ON AN ANALOG OF REGRESSION ANALYSIS

By P. K. Bhattacharya[1]

University of North Carolina

1. Introduction. Suppose $(X_1, \cdots, X_h, Y)$ follow an unknown multivariate distribution on which independent observations are made. The nature of dependence of Y on $(X_1, \cdots, X_h)$ is fully understood only when we know how the conditional distribution of Y given $X_1 = x_1, \cdots, X_h = x_h$, changes with $(x_1, \cdots, x_h)$. This can be attempted in two different ways. One approach is to make inference about the functional relation between the conditional moments of Y given $X_1 = x_1, \cdots, X_h = x_h$, and $(x_1, \cdots, x_h)$, and a special case of this (when the behavior of only the conditional expectation of Y is studied) is known in statistical literature as regression analysis. The classical methods of regression analysis are based on the assumption that $E[Y \mid X_1 = x_1, \cdots, X_h = x_h]$ is a linear function of $x_1, \cdots, x_h$. Mahalanobis [3] and Parthasarathy and Bhattacharya [4] have proposed some methods of regression analysis which do not involve any such linearity assumption. The methods proposed by Parthasarathy and Bhattacharya can be generalized in a straightforward manner for estimating and testing hypotheses about higher order conditional moments of Y, but in order to prove the consistency of these estimates and tests for the first m conditional moments, one should assume the existence of at least first $3m$ conditional absolute moments. A second approach to this problem is to make inference about the functional relation between the quantiles of the conditional distribution of Y given $X_1 = x_1, \cdots, X_h = x_h$, and $(x_1, \cdots, x_h)$. This kind of an analog of the variance components analysis has been considered by Roy and Cobb [5], while Sathe [6] has given a test for the conditional median of Y given X under a restricted model.

In this paper, estimates have been proposed for conditional quantile functions of Y given $X_1, \cdots, X_h$, and the simultaneous uniform convergence of any number of such estimates to the corresponding conditional quantile functions has been studied. A large sample test for the hypothesis that certain conditional quantile functions are equal to specified functions, has been suggested and proved to be consistent. In order to avoid complicated notations, the methods and their properties will be discussed for the bivariate case in Sections 2, 3 and 4, while in Section 5, the corresponding methods for the multivariate case will be explained and their properties will be stated without using too many symbols.

2. Problem, assumptions and notations. (X, Y) has a bivariate distribution. F is the marginal distribution function of X and G_x is the conditional distribu-

Received September 7, 1962; revised May 30, 1963.

[1] This research was supported by the Air Force Office of Scientific Research.

tion function of Y given $X = x$. For any $0 < p < 1$, $\phi_p(x)$ is the solution of the equation

$$G_x(\phi_p(x)) = p. \tag{1}$$

In other words, $\varphi_p(x)$ is the conditional p-quantile of Y given $X =$

On the basis of independent observations on (X, Y), we want to estimate the function ϕ_p for a given p, and for a specified real valued function μ defined on the range of X, we want to test the hypothesis $H_0:\phi_p = \mu$.

In what follows, $0 < p < 1$ is always a specified number for which we are interested in ϕ_p.

We make the following assumptions about F, $\{G_x\}$ and ϕ_p:

(i) The range of X is bounded; for simplicity $0 \leqq X \leqq 1$.

(ii) F is continuous and strictly increasing.

(iia) For each x, $0 \leqq x \leqq 1$, G_x is continuous and strictly increasing.

(iii) ϕ_p is continuous (hence uniformly continuous).

(iv) For any given $\epsilon > 0$, there exists $\delta > 0$ (not depending on x) such that $|p' - p| < \delta$ implies $|\phi_p(x) - \phi_{p'}(x)| < \epsilon$ for all x.

It should be noted that when we say that a univariate distribution function H is strictly increasing, we mean thereby, that H is strictly increasing on an interval (a, b), where

$$\begin{aligned} a &= \text{the l.u.b. of the set } \{x:H(x) = 0\} \text{ if this set is non-empty} \\ &= -\infty \text{ otherwise,} \end{aligned}$$

and

$$\begin{aligned} b &= \text{the g.l.b. of the set } \{x:H(x) = 1\} \text{ if this set is non-empty} \\ &= \infty \text{ otherwise.} \end{aligned}$$

Another point to note is that condition (iv) is satisfied if there exists a function ψ on $[0, 1]$ such that $G_x(y) = G_0(y + \psi(x))$.

The methods suggested in this paper can be outlined as follows. Suppose we have nk independent observations on (X, Y). With the help of the X observations alone we divide this entire sample into k fractile groups, the 1st fractile group consisting of the samples with n smallest X observations, the 2nd fractile group consisting of the samples with the n smallest X observations among the rest, and so on. The extreme X observations in these fractile groups will divide the range of X into k disjoint random intervals. For large values of k, the lengths of all these random intervals will simultaneously become very small with high probability because of assumptions (i) and (ii). By assumption (iia), (iii) and (iv), the conditional quantile function on each of these small intervals will lie within a band contained in a small rectangle with high probability. On the other hand if the number of observations in each of these intervals is also large, the p-quantile of the Y observations in each fractile group plotted against any point in the corresponding fractile interval of X will be contained in the corre sponding small rectangle with high probability. This will allow us to estimate the conditional quantile function by a step-function. The hypothesis that the conditional quantile function is equal to a specified function will be accepted only if the data in each fractile group are found to be in agreement with the hypothesis. Thus if there is any deviation from the null hypothesis, it will be detected with high probability in at least one of these fractile groups.

$(X_1, Y_1), \cdots, (X_{nk}, Y_{nk})$ are independent observations on (X, Y). Let $X_{(1)} < \cdots < X_{(nk)}$ be the ordered values of $X_1, \cdots, X_{nk}$. $Y_{(i)} = Y_j$ if $X_{(i)} = X_j$. For $r = 1, \cdots, k$ and for $s = 1, \cdots, n$,

$$X_{(\overline{r-1}\cdot n+s)} = X_{r,s}, \qquad Y_{(\overline{r-1}\cdot n+s)} = Y_{r,s}.$$

For any given integer k, define a set of random intervals as follows:

$$I_{k1} = [0, X_{1,n}], I_{kr} = (X_{r-1,n}, X_{r,n}], \qquad r = 2, \cdots, k-1, \qquad \text{and}$$

$$I_{kk} = (X_{k-1,n}, 1].$$

Next let $Y_{r(1)} < \cdots < Y_{r(n)}$ be the ordered values of $Y_{r,1}, \cdots, Y_{r,n}$, and denote by $[a]$ the largest integer $\leqq a$. Now define a random step-function f_{nk} on $[0, 1]$ as follows: $f_{nk}(x) = Y_{r([np])}$ if $x \in I_{kr}$, $r = 1, \cdots, k$.

