ON LARGE DEVIATIONS OF THE EMPIRIC D.F. OF VECTOR CHANCE VARIABLES AND A LAW OF THE ITERATED LOGARITHM

J. KIEFER

1. Introduction and preliminaries. Let F be a distribution function (d.f.) on m-dimensional Euclidean space R^m, and let $X_1, \cdots, X_n$ be independent chance vectors with common d.f. F. The empiric d.f. S_n is a chance d.f. on R^m defined as follows: if $x = (x_1, \cdots, x_m)$, $nS_n(x)$ is the number of X_i's, $1 \leqq i \leqq n$, such that, for $j = 1, \cdots, m$, the jth component $X_i^{(j)}$ of X_i is less than or equal to x_i.

When $m = 1$, the distribution of $D_n = \sup_x |S_n(x) - F(x)|$ is the same for all continuous F, and Kolomogorov [5] first computed the limiting distribution of $n^{1/2}D_n$ as $n \to \infty$. Chung [1] gave a bound on the error term which was sharp enough to yield a law of the iterated logarithm for the empiric d.f. and, in fact, the more precise complete characterization of monotone functions of upper and lower class. (The more recent literature contains several asymptotic expansions of Kolmogorov's distribution.) It was proved by Dvoretzky, Kiefer and Wolfowitz [2] that there is a universal constant C such that, for all $n > 0$ and $r \geqq 0$,

$$P\{n^{1/2}D_n \geqq r\} \leqq Ce^{-2r^2}\ ; \tag{1.1}$$

since $\lim_n P\{n^{1/2}D_n \geqq r\}$ is asymptotically $2e^{-2r^2}(1 + o(1))$ as $r \to \infty$, the estimate (1.1) cannot be improved upon in this general form.

Much less is known when $m > 1$. The limiting d.f. of $n^{1/2}D_n$ was proved to exist by Kiefer and Wolfowitz [4]; of course, its form depends on F (and is unknown except in a few trivial cases), unlike the case $m = 1$. It was also proved in [4] that there exist positive constants c_m and c'_m such that, for all F, $n > 0$, and $r \geqq 0$,

$$P_F\{n^{1/2}D_n \geqq r\} \leqq c'_m e^{-c_m r^2}\ , \tag{1.2}$$

whereby $P_F\{A\}$ we denote the probability of the event A when X_1 has d.f. F. Possible choices of the constants c_m were shown to be

$$c_2 = .0157,\ c_3 = .000107,\ \cdots\ (\lim_m c_m = 0)\ . \tag{1.3}$$

It was also shown in [4] that, for $m > 1$, one cannot have $c_m = 2$ in (1.2); specifically, if $m = 2$ and F^* distributes probability uniformly on the line segment $\{(x_1, x_2): x_1 \geqq 0, x_2 \geqq 0, x_1 + x_2 = 1\}$, then as $r \to \infty$ we have

$$\lim_n P_{F^*}\{n^{1/2}D_n \geqq r\} = 8r^2e^{-2r^2}(1 + o(1))\ . \tag{1.4}$$

Received May 9, 1960. Research sponsored by the Office of Naval Research.