only to enlarge c_0, if necessary, so that $c_0 e^{-cR^2} > 1$; the inequalities (1.1), (1.2), and (1.3) are then trivially true. It will therefore be sufficient to prove the theorem for all r sufficiently large, say $> R > 3$. Then $n > R^2$.

Since n and r will be fixed in the present discussion we may allow ourselves the luxury of a notation simpler than that of the next section and not display all dependences on n and r. (We remind the reader that c and c_0 will not depend on n and r.) Define the events

$$L = \left\{ S_n(x_1, x_2) - F(x_1, x_2) > \frac{r}{n^{1/2}} \text{ for some } x_1 \leq \frac{1}{2} \right\},$$

$$L' = \left\{ S_n(x_1, x_2) - F(x_1, x_2) < -\frac{r}{n^{1/2}} \text{ for some } x_1 \leq \frac{1}{2} \right\},$$

$$B = \left\{ S_n\left(\frac{1}{2}, x_2\right) - F\left(\frac{1}{2}, x_2\right) > \frac{r}{4n^{1/2}} \text{ for some } x_2 \right\},$$

$$\overline{L} = \left\{ S_n(x_1, x_2) - F(x_1, x_2) > \frac{r}{n^{1/2}} \text{ for some } x_1 > \frac{1}{2} \right\},$$

$$L^* = \left\{ S_n(x_1, x_2) - F(x_1, x_2) > \frac{r}{n^{1/2}} \text{ for some } x_1 > \frac{1}{2}, \ x_2 > \frac{1}{2} \right\}.$$

Define the chance variables (z_1, z_2) when the event L occurs (we shall not need them when L does not occur) as follows: First, z_1 is the infimum of those values x_1 ($\leq 1/2$) for which $\sup_{x_2}[S_n(x_1, x_2) - F(x_1, x_2)] > r/n^{1/2}$. There is then an i such that $X_i^{(1)} = z_1$. We define $z_2 = X_i^{(2)}$. We now define the event $L(x_1, x_2)$ for any pair x_1, x_2, $0 \leq x_1 \leq 1/2$, $0 \leq x_2 \leq 1$, as follows: $L(x_1, x_2)$ is the subset of L where $z_1 = x_1$, $z_2 = x_2$.

Define $r(x_1, x_2)$ as

$$\frac{1}{n^{1/2}} [(\text{least integer} > nF(x_1, x_2) + n^{1/2}r) - nF(x_1, x_2)].$$

Then, for almost all values (x_1, x_2) of (z_1, z_2), the event $L(x_1, x_2)$ implies the event

$$(2.1) \qquad \left\{ S_n(x_1, x_2) - F(x_1, x_2) = \frac{r(x_1, x_2)}{n^{1/2}} \right\}.$$

Define, for any x_2, $0 \leq x_2 \leq 1$, the events

$$B(x_2) = \left\{ S_n\left(\frac{1}{2}, x_2\right) - F\left(\frac{1}{2}, x_2\right) > \frac{r}{4n^{1/2}} \right\}$$

and

$$J = \bigcup_{0 \leq x_1 \leq 1/2} \bigcup_{0 \leq x_2 \leq 1} L(x_1, x_2) B(x_2).$$

Obviously

$$1 - G_n^+(r) \leq P\{L\} + P\{\overline{L}\}.$$

Our immediate goal will now be to prove

$$(2.2) \qquad P\{L\} < c_0 e^{-cr^2}$$

for all F, and for r sufficiently large, say $> R$.

We have

$$P\{L\} \cdot \text{ess. inf. } P\{J \mid z_1, z_2\} \leq \int_{0 \leq x_1 \leq 1/2; 0 \leq x_2 \leq 1} P\{J \mid x_1, x_2\} d_{x_1, x_2} P\{z_1 < x_1, z_2 < x_2\}$$

$$= P\{J \cap L\} \leq P\{J\} \leq P\{B\}.$$

Hence

$$(2.3) \qquad P\{L\} \leq \frac{P\{B\}}{\text{ess. inf. } P\{J \mid z_1, z_2\}}.$$

Our plan to prove (2.2) is as follows: First, we shall prove that

$$(2.4) \qquad \text{ess. inf. } P\{J \mid z_1, z_2\} \geq 1/2.$$

Then we shall prove that

$$(2.5) \qquad P\{B\} < c_0 e^{-cr^2}$$

for $r > R$ and $n > r^2$.