In course of the analysis carried out in the next two sections, we shall make use of an upper bound for the tail probabilities of sums of independent and bounded random variables due to Hoeffding [2]. We shall also make use of an upper bound for the error of approximation by the Central Limit Theorem and an upper bound of error involved in the usual approximation of the distribution function of "frequence χ^2" for a simple hypothesis regarding a multinomial distribution, both due to Esseen [1]. For the sake of completeness, these are stated below.

THEOREM (Hoeffding). *If $X_1, \cdots, X_n$ are independent random variables, $0 \leqq X_i \leqq 1$, $EX_i = \mu$, $S_n = X_1 + \cdots + X_n$, and if $0 < t < 1 - \mu$, then*

$$P[S_n - ES_n \geqq nt] \leqq e^{-2nt^2}. \tag{2}$$

THEOREM (Esseen). *If $X_1, \cdots, X_n$ are a sequence of independent random variables with the same distribution function F, the mean value zero, the variance $\sigma^2 \neq 0$ and the finite absolute moments $\beta_1, \cdots, \beta_\nu$ (ν is an integer $\geqq 3$), then for $|x| < \{(1+\delta)(\nu-2)\log n\}^{1/2}$,*

$$|F_n(x) - \Phi(x)| \leqq C_1(\delta, \beta)n^{-1/2}(1 + |x|^3)e^{-x^2/2} + C_2(\delta, \beta)n^{-(\nu-2)/2}, \tag{3}$$

where F_n is the distribution function of $(X_1 + \cdots + X_n)/n^{1/2}\sigma$, $\Phi(x) = \int_{-\infty}^{x} (2\pi)^{-1/2}e^{-t^2/2}\,dt$, $C_1(\delta, \beta)$ and $C_2(\delta, \beta)$ are finite constants, depending only on $0 < \delta < 1$, and the moments $\beta_1, \cdots, \beta_\nu$.

THEOREM (Esseen). *If n independent observations are taken on a multinomial distribution with $(m + 1)$ classes and with class probabilities $q_1, \cdots, q_{m+1}$, $\sum_1^{m+1} q_i = 1$, and if $n_1, \cdots, n_{m+1}$ are the respective observed frequencies, $\sum_1^{m+1} n_i = n$, then*

$$P\left[\sum_1^{m+1} (n_i - nq_i)^2/nq_i \leq \chi^2\right] = \frac{1}{2^{m/2}\Gamma(m/2)} \int_0^{\chi^2} e^{-w/2} w^{m/2-1}\, dw + \frac{\theta(q_1, \cdots, q_{m+1})}{n^{m/(m+1)}}, \tag{4}$$

where $\theta(q_1, \cdots, q_{m+1})$ *is a finite constant, depending only on* $q_1, \cdots, q_{m+1}$.

3. Convergence of f_{nk}. In this section we shall study the convergence in probability and almost sure convergence of f_{nk} to ϕ_p, uniformly. It is obvious that both n and k should tend to infinity for such convergences to take place, but the crucial point in our analysis is to find out how n and k should depend on each other as they tend to infinity.

We shall first prove a probability inequality for the event that the random variables $\{X_{rn}\}$, $r = 1, \cdots, k - 1$, lie in some neighborhoods of $F^{-1}(r/k)$, $r = 1, \cdots, k - 1$ respectively. The following lemma is a modification of a similar probability inequality given by Parthasarathy and Bhattacharya [4].

LEMMA 1. *Let* $0 < a_k < 1$, $k = 1, 2, \cdots$ *ad inf., be a sequence converging to zero. Under assumption* (ii),

$$P[F^{-1}(r/k - a_k) \leq X_{rn} \leq F^{-1}(r/k + a_k), r = 1, \cdots, k - 1] > 1 - 2ke^{-2nka_k^2}, \tag{5}$$

where $F^{-1}(a)$ *for* $a \leq 0$ *is defined to be* 0 *and* $F^{-1}(a)$ *for* $a \geq 1$ *is defined to be* 1.

PROOF. The left side of (5) is

$$\geq 1 - \sum_{r=1}^{k-1} P[X_{rn} < F^{-1}(r/k - a_k)] - \sum_{r=1}^{k-1} P[X_{rn} > F^{-1}(r/k + a_k)]$$

$$= 1 - \sum_{r=1}^{k-1} \sum_{s \geq rn} \binom{kn}{s} (r/k - a_k)^s (1 - r/k + a_k)^{kn-s} - \sum_{r=1}^{k-1} \sum_{s \geq kn - rn} \binom{kn}{s} (1 - r/k - a_k)^s (r/k + a_k)^{kn-s}$$

$$= 1 - \sum_{r=1}^{k-1} \sum_{s \geq kn[(r/k - a_k) + a_k]} \binom{kn}{s} (r/k - a_k)^s (1 - r/k + a_k)^{kn-s} - \sum_{r=1}^{k-1} \sum_{s \geq kn[(1 - r/k - a_k) + a_k]} \binom{kn}{s} (1 - r/k - a_k)^s (r/k + a_k)^{kn-s}$$

$$\geq 1 - 2(k-1)e^{-2nka_k^2} > 1 - 2ke^{-2nka_k^2}, \quad \text{by (2)}.$$

We now find a probability inequality for the event that the length of each of the intervals $I_{k1}, \cdots, I_{kk}$ is less than a specified positive number. Let $L(I)$ denote the length of an interval I on the real line.

LEMMA 2. *Under assumptions* (i) *and* (ii), *for any* $\delta > 0$,

$$P[L(I_{kr}) < \delta, r = 1, \cdots, k] > 1 - 2ke^{-2kna_k^2}$$

for sufficiently large k.

PROOF. We first note that under assumptions (i) and (ii), for any given $\delta > 0$, there exists an integer k_0 such that for $k > k_0$,

$$\text{(6)}\quad \max[F^{-1}(1/k + a_k), F^{-1}((r+1)/k + a_k) - F^{-1}(r/k - a_k), \quad r = 1, \cdots, k-2, \quad 1 - F^{-1}((k-1)/k - a_k)] < \delta.$$

We shall show that (6) along with

$$\text{(7)}\qquad F^{-1}(r/k - a_k) \leqq X_{r,n} \leqq F^{-1}(r/k + a_k), r = 1, \cdots, k-1,$$

implies $L(I_{kr}) < \delta, r = 1, \cdots, k$.

