Sankhyā : The Indian Journal of Statistics
1986, Volume 48, Series A, Pt. 2, pp. 142–149.

ESTIMATION OF DENSITY QUANTILE FUNCTION

By GUTTI JOGESH BABU

Indian Statistical Institute

SUMMARY. Let F be a distribution function with density f at p-th quantile $F^{-1}(p)$. Some results on estimation of density quantile are obtained in this paper. The density quantile function is estimated uniformly in an interval and almost sure bounds for such estimators are obtained in the dependent case also.

1. Introduction

Let F be a distribution function with density f. Parzen (1979) attaches great importance to the *density quantile function* $f(F^{-1}(t))$ in statistical data modelling, when he proposes, "... greater insight will be obtained by formulating conclusions in terms of qualitative and quantitative behaviour of the quantile and the density quantile function." Some results on estimation of density quantile function are given in the next section. In the last section, almost sure bounds for density quantile estimators are obtained for the dependent case.

The following preliminary results are needed to state the main results.

Lemma 1 : *For any integer $k \geqslant 2$, there exists a polynomial h_k of degree $\leqslant k$ such that*

$$\int_0^\infty y^j h_k(y) e^{-y} dy = \begin{cases} 0 & \text{if } j = 2, \ldots, k \\ 1 & \text{if } j = 0, 1. \end{cases} \qquad \ldots (1)$$

Proof : Let E be the $(k+1)\times(k+1)$ matrix whose (i, j)-th element is $(i+j)!$, $i, j = 0, 1, \ldots, k$. For any real numbers $c_0, c_1, \ldots, c_k$, not all zero, we have

$$\sum_{i,j=0}^{k} (i+j)!\, c_i c_j = \sum_{i,j=0}^{k} c_i c_j \int_0^\infty y^{i+j} e^{-y}\, dy$$

$$= \int_0^\infty \Big(\sum_{i=0}^{k} y^i c_i \Big)^2 e^{-y}\, dy > 0.$$

So, E is a positive definite matrix. Let

$$(0_0, \ldots, 0_k) = (1, 1, 0, \ldots, 0) E^{-1}.$$

Clearly, (1) holds if we take $h_k(x) = \sum_{i=0}^{k} x^i 0_i$. This completes the proof.

AMS (1980) *subject classification* : 62G30, 62G05, 62E20.

Key words and phrases : Quantile, density quantile function, Bahadur-Kiefer representation of quantiles, ϕ -mixing, Borel–Cantelli lemma.

Lemma 2: *For any $k \geqslant 2$, let h_k be as in Lemma 1. Then*

$$\int_0^\infty \int_0^\infty h_k(x)h_k(y)\min\{x, y\}e^{-x-y}\,dx\,dy = \sigma_k^2 > 0, \qquad \ldots (2)$$

where $\sigma_k^2 = \int_0^\infty (g_k(x))^2 dx$ *and* $g_k(x) = \int_x^\infty h_k(y)e^{-y}\,dy$.

Proof: Let $I(B)$ denote the indicator of B. Let $h_k^*(x) = h_k(x)e^{-x}$. Since

$$\int_0^\infty \int_0^\infty x\,|h_k^*(x)h_k^*(y)|\,dx\,dy < \infty,$$

using the dominated convergence theorem and the Fubini's theorem we get that the l.h.s. of (2) equals

$$\int_0^\infty \int_0^\infty h_k^*(x)h_k^*(y)\,\Big(\int_0^\infty I(u \leqslant x)I(u \leqslant y)du\Big)dx\,dy$$

$$= \lim_{v\to\infty} \int_0^\infty \int_0^\infty h_k^*(x)h_k^*(y)\Big(\int_0^v I(u \leqslant x)I(u \leqslant y)du\Big)dx\,dy$$

$$= \lim_{v\to\infty} \int_0^v \Big(\int_u^\infty h_k^*(x)dx\Big)^2 du = \sigma_k^2.$$

Clearly $\sigma_k > 0$. Note that

$$|g_k(x)| \ll \int_x^\infty y^k e^{-y}dy \ll (1+x^k)e^{-x}.$$

So g_k^2 is integrable, as a result σ_k is finite. This completes the proof.

For $k = 2, 3$ we can take

$$h_2(x) = (4x-x^2)/2$$

and

$$h_3(x) = (-12+48x-21x^2+2x^3)/6.$$

For these choices $\sigma_2^2 = 13/16$ and $\sigma_3^2 = 63/32$.

