The Annals of Probability
1975, Vol. 3, No. 3, 526-533

ORDER OF NORMAL APPROXIMATION FOR RANK TEST STATISTICS DISTRIBUTION

BY JANA JUREČKOVÁ AND MADAN L. PURI[1]

Charles University and Indiana University

0. Summary. Under suitable assumptions, it is established that the rate of convergence of the cdf (cumulative distribution function) of the simple linear rank statistics

$$S_N = \sum_{i=1}^N C_{Ni}\varphi\left(\frac{R_{Ni}}{N+1}\right)$$

to the normal one is $O(N^{-\frac{1}{2}+\delta})$ for any $\delta > 0$. Here $C_{N1}, \cdots, C_{NN}$ are known constants, $R_{N1}, \cdots, R_{NN}$ are the ranks of independent observations $X_{N1}, \cdots, X_{NN}$, and φ is a score generating function defined in Section 1.

1. Introduction. Let X_{Ni}, $i = 1, \cdots, N$ be independent rvs distributed according to the cdf $F_i(x) = F(x - \Delta d_{Ni})$, $i = 1, \cdots, N$. We assumed that $F(x)$ is absolutely continuous having the density function $f(x)$ whose derivative $f'(x)$ exists. Furthermore, $F(x)$ is assumed to have the finite Fisher information, that is,

$$I(f) = \int_{-\infty}^{\infty} [f'(x)/f(x)]^2 f(x)\, dx < \infty . \tag{1.1}$$

Δ is an unknown parameter, and d_{Ni}, $i = 1, \cdots, N$ are known constants. Let R_{Ni} be the rank of X_{Ni} among $X_{N1}, \cdots, X_{NN}$. Setting $u(x) = 1$ if $x \geqq 0$, and $u(x) = 0$ otherwise, we can write

$$R_{Ni} = \sum_{j=1}^N u(X_{Ni} - X_{Nj}) , \qquad i = 1, \cdots, N . \tag{1.2}$$

Consider now the simple linear rank statistics

$$S_N = \sum_{i=1}^N C_{Ni}\, a_N(R_{Ni}) \tag{1.3}$$

where $C_{N1}, \cdots, C_{NN}$ are known constants, and $a_N(i)$, $i = 1, \cdots, N$ are "scores" generated by a function $\varphi(t)$ in the following manner:

$$a_N(i) = \varphi\left(\frac{i}{N+1}\right), \qquad 1 \leqq i \leqq N . \tag{1.4}$$

Statistics of the type (1.3) play an important role in the theory of nonparametric inference. For example, in the two sample problem where $F_1 = \cdots = F_m \equiv F$, and

$$F_{m+1} = \cdots = F_N \equiv G ,$$

Received March 26, 1974; revised October 1, 1974.

[1] Work supported by the Air Force Office of Scientific Research, AFSC, USAF, under Grant No. AFOSR 71-2009. Reproduction in whole or in part is permitted for any purpose of the U.S. Government. Part of this work was done while the author was the recipient of a Senior U.S. Scientist Award from the Alexander-von-Humboldt-Foundation.

AMS 1970 *subject classifications.* Primary 62E20; Secondary 60F05, 60F99.

Key words and phrases. Linear rank statistics, score generating function, rate of convergence.