Suppose the event $L(x_1, x_2)$ has occurred. Since there is exactly one i such that $X_i^{(1)} = x_1$, and since

$$\sup_{x_2} [S_n(x_1', x_2) - F(x_1', x_2)] \leq \frac{r}{n^{1/2}}$$

for $x_1' < x_1$, we have

(2.6) $$nS_n(x_1, 1) < nF(x_1, 1) + rn^{1/2} + 1.$$

Hence the number N of $X_1, \cdots, X_n$ which have first coordinate greater than x_1 is at least

(2.7) $$M = n(1 - x_1) - rn^{1/2} - 1.$$

First suppose that $M \leq 0$. Then

(2.8) $$\frac{r}{n^{1/2}} \geq -\frac{1}{n} - (1 + x_1).$$

Obviously $0 \leq r(x_1, x_2) - r \leq 1/n^{1/2}$. The event $B(x_2)$ occurs when

(2.9) $$S_n\left(\frac{1}{2}, x_2\right) > F\left(\frac{1}{2}, x_2\right) + \frac{r}{4n^{1/2}}.$$

From (2.1) and (2.8) we obtain that

$$S_n\left(\frac{1}{2}, x_2\right) \geq S_n(x_1, x_2) \geq F(x_1, x_2) + \frac{r}{n^{1/2}}$$

$$\geq F(x_1, x_2) + (1 - x_1) - \frac{1}{n}$$

(2.10) $$= F\left(\frac{1}{2}, x_2\right) - \left[F\left(\frac{1}{2}, x_2\right) - F(x_1, x_2)\right] + (1 - x_1) - \frac{1}{n}$$

$$\geq F\left(\frac{1}{2}, x_2\right) - \left(\frac{1}{2} - x_1\right) + (1 - x_1) - \frac{1}{n}$$

$$= F\left(\frac{1}{2}, x_2\right) + \frac{1}{2} - \frac{1}{n}.$$

Since $r/4n^{1/2} < 1/4$, it follows that, for $n > R^2 > 9$ (which is all we need consider), (2.9) holds.

Suppose now that $M > 0$. Let R_0 be the region in the x_1', x_2' plane defined by the inequalities

$$x_1 < x_1' \leq 1/2, \qquad 0 \leq x_2' \leq x_2.$$

In order for $B(x_2)$ to occur it is sufficient, by (2.9) and (2.1), that, of the N chance variables among $X_1, \cdots, X_n$ whose first coordinates are greater than x_1, at least

(2.11) $$n(F(1/2, x_2) - F(x_1, x_2)) - 3rn^{1/2}/4$$

take on values in R_0. We shall compute a lower bound for the probability of this under the assumption that $N = M \ (>0)$. It will be easy to see that, if

$N > M$ or M is not an integer, the probability is a fortiori greater than this lower bound, which is $> 1/2$ for $r > R$. This will prove (2.4).

If we define

$$(2.12) \qquad p = \frac{F(1/2, x_2) - F(x_1, x_2)}{1 - x_1}$$

and

$$(2.13) \qquad t = \frac{n(F(1/2, x_2) - F(x_1, x_2)) - 3rn^{1/2}/4 - Mp}{(Mp(1 - p))^{1/2}},$$

we obtain that

$$t = \frac{r(p - 3/4) + p/n^{1/2}}{\left(\dfrac{Mp(1 - p)}{n}\right)^{1/2}}.$$

Since $p < 1/2$ and $M/n \leq 1$ it follows that $t < -r/4$ for $r > R$. The probability in question is the probability that, of M independent Bernoulli chance variables with common probability p of a "success," the number N^* of "successes" satisfy the inequality

$$(2.14) \qquad \frac{N^* - EN^*}{(E(N^* - EN^*)^2)^{1/2}} > t.$$

This probability is greater than

$$(2.15) \qquad P\left\{\frac{N^* - EN^*}{(E(N^* - EN^*)^2)^{1/2}} > -\frac{r}{4}\right\}$$

which, by Chebyshev's inequality, is greater than $1 - 16/r^2$, which, for $r > R$, is $> 1/2$, as was to be proved. This proves (2.4).