For $r = 2, \cdots, k-1$,

$$\begin{aligned} L(I_{kr}) &= X_{r,n} - X_{r-1,n} \\ &\leqq F^{-1}(r/k + a_k) - F^{-1}((r-1)/k - a_k), \quad \text{by (7)} \\ &< \delta, \quad \text{by (6)}. \end{aligned}$$

Also,

$$\begin{aligned} L(I_{k1}) &< X_{1,n} \\ &\leqq F^{-1}(1/k + a_k), \quad \text{by (7)} \\ &< \delta, \quad \text{by (6)}, \end{aligned}$$

and

$$\begin{aligned} L(I_{kk}) &\leqq 1 - F^{-1}((k-1)/k - a_k), \quad \text{by (7)} \\ &< \delta, \quad \text{by (6)}. \end{aligned}$$

An application of Lemma 1 now completes the proof.

LEMMA 3. *Under assumptions* (iia), (iii) *and* (iv), *for sufficiently large k and for any given $\eta > 0$,*

$$P[\text{Sup}_{x \epsilon I_{kr}} \phi_{p+\epsilon}(x) - \text{Inf}_{x \epsilon I_{kr}} \phi_{p-\epsilon}(x) < \eta, r = 1, \cdots, k] > 1 - 2ke^{-2nka_k^2},$$

if $\epsilon > 0$ is such that $|p' - p| < \epsilon$ implies $|\phi_p(x) - \phi_{p'}(x)| < \eta/3$ for all x. (By assumption (iv) such an $\epsilon > 0$ exists for any given $\eta > 0$.)

PROOF. Choose $\delta > 0$ such that $x_1 \epsilon [0, 1]$, $x_2 \epsilon [0, 1]$, $|x_1 - x_2| < \delta$ together imply $|\phi_p(x_1) - \phi_p(x_2)| < \eta/3$. By assumption (iii) such a $\delta > 0$ exists for any given $\eta > 0$. Now,

$$\begin{aligned} &P[\text{Sup}_{x \epsilon I_{kr}} \phi_{p+\epsilon}(x) - \text{Inf}_{x \epsilon I_{kr}} \phi_{p-\epsilon}(x) < \eta, r = 1, \cdots, k] \\ &\qquad \geqq P[\text{Sup}_{x \epsilon I_{kr}} \phi_{p+\epsilon}(x) - \text{Inf}_{x \epsilon I_{kr}} \phi_{p-\epsilon}(x) < \eta, L(I_{kr}) < \delta, r = 1, \cdots, k]. \end{aligned}$$

But

$$\begin{aligned} &P[\text{Sup}_{x \epsilon I_{kr}} \phi_{p+\epsilon}(x) - \text{Inf}_{x \epsilon I_{kr}} \phi_{p-\epsilon}(x) < \eta, \\ &\qquad\qquad r = 1, \cdots, k \mid L(I_{kr}) < \delta, r = 1, \cdots, k] \\ &\quad = P[\text{Sup}_{x_1, x_2 \epsilon I_{kr}} \{\phi_{p+\epsilon}(x_1) - \phi_{p-\epsilon}(x_2)\} < \eta, \\ &\qquad\qquad r = 1, \cdots, k \mid L(I_{kr}) < \delta, r = 1, \cdots, k] \end{aligned}$$

$$\geqq P[\mathrm{Sup}_{x_1, x_2 \epsilon I_{kr}} |\phi_p(x_1) - \phi_p(x_2)| < \eta/3, \mathrm{Sup}_{x \epsilon I_{kr}} |\phi_{p+\epsilon}(x) - \phi_p(x)| < \eta/3,$$

$$\mathrm{Sup}_{x \epsilon I_{kr}} |\phi_{p-\epsilon}(x) - \phi_p(x)| < \eta/3,$$

$$r = 1, \cdots, k \mid L(I_{kr}) < \delta, r = 1, \cdots, k] = 1,$$

by the choice of δ and ϵ. An application of Lemma 2 now completes the proof.

LEMMA 4. *Under assumptions* (iia), (iii) *and* (iv), *for arbitrary* $\epsilon > 0$ *and for large* n,

$$P[\mathrm{Inf}_{x \epsilon I_{kr}} \phi_{p-\epsilon}(x) < Y_{r([np])} < \mathrm{Sup}_{x \epsilon I_{kr}} \phi_{p+\epsilon}(x),$$

$$r = 1, \cdots, k \mid X_1 = x_1, \cdots, X_{nk} = x_{nk}] \geqq 1 - 2ke^{-n\epsilon^2},$$

for any given $x_1, \cdots, x_{nk}$ *in* [0, 1].

PROOF. Suppose $x_1, \cdots, x_n$ are points in an interval $[a, b] \subset [0, 1]$. Let $Y_1, \cdots, Y_n$ be mutually independent random variables, Y_i having distribution function G_{x_i}, $i = 1, \cdots, n$. Also let $Y_{(1)} < \cdots < Y_{(n)}$ be the ordered values of $Y_1, \cdots, Y_n$. Then

$$P[Y_{([pn])} \leqq \mathrm{Inf}_{x \epsilon [a,b]} \phi_{p-\epsilon}(x)]$$

$$= P[\text{at least } [np] \text{ of } Y_1, \cdots, Y_n \text{ are } \leqq \mathrm{Inf}_{x \epsilon [a,b]} \phi_{p-\epsilon}(x)]$$

$$= P[U_1 + \cdots U_n \geqq [np] - \sum_{i=1}^{n} G_{x_i}(\mathrm{Inf}_{x \epsilon [a,b]} \phi_{p-\epsilon}(x))],$$

where $U_1, \cdots, U_n$ are mutually independent random variables with

$$P[U_i = 1 - G_{x_i}(\mathrm{Inf}_{x \epsilon [a,b]} \phi_{p-\epsilon}(x))] = G_{x_i}(\mathrm{Inf}_{x \epsilon [a,b]} \phi_{p-\epsilon}(x))$$

and

$$P[U_i = -G_{x_i}(\mathrm{Inf}_{x \epsilon [a,b]} \phi_{p-\epsilon}(x))]$$

$$= 1 - G_{x_i}(\mathrm{Inf}_{x \epsilon [a,b]} \phi_{p-\epsilon}(x)), \qquad i = 1, \cdots, n.$$

Now,

$$G_{x_i}(\mathrm{Inf}_{x \epsilon [a,b]} \phi_{p-\epsilon}(x))$$

$$\leqq G_{x_i}(\phi_{p-\epsilon}(x_i)), \text{ since } x_i \epsilon [a, b],$$

$$= p - \epsilon \text{ by } (1).$$

Hence

$$[np] - \sum_{i=1}^{n} G_{x_i}(\mathrm{Inf}_{x \epsilon [a,b]} \phi_{p-\epsilon}(x))$$

$$\geqq np - 1 - n(p - \epsilon) = n\epsilon - 1 > n\epsilon/2^{\frac{1}{2}} \quad \text{for large } n.$$

Thus,

$$P[U_1 + \cdots + U_n \geqq [np] - \sum_{i=1}^{n} G_{x_i}(\mathrm{Inf}_{x\epsilon[a,b]}\, \phi_{p-\epsilon}(x))]$$
$$\leqq P[U_1 + \cdots + U_n \geqq n\epsilon/2^{\frac{1}{2}}] \leqq e^{-n\epsilon^2}, \quad \text{by (2)}.$$

Similarly, $P[Y_{([np])} \geqq \mathrm{Sup}_{x\epsilon[a,b]}\, \phi_{p+\epsilon}(x)] \leqq e^{-n\epsilon^2}$.

