ON BAHADUR'S REPRESENTATION OF SAMPLE QUANTILES

By J. Kiefer[1]

Cornell University

1. Introduction and summary. Let $X_1, X_2, \cdots$ be independent and identically distributed real random variables with common df F. Suppose that $0 < p < 1$, that $F(\xi_p) = p$, that F is twice differentiable in a neighborhood of p, and that F'' is bounded in that neighborhood and $F'(\xi_p) > 0$. Let S_n be the sample df based on $(X_1, \cdots, X_n)$; i.e., $nS_n(x) =$ number of $X_i \leqq x$, $1 \leqq i \leqq n$. Let $Y_{p,n}$ be a sample p-quantile based on $(X_1, \cdots X_n)$; i.e., $S_n(Y_{p,n}-) \leqq p \leqq S_n(Y_{p,n})$; if np is an integer, so that $Y_{p,n}$ is not unique, it will be seen that any measurable definition can be used in the sequel, and for the sake of definiteness we shall take the smallest possible value. We shall write $\sigma_p = [p(1-p)]^{\frac{1}{2}}$.

Let

$$R_n(p) = Y_{p,n} - \xi_p + [S_n(\xi_p) - p]/F'(\xi_p). \tag{1.1}$$

Bahadur (1966) initiated the study of $R_n(p)$ and showed that

$$R_n(p) = O(n^{-3/4}(\log n)^{\frac{1}{2}}(\log \log n)^{\frac{1}{4}}) \tag{1.2}$$

wp 1 as $n \to \infty$. He also raised the question of finding the exact order of $R_n(p)$. In the present paper we answer this by proving

Theorem 1. *For either choice of sign,*

$$\limsup_{n\to\infty} \pm F'(\xi_p)R_n(p)/[2^{5/4}3^{-3/4}\sigma_p^{\frac{1}{2}}n^{-3/4}(\log\log n)^{3/4}] = 1 \tag{1.3}$$

wp 1.

Later in this section we shall sketch the rationale behind this result, whose proof will occupy most of the paper.

Let τ be any positive number. For $t \in [-\tau, \tau]$, we write

$$U_n(t) = n^{3/4}F'(\xi_{p+n^{-\frac{1}{2}}t})R_n(p + n^{-\frac{1}{2}}t), \tag{1.4}$$
$$U_n \equiv U_n(0) = n^{3/4}F'(\xi_p)R_n(p),$$

which by assumption on F make sense for n large. Also write

$$K_n = n^{\frac{1}{2}}(p - S_n(\xi_p)). \tag{1.5}$$

As the discussion later in this section shows, it is trivial that, for each $b > 0$, uniformly in $b^{-1} < |k_n| < b$, the conditional law of $U_n/|k_n|^{\frac{1}{2}}$, given $K_n = k_n$, is asymptotically $N(0, 1)$ as $n \to \infty$. Since K_n/σ_p is also asymptotically $N(0, 1)$, it is obvious that

$$\lim_{n\to\infty} P\{U_n \leqq u\} = 2\sigma_p^{-1}\int_0^\infty \Phi(k^{-\frac{1}{2}}u)\phi(k/\sigma_p)\,dk, \tag{1.6}$$

where Φ and ϕ are the standard $N(0, 1)$ df and density. More generally, we shall

Received 22 December 1966.

[1] Research supported by the Office of Naval Research under Contract No. NONR 401 (50). Reproduction in whole or in part is permitted for any purpose of the United States Government.