From [6, p. 288, Equation (96)], it follows that, for $r > R$,

$$(2.16) \qquad P\left\{\left|S_n\left(\frac{1}{2}, 1\right) - \frac{1}{2}\right| > \frac{r}{16n^{1/2}}\right\} < c_0 e^{-cr^2}.$$

Suppose now that the event $\{|S_n(1/2, 1) - 1/2| \leq r/16n^{1/2}\}$ occurs. Then the number n_1 of chance variables $X_1, \cdots, X_n$ with first coördinate not greater than $1/2$ satisfies

$$\frac{n}{2} - \frac{rn^{1/2}}{16} = n_2 < n_1 < n_3 = \frac{n}{2} + \frac{rn^{1/2}}{16} < \frac{9n}{16}.$$

Let $T_{n_1}(x_2)$ denote $1/n_1$ multiplied by the number of chance variables $X_1, \cdots, X_n$ whose first coordinates are less than $1/2$ and whose second coördinates are less than x_2. The relation

$$(2.17) \qquad nS_n\left(\frac{1}{2}, x_2\right) - nF\left(\frac{1}{2}, x_2\right) > \frac{rn^{1/2}}{4}$$

implies the relation

$$(2.18) \qquad n_1 T_{n_1}(x_2) - n_1\left[2F\left(\frac{1}{2}, x_2\right)\right] > \frac{rn^{1/2}}{4} - \frac{rn^{1/2}}{8} F\left(\frac{1}{2}, x_2\right),$$

whose right member is not less than

$$(2.19) \qquad \frac{3rn^{1/2}}{16} > \frac{r(n_3)^{1/2}}{4} > \frac{r(n_1)^{1/2}}{4}.$$

The theorem for the case $m = 1$ implies that

$$(2.20) \qquad P\left\{\sup_{x_2}\left[n_1 T_{n_1}(x_2) - n_1\left(2F\left(\frac{1}{2}, x_2\right)\right)\right] > \frac{r(n_1)^{1/2}}{4}\right\} < c_0 e^{-cr^2}.$$

Equations (2.16) to (2.20) prove (2.5) and hence (2.2).

In the proof of Theorem 1 for general m the induction on m would occur at this point. We have just used the theorem for $m = 1$ to prove (2.5) for $m = 2$. We can then use this to prove the result corresponding to (2.5) for $m = 3$, and so on (x_2 represents all variables other than x_1 in this proof, when $m > 2$).

Returning to the case $m = 2$, the proof of

$$(2.21) \qquad P\{L'\} < c_0 e^{-cr^2}$$

is practically the same as that of (2.2), and will be omitted. We shall henceforth assume that (2.21) holds, and use this fact to prove that

$$(2.22) \qquad P\{\bar{L}\} < c_0 e^{-cr^2}.$$

First, applying the result (2.2) to the chance variables $X_1^*, X_2^*, \cdots$ defined by $X_i^* = (X_i^{(2)}, X_i^{(1)})$, we obtain for the original sequence $X_1, X_2, \cdots$ that

$$(2.23) \qquad P\left\{S_n(x_1, x_2) - F(x_1, x_2) > \frac{r}{n^{1/2}} \text{ for some } x_2 \leq \frac{1}{2}\right\} < c_0 e^{-cr^2}.$$

For any pair (x_1, x_2), $1/2 < x_1 \leq 1$, $1/2 < x_2 \leq 1$, we define the following regions in the x_1', x_2' plane:

$$U_1(x_1, x_2) = \{x_1', x_2' \mid x_1 < x_1' \leq 1, x_2 < x_2' \leq 1\},$$
$$U_2(x_1, x_2) = \{x_1', x_2' \mid x_1 < x_1' \leq 1, 0 \leq x_2' < x_2\},$$
$$U_3(x_1, x_2) = \{x_1', x_2' \mid 0 \leq x_1' < x_1, x_2 < x_2' \leq 1\},$$

and the following events for $i = 1, 2, 3$:

$$Q_i = \left\{ \text{for some } (x_1, x_2), \frac{1}{2} < x_1 \leqq 1, \frac{1}{2} < x_2 \leqq 1, \text{ the number of } (X_1, \cdots, X_n) \right.$$

$$\left. \text{in } U_i(x_1, x_2) \text{ minus expected number } < -\frac{rn^{1/2}}{3} \right\}.$$

Obviously

$$L^* \subset Q_1 \cup Q_2 \cup Q_3.$$

We shall prove

$$(2.24) \qquad\qquad P\{Q_v\} < c_0 e^{-cr^2}, \qquad\qquad v = 1, 2, 3.$$

The result (2.24) follows for $v=1$ from the application of (2.21) to the sequence of chance variables $(1-X_i^{(2)}, 1-X_i^{(1)})$, for $v=2$ by the application of (2.21) to the sequence of chance variables $(1-X_i^{(1)}, X_i^{(2)})$, and for $v=3$ by the application of (2.21) (in the form (2.23)) to the sequence of chance variables $(X_i^{(1)}, 1-X_i^{(2)})$. Thus, (2.24) is proved, and this and (2.23) imply (2.22) and hence (1.2).