The desired probability inequality is now obtained.

We are now in a position to prove the following theorem about uniform convergence of f_{nk} to ϕ_p.

THEOREM 1. *If F, $\{G_x\}$ and ϕ_p satisfy conditions* (i)–(iv), *then for $n \geqq k^\gamma$, $\gamma > 0$, $d_{nk} = \mathrm{Sup}_{0 \leqq x \leqq 1} |f_{nk}(x) - \phi_p(x)|$ converges to zero in probability as $k \to \infty$ and for $n \geqq k$, d_{nk} converges to zero with probability one as $k \to \infty$.*

PROOF. Let $\eta > 0$ be given.

$$\begin{aligned} P[\mathrm{Sup}_{0\leqq x\leqq 1} &|f_{nk}(x) - \phi_p(x)| < \eta] \\ &\geqq P[\mathrm{Inf}_{x\epsilon I_{kr}}\, \phi_{p-\epsilon}(x) < Y_{r([np])} < \mathrm{Sup}_{x\epsilon I_{kr}}\, \phi_{p+\epsilon}(x), \mathrm{Sup}_{x\epsilon I_{kr}}\, \phi_{p+\epsilon}(x) \\ &\qquad - \mathrm{Inf}_{x\epsilon I_{kr}}\, \phi_{p-\epsilon}(x) < \eta, r = 1, \cdots, k], \text{ for arbitrary } \epsilon > 0, \\ &= P[\mathrm{Inf}_{x\epsilon I_{kr}}\, \phi_{p-\epsilon}(x) < Y_{r([np])} < \mathrm{Sup}_{x\epsilon I_{kr}}\, \phi_{p+\epsilon}(x), \\ &\qquad r = 1, \cdots, k \mid \mathrm{Sup}_{x\epsilon I_{kr}}\, \phi_{p+\epsilon}(x) - \mathrm{Inf}_{x\epsilon I_{kr}}\, \phi_{p-\epsilon}(x) < \eta, r = 1, \cdots, k] \\ &\qquad \times P[\mathrm{Sup}_{x\epsilon I_{kr}}\, \phi_{p+\epsilon}(x) - \mathrm{Inf}_{x\epsilon I_{kr}}\, \phi_{p-\epsilon}(x) < \eta, r = 1, \cdots, k] \\ &> (1 - 2ke^{-n\epsilon^2})\cdot(1 - 2ke^{-2nka_k^2}), \end{aligned}$$

by Lemmas 3 and 4 if $\epsilon > 0$ is so chosen as to satisfy the condition of Lemma 3. Both the factors in the last term tend to one as $k \to \infty$ if $n \geqq k^\gamma$, $\gamma > 0$ and $a_k = k^{-\gamma/2}$, which proves the first part of the theorem. Again,

$$\sum_{k=1}^{\infty} [1 - (1 - 2ke^{-n\epsilon^2})(1 - 2ke^{-2nka_k^2})]$$

converges if $n \geqq k$ and $a_k = k^{-\frac{1}{2}}$. The second part of the theorem now follows from Borel-Cantelli lemma.

If we have $0 < p_1 < \cdots < p_m < 1$, and if we define $f_{nk}^{(i)}(x) = Y_{r([np_i])}$ for $x \,\epsilon\, I_{kr}$, $r = 1, \cdots, k$, $i = 1, \cdots, m$, then Theorem 1 can be immediately extended for the simultaneous uniform convergence of $f_{nk}^{(1)}, \cdots, f_{nk}^{(m)}$ to $\phi_{p_1}, \cdots, \phi_{p_m}$ respectively.

4. Large sample tests for specified conditional quantile functions. Let μ be a specified real-valued function on [0, 1] and consider the problem of testing the hypothesis $H_0{:}\phi_p = \mu$ against the alternative $H_1{:}\phi_p \neq \mu$.

We define random variables $\{U_{rs}\}$, $r = 1, \cdots, k$; $s = 1, \cdots, n$, as follows:

$$\begin{aligned} U_{rs} &= 1 \quad \text{if } Y_{rs} \leqq \mu(X_{rs}) \\ &= 0 \quad \text{otherwise.} \end{aligned}$$

These are mutually independent random variables and under H_0, each of them takes values 1 and 0 with probabilities p and $1 - p$ respectively. Let $\bar{U}_r = \sum_{s=1}^{n} U_{rs}/n$, $r = 1, \cdots, k$, and

$$\tau_{nk} = \operatorname{Sup}_{r=1,\cdots,k} |\bar{U}_r - p|/\{p(1-p)/n\}^{\frac{1}{2}}.$$

If we can find the limiting distribution of τ_{nk} (suitably standardized) under H_0, then we can test H_0 at any given level of significance $0 < \alpha < 1$, as follows: Reject H_0 if and only if $\tau_{nk} > \tau_{nk}(\alpha)$, where $\lim_{k\to\infty} P[\tau_{nk} \leqq \tau_{nk}(\alpha) \mid H_0] = 1 - \alpha$.

Again, suppose $0 < p_1 < \cdots < p_m < 1$ are given numbers and $\mu_1 \cdots, \mu_m$ are specified real-valued functions on $[0, 1]$, satisfying the condition that $\mu_1(x) < \cdots < \mu_m(x)$ for all $x \epsilon [0, 1]$. Consider the problem of testing the hypothesis $H_0^{(m)}: (\phi_{p_1} = \mu_1, \cdots, \phi_{p_m} = \mu_m)$ against the alternative $H_1^{(m)}$ that $\phi_{p_i} \neq \mu_i$ for at least one i, $i = 1, \cdots, m$.