The proof of (1.3) is completed in a similar manner. Obviously (1.2) and (1.3) imply (1.1). This completes the proof of Theorem 1.

We shall now obtain explicit possible values for the constant c (c_0 could be obtained similarly, but this is of less interest). First consider the case $m=2$. In the definition of the set B, let us replace $r/4$ by $r/(2+\epsilon)$ with $\epsilon > 0$; the proof of Theorem 1 then still holds, but will yield a larger value of c. Making appropriate changes in the argument, an analogue of (2.15) holds, as before. In (2.16) and what follows, put λ for $1/16$. The constant c on the right side of (2.16) then becomes $2\lambda^2$. The displayed inequality on n_1 becomes

$$(2.25) \quad n\left(\frac{1}{2} - \lambda\right) < \frac{n}{2} - r\lambda n^{1/2} < n_1 < \frac{n}{2} + r\lambda n^{1/2} < n\left(\lambda + \frac{1}{2}\right).$$

Equation (2.17) (with $1/(2+\epsilon)$ for $1/4$) and Equation (2.25) imply an analogue of (2.18) with (2.19) replaced by

$$rn^{1/2}\left(\frac{1}{2+\epsilon} - \lambda\right) > rn_1^{1/2}\left(\frac{1}{2+\epsilon} - \lambda\right) \Big/ \left(\lambda + \frac{1}{2}\right)^{1/2}.$$

The fact that we can take $c=2$ for $m=1$ implies, in place of (2.20),

$$P\left\{ \sup_{x_2} \left| n_1 T_{n_1}(x_2) - 2n_1 F\left(\frac{1}{2}, x_2\right) \right| > rn^{1/2}\left(\frac{1}{2+\epsilon} - \lambda\right) \Big/ \left(\lambda + \frac{1}{2}\right)^{1/2} \right\}$$

$$(2.26)$$

$$< c_0 \exp\left(-2r^2\left(\frac{1}{2+\epsilon} - \lambda\right)^2 \Big/ \left(\lambda + \frac{1}{2}\right)\right).$$

The minimum of the coefficients of r^2 in the exponents of (2.26) and the

analogue of (2.16) is maximized when ϵ is small by taking $\lambda \doteq .266$; this gives $c > .142 + o(1)$ as $\epsilon \to 0$, in (2.5). The same value of c may be obtained similarly for L' in (2.21), and also in (2.23). If this value of c is multiplied by $1/9$, we obtain a value applicable in (2.24), (2.22), and (1.2); a similar argument applies for (1.1). Thus, we also obtain $c > .0157 - \epsilon'$ in (1.3) for $m = 2$, where ϵ' is an arbitrary positive value. Thus, $c = .0157$ is a possible value in (1.3) for $m = 2$.

For general m, we may similarly obtain a possible value for c. Let $2d_m^2$ be the value for c obtained by this argument for dimension m, with $d_m > 0$. For dimension m we then obtain $2d_{m-1}^2(1/2 - \lambda)^2$ for the coefficient of r^2 in (2.26) and thus the solution λ of the equation $2\lambda^2 = 2d_{m-1}^2(1/2 - \lambda)^2/(\lambda + 1/2)$ is the value of λ which maximizes the minimum of the coefficients of r^2 in (2.16) and (2.26). Rather than carry out the obvious analogue of the case $m = 2$ in terms of this inexplicit λ, we shall obtain explicitly a slightly smaller value of c. This value is suggested by the fact that d_{m-1}, and hence the above λ, is small for $m > 2$. Taking then for λ the value $d_{m-1}/2^{1/2}$, the two coefficients of r^2 are almost equal, the smaller (that of (2.26)) being

$$2d_{m-1}^2(1/2 - d_{m-1}/2^{1/2})^2/(1/2 + d_{m-1}/2^{1/2}).$$

The factor $1/9$ above must be replaced by $(2^m - 1)^{-2}$. Thus, we obtain for possible values of d_m and c (the ϵ' no longer being needed):