Define

$$\begin{aligned} U_{rs}^{(1)} &= 1 \quad \text{if } Y_{rs} \leqq \mu_1(X_{rs}) \\ &= 0 \quad \text{otherwise,} \\ U_{rs}^{(i)} &= 1 \quad \text{if } \mu_{i-1}(X_{rs}) < Y_{rs} \leqq \mu_i(X_{rs}) \\ &= 0 \quad \text{otherwise,} \quad i = 2, \cdots, m, \\ U_{rs}^{(m+1)} &= 1 \quad \text{if } Y_{rs} > \mu_m(X_{rs}) \\ &= 0 \quad \text{otherwise,} \end{aligned}$$

and $U_{r0}^{(i)} = \sum_{s=1}^{n} U_{rs}^{(i)}$, $r = 1, \cdots, k$, $i = 1, \cdots, m + 1$. Let $q_1 = p_1$, $q_i = p_i - p_{i-1}$, $i = 2, \cdots, m$; $q_{m+1} = 1 - p_m$, and consider the statistic

$$\tau_{nk}^{(m)} = \operatorname{Sup}_{r=1,\cdots,k} \sum_{i=1}^{m+1} [(U_{r0}^{(i)} - nq_i)^2/nq_i].$$

If we can find the limiting distribution of $\tau_{nk}^{(m)}$ (suitably standardized) under $H_0^{(m)}$, then we can test $H_0^{(m)}$ with the help of $\tau_{nk}^{(m)}$ in the same way as we hope to test H_0 with the help of τ_{nk}.

The following theorem gives us the limiting distributions of τ_{nk} and $\tau_{nk}^{(m)}$ under H_0 and $H_0^{(m)}$ respectively.

For any real θ and for any integer k, define

$$\begin{aligned} \lambda_k(\theta) &= [2(\theta + \log k - \tfrac{1}{2} \log\log k)]^{\frac{1}{2}} \quad \text{if} \quad \theta + \log k - \tfrac{1}{2}\log\log k > 0 \\ &= 0 \quad \text{otherwise,} \end{aligned}$$

and

$$\begin{aligned} \lambda_k^{(m)}(\theta) &= 2[\theta + \log k + (m/2 - 1)\log\log k] \\ &\qquad \text{if} \quad \theta + \log k + (m/2 - 1)\log\log k > 0 \\ &= 0 \quad \text{otherwise.} \end{aligned}$$

THEOREM 2.

(a) *If* $n \geqq k^{\gamma}, \gamma > 0$, *then*

$$\lim_{k\to\infty} P[\tau_{nk} \leqq \lambda_k(\theta) \mid H_0] = \exp[-\pi^{-\frac{1}{2}} e^{-\theta}].$$

(b) *If* $n \geqq (k \log k)^{1+1/m}$, *then*

$$\lim_{k\to\infty} P[\tau_{nk}^{(m)} \leqq \lambda_k^{(m)}(\theta) \mid H_0^{(m)}] = \exp[-(\Gamma(m/2))^{-1} e^{-\theta}].$$

PROOF.

(a) Under H_0, $\{U_{rs}\}$ are independently and identically distributed with mean p, variance $p(1-p)$, and with finite moments of all orders which are functions of p only. Let ν be so large that

$$\lambda_k(\theta) < \{(1+\delta)(\nu-2)\gamma \log k\}^{\frac{1}{2}}, \quad 0 < \delta < 1, \quad \text{and} \tag{8}$$

$$\gamma(\nu - 2) > 2. \tag{9}$$

Making use of (8) and (3) and denoting $C = \max\{C_1(\delta, \beta), C_2(\delta, \beta)\}$ in (3), we get

$$P\left[\frac{|\bar{U}_r - p|}{\{p(1-p)/n\}^{\frac{1}{2}}} \leqq \lambda_k(\theta) \mid H_0\right] = \Phi(\lambda_k(\theta) - \Phi(-\lambda_k(\theta)) + e_{kr},$$

where

$$|e_{kr}| \leqq C[|1 + (\lambda_k(\theta))^3| n^{-\frac{1}{2}} \exp[-(\lambda_k(\theta))^2/2] + n^{-(\nu-2)/2}]$$
$$\leqq C'[e^{-\theta}(\log k)^2 k^{-(1+\gamma/2)} + k^{-\gamma(\nu-2)/2}],$$

since $n \geqq k^{\gamma}$, C' being a finite constant, depending on p and δ. Now,

$$\log P[\tau_{nk} \leqq \lambda_k(\theta) \mid H_0] = \sum_{r=1}^{k} \log P\left[\frac{|\bar{U}_r - p|}{\{p(1-p)/n\}^{\frac{1}{2}}} \leqq \lambda_k(\theta) \mid H_0\right]$$
$$= \sum_{r=1}^{k} \log[\Phi(\lambda_k(\theta)) - \Phi(-\lambda_k(\theta)) + e_{kr}]$$
$$= k \log[\Phi(\lambda_k(\theta)) - \Phi(-\lambda_k(\theta))] + Z_k,$$

where $|Z_k| \leqq \sum_{r=1}^{k} |\log[1 + e_{kr}/[\Phi(\lambda_k(\theta)) - \Phi(-\lambda_k(\theta))]]|$. Since $e_{kr} \to 0$ and $\Phi(\lambda_k(\theta)) - \Phi(-\lambda_k(\theta)) \to 1$ as $k \to \infty$, for sufficiently large k,

$$|e_{kr}|/[\Phi(\lambda_k(\theta)) - \Phi(-\lambda_k(\theta))] < \tfrac{1}{2}.$$

Now $\log(1 + x) = x + vx^2$, $|v| < 1$ for $|x| < \frac{1}{2}$. Hence,

$$|Z_k| \leqq C'' \cdot k[e^{-\theta}(\log k)^2 k^{-(1+\gamma/2)} + k^{-\gamma(\nu-2)/2}],$$

where C'' is a finite constant depending on p and δ. It now follows from (9) that $Z_k \to 0$ as $k \to \infty$. To complete the proof we have only to verify that

$$\lim_{k\to\infty} k \log[\Phi(\lambda_k(\theta)) - \Phi(-\lambda_k(\theta))] = -\pi^{-\frac{1}{2}} e^{-\theta}.$$

(b) It can be shown in the same way as in proving (a) that

$$\log [\tau_{nk}^{(m)} \leqq \lambda_k^{(m)}(\theta) \mid H_0^{(m)}] = k \log \left[\frac{1}{2^{m/2}\Gamma(m/2)} \int_0^{\lambda_k^{(m)}(\theta)} e^{-w/2} w^{m/2-1}\, dw \right] + Z_k^{(m)},$$

where $\lim_{k \to \infty} Z_k^{(m)}$ can be shown to be zero if $n \geqq (k \log k)^{1+1/m}$, by an application of (4). The rest of the proof follows from the fact that

$$\lim_{k \to \infty} k \log \left[\frac{1}{2^{m/2}\Gamma(m/2)} \int_0^{\lambda_k^{(m)}(\theta)} e^{-w/2} w^{m/2-1}\, dw \right] = -\frac{1}{\Gamma(m/2)} e^{-\theta}$$