(2.27)
$$c = 2d_m^2 \quad \text{(dimension } m),$$
$$d_m = \frac{d_{m-1}(1/2 - d_{m-1}/2^{1/2})}{(1/2 + d_{m-1}/2^{1/2})^{1/2}(2^m - 1)} \qquad (d_2 = .088).$$

The above possible value for c is probably not a very good one ($c = .0157$ for $m = 2$ and $c = .000107$ for $m = 3$). It could be improved by considering $S_n(x) - F(x)$ at a large finite number of lines (in the case $m = 2$, for example) instead of just on the line $x_1 = 1/2$; but this would be at the expense of more tedious computations. The value $c = 2$ obtained in [2] for the case $m = 1$ is the best possible in the sense that (1.3) is clearly false for any $c > 2$ and any c_0. We next show that $c < 2$ for $m > 1$; i.e., (1.3) is false for $c = 2$ and any c_0 when $m > 1$ (in fact, this is so even if c_0 is permitted to depend on F).

In fact, consider the case $m = 2$ and suppose F_1 is the d.f. which distributes all probability uniformly on the line $L: x_1 + x_2 = 1$. Then $S_n(x, y) = 0$ w.p. 1 if $x + y \leq 1$, and for $x + y > 1$ we have $nS_n(x, y) = $ number of observations on L between $(1 - y, y)$ and $(x, 1 - x)$. Let $S_n^*(u) = S_n(u, 1)$. Of course, $S_n^*(u)$ is a univariate sample d.f. corresponding to the uniform d.f. on the unit interval. Denoting by D_n^+ and D_n^- the supremum positive and negative deviations of $S_n^*(u)$ from the function u, $0 \leq u \leq 1$, we have, w.p.1.,

$$\sup_{x,y} |S_n(x, y) - F_1(x, y)| = \sup_{u,v} |(S_n^*(u) - u) - (S_n^*(v) - v)| = D_n^+ + D_n^-.$$

From [5] we obtain for the limiting d.f. of $n^{1/2}(D_n^+ + D_n^-)$ (note, e.g., Equation (4.6) of [7], which gives the limiting d.f. of $n^{1/2}(D_n^+ + D_n^-)/2$), as $r \to \infty$,

$$\lim_{n \to \infty} P\{n^{1/2}(D_n^+ + D_n^-) > r\} \sim 8r^2 e^{-2r^2},$$

which demonstrates the impossibility of taking $c = 2$. We note, in fact, that (1.3) cannot hold for $c = 2$ and any c_0 and for all absolutely continuous F 'or, instead, for all *discrete F*); this is obtained easily from the above result by taking a fixed r so large that $4r^2 > c_0$, a k so large that the above limiting probability for the case of the discrete approximation of F_1 is $> 6r^2 e^{-2r^2}$, and an absolutely continuous d.f. F_2 whose probability is concentrated on such small spheres about the discrete points that the probability of a deviation $> r$ cannot be smaller for F_2 than for F_1.

The supremum operation involved in the definition of D_n, D_n^+, and D_n^- is over all sets of the form $x_j \leq a_j$, $j = 1, \cdots, m$, for all $a = (a_1, \cdots, a_m)$ in m-space. It is obvious that Theorem 1 applies also to the case when the supremum is taken over any of several larger classes of sets such as, for example, that which consists of all rectangular parallelepipeds with sides parallel to the coördinate planes. This will be of interest in statistical applications where it is often required or desired that the results be invariant under certain transformations of the chance variables, e.g., $X \to -X$.

3. **Proof of Theorem** 2. Let I^m denote the closed unit m-cell $\{x \mid 0 \leq x_1, \cdots, x_m \leq 1\}$. We shall first prove Theorem 2 for the case when F is continuous, and then, at the conclusion of the proof, we shall remark on how the proof proceeds for discontinuous F. As in §2, since F is now assumed continuous, it suffices to consider the case where all F_j's are uniform on $[0, 1]$, and we hereafter assume we are in this case. Write $Q_{k,0} = I^m$, and for $j > 0$ let $Q_{k,j}$ be the subset of I^m whose first j coördinates are integral multiples of $1/k$ (thus, $Q_{k,m} = A_k^m$). Write

$$D_{n,k,j}^+ = \sup_{x \in Q_{k,j}} [S_n(x) - F(x)],$$

$$D_{n,k,j}^- = \sup_{x \in Q_{k,j}} [F(x) - S_n(x)].$$

For fixed $d > 0$, r, and r', we shall show in the succeeding paragraphs that

$$\limsup_{k \to \infty} \limsup_{n \to \infty} [P\{D_{n,k,j}^+ < r/n^{1/2}; D_{n,k,j}^- < r'/n^{1/2}\}$$

(3.1)
$$- P\{D_{n,k,j-1}^+ < (r+d)/n^{1/2}; D_{n,k,j-1}^- < (r'+d)/n^{1/2}\}] \leq 0$$

for $1 \leq j \leq m$. Adding these inequalities over j yields

$$(3.2) \qquad \limsup_{k \to \infty} \limsup_{n \to \infty} \left[H_{n,k}(r, r') - H_n(r + md, r' + md) \right] \leqq 0.$$

We have remarked in §1 that $\lim_n H_{n,k} = H_{\infty,k}$ exists. Hence, writing r for $r + md$ and r' for $r' + md$, (3.2) becomes

$$(3.3) \qquad \limsup_{k \to \infty} H_{\infty,k}(r - md, r' - md) \leqq \liminf_{n \to \infty} H_n(r, r').$$

Write $H^*(r, r') = \limsup_k H_{\infty,k}(r, r')$. Since obviously $H_n(r, r') \leqq H_{n,k}(r, r')$, we obtain from (3.3),

$$(3.4) \quad H^*(r - md, r' - md) \leqq \liminf_{n \to \infty} H_n(r, r') \leqq \limsup_{n \to \infty} H_n(r, r') \leqq H^*(r, r').$$

Since H^* is clearly monotone and bounded, it is continuous except possibly on a denumerable set of lines parallel to the coördinate axes. Letting d tend to zero in (3.4) at continuity points (r, r') of H^*, we see that $\lim_n H_n(r, r')$ exists for all points (r, r') in the plane, except possibly on a denumerable set of lines parallel to the coördinate axes. This limit determines a left-continuous function H (say) which has variation one by Theorem 1 and which is clearly a d.f. Hence the sequence H_n converges to a d.f. H at every continuity point of the latter. Finally, since clearly we can also write, for all continuity points (r, r') of H,

$$(3.5) \qquad \begin{aligned} H(r, r') &\leqq \liminf_{k \to \infty} H_{\infty,k}(r, r') \\ &\leqq \limsup_{k \to \infty} H_{\infty,k}(r, r') \leqq H(r + md, r' + md), \end{aligned}$$

letting $d \to 0$ shows that $\lim_k H_{\infty,k}(r, r') = H(r, r')$ at all continuity points of the latter, and hence that $H_{\infty,k}$ converges to H at every continuity point of the latter as $k \to \infty$. Thus, the theorem for H will be proved if we can show (3.1), and the result for G, G^+, and G^- can be obtained easily from this result or else can be proved directly in the same manner as the result for H.

We now prove (3.1). Fix $d > 0$, r, r', and j. For h an integer, write $V_{k,j,h}$ for the subset of I^m where $h/k < x_j \leqq (h+1)/k$ and write $x^{(i)} = (x_1, \cdots, x_{j-1}, h/k, x_{j+1}, \cdots, x_m)$ if $x \in V_{k,j,h}$. Clearly, if the event

$$\Lambda_{n,k} = \left\{ D^+_{n,k,j} < r/n^{1/2}, \ D^-_{n,k,j} < r'/n^{1/2} \right\}$$

occurs and the event

$$\Lambda^*_{n,k} = \left\{ D^+_{n,k,j-1} < (r + d)/n^{1/2}, \ D^-_{n,k,j-1} < (r' + d)/n^{1/2} \right\}$$

does not occur, it is necessary that for some h with $0 \leqq h \leqq k - 1$ the event

$$\Gamma_{n,k,h} = \left\{ \sup_{x \in V_{k,j,h}} \left| [S_n(x) - F(x)] - [S_n(x^{(i)}) - F(x^{(i)})] \right| > d/n^{1/2} \right\}$$

occurs. Write N_{nkjh} = number of X_i $(1 \leq i \leq n)$ whose values are in $V_{k,j,h}$. Now,