For any $0 < \alpha < 1$, let

$$\tau_{nk}(\alpha) = [\log (k^2/\pi) - \log \log k - 2 \log \log (1 - \alpha)^{-1}]^{\frac{1}{2}}$$

and

$$\tau_{nk}^{(m)}(\alpha) = 2\{\log (k/\Gamma(m/2)) + (m/2 - 1)\log \log k - \log \log (1 - \alpha)^{-1}\}.$$

The test for $H_0 : \varphi_p = \mu$ can now be explained graphically as follows: (a) Plot the graph of the function μ on the (X, Y) plane along with the sample points $(X_1, Y_1), \cdots, (X_{nk}, Y_{nk})$. (b) Of the n samples in the rth fractile group, find the proportion of Y observations below the graph of μ (this is $\bar{U}_r$). (c) Compute $|\bar{U}_r - p|\{n/p(1 - p)\}^{\frac{1}{2}}$ for $r = 1, \cdots, k$. The largest of these quantities is the test statistic τ_{nk}. (d) Reject H_0 if $\tau_{nk} > \tau_{nk}(\alpha)$. The test for

$$H_0^{(m)} : (\varphi_{p_1} = \mu_1, \cdots, \varphi_{p_m} = \mu_m)$$

can also be explained similarly. Here we plot all the functions $\mu_1, \cdots, \mu_m$ along with the sample points and for each fractile group we find the proportion of Y observations in each of the $m + 1$ regions into which the (X, Y) plane is divided by the graphs of $\mu_1, \cdots, \mu_m$. This gives us $U_{r0}^{(i)}$ for $i = 1, \cdots, m + 1$ and $r = 1, \cdots, k$. For each r, we compute the frequency χ^2 from observed frequencies $U_{r0}^{(1)}, \cdots, U_{r0}^{(m+1)}$ and expected frequencies $np_1, n(p_2 - p_1), \cdots, n(p_m - p_{m-1}), n(1 - p_m)$ respectively to measure the discrepancy between the hypothesis and the data in the rth fractile group. The largest of these quantities is test statistic $\tau_{nk}^{(m)}$ and we reject $H_0^{(m)}$ if $\tau_{nk}^{(m)} > \tau_{nk}^{(m)}(\alpha)$. The following theorem asserts that these tests are asymptotically of level α and that they are consistent.

THEOREM 3. *Under assumptions* (i)–(iv),

(a) *If* $n \geqq k^\gamma, \gamma > 0$, *then the test with critical region* $\tau_{nk} > \tau_{nk}(\alpha)$, $0 < \alpha < 1$, *is a large sample size* α *test for the null hypothesis* H_0, *and is consistent against the alternative* H_1.

(b) *If* $n \geqq (k \log k)^{1+1/m}$, *then the test with critical region* $\tau_{nk}^{(m)} > \tau_{nk}^{(m)}(\alpha)$, $0 < \alpha < 1$, *is a large sample size* α *test for the null hypothesis* $H_0^{(m)}$, *and is consistent against the alternative* $H_1^{(m)}$.

PROOF. The first parts of (a) and (b) are immediate consequences of Theorem 2. We shall prove here the second part of (b), and the proof for the second part of (a) will follow on exactly similar lines.

Since we are assuming $\phi_{p_i}, i = 1, \cdots, m$ to be continuous, we need consider only the case when $\mu_1, \cdots, \mu_m$ are continuous. Suppose $H_0^{(m)}$ does not hold. Then $\phi_{p_i} \neq \mu_i$ for at least one integer i between 1 and m. Let j be the smallest integer for which $\phi_{p_j} \neq \mu_j$. Since ϕ_{p_j} and μ_j are both continuous, there exists some $\delta > 0$ and an interval $(c_1, c_2) \subset [0, 1]$, such that either $\mu_j(x) > \phi_{p_j}(x) + \delta$ for all $x \,\varepsilon\, (c_1, c_2)$ or $\mu_j(x) < \phi_{p_j}(x) - \delta$ for all $x \,\varepsilon\, (c_1, c_2)$. Suppose the first is true. (The other case can be treated similarly.) By assumption (iv), for given $\delta > 0$, there exists an $\epsilon > 0$ such that $p_j < p < p_j + 2\epsilon$ implies $\phi_p(x) - \phi_{p_j}(x) < \delta$ for all x. Choose $p = p_j + \epsilon$. Then $\phi_{p_j+\epsilon}(x) < \phi_{p_j}(x) + \delta$. Hence, $G_x(\phi_{p_j}(x) + \delta) > G_x(\phi_{p_j+\epsilon}(x)) = p_j + \epsilon$.

For convenience of notation, let $p_0 = 0$, and $\phi_{p_0}(x) = \mu_0(x) = -\infty$ for all x. Now for any $x \,\varepsilon\, (c_1, c_2)$,

$$\begin{aligned} P[U_{r,s}^{(j)} = 1 \mid X_{r,s} = x] &= P[\mu_{j-1}(x) < Y_{rs} \leqq \mu_j(x) \mid X_{rs} = x] \\ &= P[\phi_{p_{j-1}}(x) < Y_{rs} \leqq \mu_j(x) \mid X_{rs} = x], \text{ since } j \text{ is} \\ &\qquad \text{the smallest integer for which } \phi_{p_j} \neq \mu_j \\ &\geqq P[\phi_{p_{j-1}}(x) < Y_{rs} \leqq \phi_{p_j}(x) + \delta \mid X_{rs} = x] \\ &= G_x(\phi_{p_j}(x) + \delta) - G_x(\phi_{p_{j-1}}(x)) \\ &> p_j + \epsilon - p_{j-1} = q_j + \epsilon. \end{aligned}$$

Now let $c_1'' < c_1' < c_2' < c_2''$ be points in $(c_1, (c_1 + c_2)/2)$ and choose k_0 sufficiently large such that for $k \geqq k_0$, $F(c_1) > 1/k$, $1 - F(c_2) > 1/k$, $F(c_2') - F(c_1') > 1/k$, $F(c_1') - F(c_1'') > a_k$ and $F(c_2'') - F(c_2') > a_k$, where $\{a_k\}$ is a sequence as in Lemma 1. Then for $k \geqq k_0$, $F^{-1}(r/k) \leqq c_1' < c_2' \leqq F^{-1}((r+1)/k)$ implies $F(c_2') - F(c_1') \leqq 1/k$, which is a contradiction. Hence for each $k \geqq k_0$, there exists at least one integer $r(k)$ such that $c_1' < F^{-1}(r(k)/k) < c_2'$. Also for $k \geqq k_0$, $F^{-1}(r(k)/k - a_k) \leqq c_1''$ implies $F(c_1') - F(c_1'') \leqq a_k$; which is a contradiction, and $F^{-1}(r(k)/k + a_k) \geqq c_2''$ implies $F(c_2'') - F(c_2') \leqq a_k$ which is a contradiction. Hence for each $k \geqq k_0$, there exists a smallest integer $r_1(k)$ such that