$$P\{\Gamma_{n,k,h}\} \leq P\left\{\sup_{V_{k,j,h}} \left|\frac{S_n(x) - S_n(x^{(i)})}{N_{nkjh}/n} - k\left[F(x) - F(x^{(i)})\right]\right| > dn^{1/2}/2N_{nkjh}\right\}$$

$$(3.6) \qquad + P\left\{\left|\frac{kN_{nkjh}}{n} - 1\right| \sup_{V_{k,j,h}} \left[F(x) - F(x^{(i)})\right] > d/2n^{1/2}\right\}$$

$$= P\{Y_{nkjh}\} + P\{Z_{nkjh}\} \text{ (say).}$$

Now, given that $N_{nkjh} = N$, the event Y_{nkjh} is a subset of the event that the maximum deviation of the empiric d.f. of N independent, identically distributed m-variate random variables from the corresponding theoretical d.f. is more than $(d/2)(n/N)^{1/2}/N^{1/2}$. Hence, by Theorem 1,

$$(3.7) \qquad P\{Y_{nkjh} \mid N_{nkjh} = N\} < c_0 \exp\{-cd^2 n/N\}$$

where c_0 and c are positive constants. Also, by Chebyshev's inequality,

$$(3.8) \qquad P\{N_{nkjh} \leq n/k + n/k^{1/2}\} > 1 - 1/n.$$

From (3.7) and (3.8) we obtain

$$(3.9) \qquad P\{Y_{nkjh}\} < n^{-1} + c_0 \exp\{-cd^2 k^{1/2}/2\}.$$

Also, an application of Chebyshev's inequality based on the fourth moment yields

$$(3.10) \qquad P\{Z_{nkjh}\} = P\left\{\left|\frac{N_{nkjh}}{n} - \frac{1}{k}\right| > d/2n^{1/2}\right\}$$

$$< \frac{3n^2/k^2 + n/k}{n^4[d^4/16n^2]} = \frac{16}{d^4}\left[\frac{3}{k^2} + \frac{1}{nk}\right].$$

From (3.6), (3.9), and (3.10) we obtain

$$(3.11) \quad P\{\Lambda_{n,k} - \Lambda_{n,k}^*\} < \frac{k}{n} + kc_0 \exp\{-cd^2 k^{1/2}/2\} + \frac{16}{d^4}\left[\frac{3}{k} + \frac{1}{n}\right],$$

which proves (3.1) and, hence, Theorem 2 in the continuous case.

We now remark on the method of proving Theorem 2 when F has discontinuities. The conclusion follows by the same method as that used to prove Theorem 2 for continuous F, upon noting the manner in which any discontinuous F can be obtained from a continuous one by "lumping together" (in the same manner as that used to obtain Theorem 1 for discontinuous F) certain points in the domain of the latter.

Generalizations of the theorem may be obtained by noting that, as in the case of Theorem 1, the conclusion of Theorem 2 holds if the supremum of ob-

served from theoretical frequency is taken over a larger class of sets than those of the form $x_j \leqq a_j$, $j = 1, \cdots, m$ (for all a in m-space).

REFERENCES

1. A. N. Kolmogorov, *Sulla determinazione empirica di una legge di distribuzione*, Inst. Ital. Atti. Giorn. vol. 4 (1933) pp. 83–91.

2. A. Dvoretzky, J. Kiefer, and J. Wolfowitz, *Asymptotic minimax character of the sample distribution function and of the classical multinomial estimator*, Ann. Math. Stat. vol. 27 (1956) pp. 642–669.

3. N. V. Smirnov, *Approximation of the distribution law of a random variable by empirical data*, Uspehi Matematičeskih Nauk vol. 10 (1944) pp. 179–206.

4. M. Donsker, *Justification and extension of Doob's heuristic approach to the Kolmogorov-Smirnov theorems*, Ann. Math. Stat. vol. 23 (1952) pp. 277–281.

5. J. L. Doob, *Heuristic approach to the Kolmogorov-Smirnov theorems*, Ann. Math. Stat. vol. 20 (1949) pp. 393–403.

6. P. Lévy, *Théorie de l'addition des variables aléatoires*, Paris, Gauthier-Villars, 1937.

7. M. Kac, J. Kiefer and J. Wolfowitz, *On tests of normality and other goodness of fit based on distance methods*, Ann. Math. Stat. vol. 26 (1955) pp. 189–211.

CORNELL UNIVERSITY,
    ITHACA, N. Y.