$$c_1 < F^{-1}(r_1(k)/k - a_k) < F^{-1}(r_1(k)/k + a_k) < (c_1 + c_2)/2.$$

Similarly, for each value of k greater than some integer, there exists a largest integer $r_2(k) > r_1(k)$ such that

$$(c_1 + c_2)/2 < F^{-1}(r_2(k)/k - a_k) < F^{-1}(r_2(k)/k + a_k) < c_2.$$

Hence for such large values of k,

$$F^{-1}(r_1(k)/k - a_k) < X_{r_1(k),n} < F^{-1}(r_1(k)/k + a_k)$$

and

$$F^{-1}(r_2(k)/k - a_k) < X_{r_2(k),n} < F^{-1}(r_2(k)/k + a_k)$$

imply that $I_{kr} \subset (c_1, c_2)$ for $r = r_1(k) + 1, \cdots, r_2(k)$, and this set of integers is non-empty. Thus for sufficiently large values of k,

$$P[I_{kr} \subset (c_1, c_2), r = r_1(k) + 1, \cdots, r_2(k)]$$
$$\geqq P[F^{-1}(r_i(k)/k - a_k) < X_{r_i(k),n} < F^{-1}(r_i(k)/k + a_k), i = 1, 2]$$
$$\geqq 1 - 2ke^{-2nka_k^2}, \quad \text{by Lemma 1.}$$

For each k greater than some sufficiently large integer, now choose and fix an integer $r(k)$ such that $r_1(k) + 1 \leqq r(k) \leqq r_2(k)$. Then

$$P[\tau_{nk}^{(m)} > \tau_{nk}^{(m)}(\alpha)] \geqq P[\tau_{nk}^{(m)} > \tau_{nk}^{(m)}(\alpha), I_{k,r(k)} \subset (c_1, c_2)]$$
$$\geqq P[\tau_{nk}^{(m)} > \tau_{nk}^{(m)}(\alpha) \mid I_{k,r(k)} \subset (c_1, c_2)] \times [1 - 2ke^{-2nka_k^2}]$$
$$\geqq P[(U_{r(k),0}^{(j)} - nq_j)^2/nq_j > \tau_{nk}^{(m)}(\alpha) \mid I_{k,r(k)} \subset (c_1, c_2)] \times [1 - 2ke^{-2nka_k^2}]$$
$$\geqq P[U_{r(k),0}^{(j)} > nq_j + \{nq_j\tau_{nk}^{(m)}(\alpha)\}^{1/2} \mid I_{k,r(k)} \subset (c_1, c_2)] \times [1 - 2ke^{-2nka_k^2}]$$
$$\geqq [1 - \exp(-2n[\epsilon - \{q_j\tau_{nk}^{(m)}(\alpha)/n\}]^2)] \times [1 - 2ke^{-2nka_k^2}],$$

by (2). To complete the proof we have only to note that if $n \geqq (k \log k)^{1+1/m}$, then with the choice of $a_k = k^{-\frac{1}{2}}$, both factors in the final expression tend to 1.

5. Generalization to the case when X is vector-valued. Suppose $(X, Y) = (X_1, \cdots, X_h, Y)$ follows an unknown multivariate distribution. Let F be the distribution function of $X = (X_1, \cdots, X_h)$ and for any $x = (x_1, \cdots, x_h)$, let G_x be the conditional distribution of Y given $X = x$. For given $0 < p < 1$, let $\phi_p(x)$ be the p-quantile of G_x.

The assumptions we make about F, $\{G_x\}$ and ϕ_p are the same as (i)–(iv) given in Section 2, with some modifications. We shall only state the modified version of assumption (ii), the modifications on the other assumptions being obvious.

Modified assumption (ii). The marginal distribution of X_1 and the conditional distribution of X_i given $X_1 = x_1, \cdots, X_{i-1} = x_{i-1} (0 \leqq x_j \leqq 1, j = 1, \cdots, i - 1), i = 2, \cdots, h$, are all continuous and strictly increasing.

In Sections 2, 3 and 4, for a sample of size nk from a bivariate (X, Y), we defined a random division of [0, 1] into k sub-intervals with the help of the fractiles of X observations. Studying the convergence of the sample fractiles to the corresponding population fractiles, we were able to give a probability inequality for the event that the length of each of these random intervals is less than a specified quantity. Due to the uniform continuity and some other regular nature of ϕ_p, the rest was accomplished by investigating the behavior of the p-quantile of the Y-observations corresponding to the X-observations belonging to each one of these random intervals, separately.

When X is vector-valued, there is only a partial ordering among the X's. We therefore modify our procedure as follows. Let $(X_{11}, \cdots, X_{h1}, Y_1), \cdots, (X_{1N}, \cdots, X_{hN}, Y_N)$ be N independent observations on $(X_1, \cdots, X_h, Y)$ where $N = nk^h$. Let $N_j = nk^{h-j}$, $j = 1, \cdots, h$. First, let us arrange the X_1-

coordinates of all observations in increasing order of magnitude and let $X_{1(i)}$ be the ith order statistic obtained from $X_{11}, \cdots, X_{1N}$. Let us divide the interval $[0, 1]$ into random sub-intervals as follows. $I_1 = [0, X_{1(N_1)}]$, $I_{r_1} = (X_{1(\overline{r_1-1}\cdot N_1)}, X_{1(r_1 N_1)}]$, $r_1 = 2, \cdots, k-1$, and $I_k = (X_{1(\overline{k-1}\cdot N_1)}, 1]$. For each $r_1 = 1, \cdots, k$, consider all samples $(X_{1i}, \cdots, X_{hi}, Y_i)$ for which $X_{1i} \varepsilon I_{r_1}$ and call them the samples belonging to the r_1th fractile group. Now for each $j = 1, \cdots, h-1$ and for each $r_1 = 1, \cdots, k; \cdots; r_j = 1, \cdots, k$, arrange the X_{j+1}-coordinates of all samples belonging to the $(r_1, \cdots, r_j)$th fractile group, in increasing order of magnitude, and with the N_{j+1}th, $2N_{j+1}$th, $\cdots$, $(k-1)N_{j+1}$th order statistics obtained from the X_{j+1}-coordinates of these samples, divide the interval $[0, 1]$ into random sub-intervals in the same way as $I_1, \cdots, I_k$ were defined, and call these intervals $I_{r_1 \cdots r_j 1}, \cdots, I_{r_1 \cdots r_j k}$. For each $r_1 = 1, \cdots, k; \cdots; r_{j+1} = 1, \cdots, k$, consider all samples $(X_{1i}, \cdots, X_{hi}, Y_i)$ which belong to the $(r_1, \cdots, r_j)$th fractile group and for which $X_{j+1,i} \varepsilon I_{r_1 \cdots r_j r_{j+1}}$, as belonging to the $(r_1, \cdots, r_j, r_{j+1})$th fractile group. Finally, when for each $r_1 = 1, \cdots, k; \cdots, r_h = 1, \cdots, k$, we have exactly n observations belonging to the $(r_1, \cdots, r_h)$th fractile group, let us arrange the Y-coordinates of all samples belonging to the $(r_1, \cdots, r_h)$th fractile group, in increasing order of magnitude, and suppose the order statistics obtained from the Y-coordinates of these samples are, $Y_{r_1 \cdots r_h(1)} < \cdots < Y_{r_1 \cdots r_h(n)}$. Also, for any specified real valued function μ defined on the h times Cartesian product of $[0, 1]$ with itself, let

$$U_i(r_1, \cdots, r_h) = 1 \quad \text{if } (X_{1i}, \cdots, X_{hi}, Y_i) \,\varepsilon\, I_{r_1} \times I_{r_1 r_2} \times \cdots \times I_{r_1 r_2 \cdots r_h}, \quad \text{and if } Y_i \leqq \mu(X_{1i}, \cdots, X_{hi})$$
$$= 0 \quad \text{otherwise,}$$

and let $\bar{U}(r_1, \cdots, r_h) = \sum_{i=1}^{N} U_i(r_1, \cdots, r_h)/n$.

Now we define a random function

$$f_{nk}(x_1, \cdots, x_h) = Y_{r_1 \cdots r_h([np])} \text{ if } x_1 \varepsilon I_{r_1}, x_2 \varepsilon I_{r_1 r_2}, \cdots, x_h \varepsilon I_{r_1 r_2 \cdots r_h},$$
$$r_1 = 1, \cdots, k; \cdots, r_h = 1, \cdots, k,$$

and a statistic

$$\tau_{nk} = \operatorname*{Sup}_{\substack{r_1 = 1, \cdots, k; \cdots, \\ r_h = 1, \cdots, k.}} \frac{|\bar{U}(r_1, \cdots, r_h) - p|}{\{p(1-p)/n\}^{\frac{1}{2}}}.$$

It can be easily verified that statements regarding the confergence of f_{nk} to ϕ_p, the limiting distribution of τ_{nk} (properly standardized) under the null hypothesis $H_0 : (\phi_p = \mu)$ and the consistency of the test "reject H_0 if and only if $\tau_{nk} > \tau_{nk}(\alpha)$, where $\tau_{nk}(\alpha)$, $0 < \alpha < 1$, is such that $\lim_{k\to\infty} P[\tau_{nk} \leqq \tau_{nk}(\alpha) \mid H_0] = 1 - \alpha$", made as in Theorems 1, 2 and 3 with k replaced by $K = k^h$, are valid.

For the case of several conditional quantile functions, the results of Sections 3 and 4 can be extended for the multivariate case in the same way.

6. Concluding remarks.

I. The methods discussed in this paper are designed for the situation when no definite prior information is available about the conditional quantile functions. The conditional quantile function is studied piecewise over each of a large number of small intervals, and as a result the inference about each such piece is based on the Y observations of only a small fraction of the total sample. If it were known for certain that the conditional quantile function belongs to a parametric family, we could use the total sample to make inference about these parameters and would do much better. However if this knowledge were not certain but only of an approximate nature, then the methods discussed in this paper would be asymptotically much better than the methods based on such approximate knowledge.

II. Suppose we have N independent observations available on (X, Y) where N is large. How should N be factorized into k and n (the number of fractile groups and the number of samples in each fractile group)? From the theorems about asymptotic properties of the proposed estimates and tests we only know that $n = N/k$ should be at least as large as k^γ for some $\gamma > 0$, or k, or $(k \log k)^{1+1/m}$ as the problem may be. The following scrutiny of the proofs of the theorems in this paper might help to make the choice of k and n more specific:

(a) A large value of k is needed to make the fractile intervals of X small so that the variation of the conditional quantile function over each of these intervals is small (Lemmas 2, 3 and Theorem 3(b)). Large k is also needed to approximate $k \log [\Phi(\lambda_k(\theta)) - \Phi(-\lambda_k(\theta))]$ by $-e^{-\theta}/(\pi)^{\frac{1}{2}}$ (Theorem 2), but this is not so important, because we could just as well compare τ_{nk} with the appropriate percentage point of the distribution function $[\Phi(x) - \Phi(-x)]^k$, for $x > 0$.

(b) A large n is needed primarily for the purpose of stabilizing the p-quantile of the Y observations in each fractile group. The rate of this stabilization depends on the (ϵ, δ) relation in condition (iv). In the simple case when there exists a function ψ on $[0, 1]$ such that $G_x(y) = G_0(y + \psi(x))$, where G_0 has a continuous density function g_0 with respect to Lebesgue measure, this (ϵ, δ) relation for small values of ϵ will depend largely on $g_0(\varphi_p(0))$. The smaller this quantity is, the slower is the rate of convergence and the larger should be the value of n (Lemma 4 and Theorem 3(b)). This also agrees with the heuristic reasoning that the asymptotic variance of a normalized sample quantile is inversely proportional to the square of the probability density at the corresponding population quantile.

Thus the proper choice of k and n is a matter of striking a balance between the two objectives mentioned above. For this, any prior knowledge about the variation of the conditional quantile function and the conditional densities at the population conditional quantiles will be useful. A preliminary study of the data may also be of some help.

7. Acknowledgment. The author is indebted to S. N. Roy and Sudhindra N. Ray for some discussions. The suggestions received from the referees and the editor in preparing the final version of the paper are also gratefully acknowledged.

REFERENCES

[1] ESSEEN, C. G. (1945). Fourier analysis of distribution functions. A mathematical study of the Laplace-Gaussian law. *Acta Math.* 77 1-125.

[2] HOEFFDING, WASSILY (1963). Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.* 58 13-30.

[3] MAHALANOBIS, P. C. (1958). Lectures in Japan: Fractile graphical analysis. Indian Statist. Inst. Monograph.

[4] PARTHASARATHY, K. R. and BHATTACHARYA, P. K. (1961). Some limit theorems in regression theory. *Sankhyā*, Series A 23 91-102.

[5] ROY, S. N. and COBB, WHITFIELD (1960). Mixed model variance analysis with normal error and possibly non-normal other random effects: Part I: The univariate case. *Ann. Math. Statist.* 31 939-957.

[6] SATHE, Y. S. (1962). Studies of some problems in nonparametric inference. Institute of Statistics, University of North Carolina, Mimeograph Series No. 325.