2. Density quantile estimation

For any distribution function G, $0 < p < 1$, let $G^{-1}(p) = \inf\{x : G(x) \geqslant p\}$. Let $\{X_1, \ldots, X_n\}$ be a collection of random variables with common distribution function F. Let F_n denote the empirical distribution function of $X_1, \ldots, X_n$. For any $0 < p < 1$ and $k \geqslant 2$, let $\delta(k) = 1/(2k-1)$, $\beta(k) = (1-\delta(k))/2$,

$$T(n, k, p) = n^{2\delta(k)-1}\Sigma^*(F_n^{-1}(p+i/n) - F_n^{-1}(p))h_k(in^{\delta(k)-1})\exp(-in^{\delta(k)-1}) \qquad (3)$$

and $$A(n, k, p) = n^{\beta(k)}\sigma_k^{-1}(T(n, k, p)f(F^{-1}(p))-1), \quad \ldots (4)$$

where Σ^* denotes the sum over all positive integers $1 \leqslant i \leqslant 3n^{1-\delta(k)}\log n$ and h_k and σ_k are as in Lemmas 1 and 2. In practice sums are easier to compute than the integrals. This is the reason for defining $T(n, k, p)$ through a sum. Note that $A(n, k, p)$ is defined whenever F has a derivative f at $F^{-1}(p)$. The following definition is useful in stating the results.

Definition. Let $k \geqslant 1$ be an integer. The distribution function F is said to satisfy condition $A(k)$ at x, if F is k-times continuously differentiable at x and $f(x) > 0$, where f is the first derivative of F.

Theorem 1 : *Let $\{X_n\}$ be i.i.d. random variables with distribution function F. Let $0 < p_1 < \ldots < p_m < 1$, $k \geqslant 2$ and let F satisfy condition $A(k)$ in a neighborhood of $F^{-1}(p_j)$, $1 \leqslant j \leqslant m$. Then $\{A(n, k, p_1), \ldots, A(n, k, p_m)\}$ are asymptotically independent and each $A(n, k, p_j)$ converges weakly to the standard normal distribution.*

Proof : To simplify the notation, we drop the subscript k in the proof. For example, we write δ, β, h instead of $\delta(k)$, $\beta(k)$ and h_k. First note that for $1 \leqslant i \leqslant 3\ n^{1-\delta}\log n$,

$$\begin{aligned}
&\left| \int_{(i-1)/n}^{i/n} [e^{-in^{\delta-1}}h(in^{\delta-1}) - e^{-xn^{\delta}}h(xn^{\delta})]dx \right| \\
&= \left| \int_{(i-1)/n}^{i/n} n^{\delta} \left(\int_{x}^{i/n} e^{-yn^{\delta}}(h(yn^{\delta}) - h'(yn^{\delta}))dy \right) dx \right| \\
&\ll n^{\delta-1}(\log n)^k \int_{(i-1)/n}^{i/n} e^{-xn^{\delta}}dx \\
&\ll n^{-1}(\log n)^k \int_{(i-1)n^{\delta-1}}^{in^{\delta-1}} e^{-x}\,dx. \qquad \ldots (5)
\end{aligned}$$

Since h is a polynomial of degree k the inequality above follows as

$$|h(y)| + |h'(y)| \ll |y|^k.$$

Let $\{U_i\}$ be i.i.d. $U[0, 1]$ random variables. Let V_n denote the empirical distribution of $U_1, \ldots, U_n$. Without loss of generality we can take $X_i = F^{-1}(U_i)$. We require the following well known result;

$$\sup_{0<s<1} |V_n^{-1}(s) - s| \ll d_n \text{ a.s.}, \qquad \ldots (6)$$

where $d_n = ((\log\log n)/n)^{1/2}$. Let for $t \in \{p_1, \ldots, p_m\}$, $a_j = a_j(t)$ denote the j-th derivative of F^{-1} at $(V_n^{-1}(t))$.

To simplify the notation we put

$$S(n, t, v) = V_n^{-1}(t+v) - V_n^{-1}(t).$$

By (6) and by Taylor's expansion we have for small v, that a.e.

$$\begin{aligned} F_n^{-1}(t+v) - F_n^{-1}(t) &= \sum_{j=1}^{k} (S(n, t, v))^j a_j + o((S(n, t, v))^k) \\ &= \sum_{j=1}^{k} (S(n, t, v) - v + v)^j a_j + o((S(n, t, v) - v + v)^k) \\ &= \sum_{j=1}^{k} (V_n^{-1}(t+v) - V_n^{-1}(t) - v)^j a_j + \sum_{j=1}^{k} v^j a_j + O(v d_n) + O(d_n^2) + o(v^k) \\ &= \sum_{j=1}^{k} a_j v^j + (V_n^{-1}(t+v) - V_n^{-1}(t) - v) a_1 + O(d_n^2) + O(v d_n) + o(v^k). \end{aligned}$$

By Bahadur-Kiefer representation of quantiles (see Kiefer 1967), we have for small v a.e.

$$\begin{aligned} F_n^{-1}(t+v) - F_n^{-1}(t) = \sum_{j=1}^{k} a_j v^j - (V_n(t+v) - V_n(t) - v) a_1 \\ + O(n^{-3/4} \log n) + O(v d_n) + o(v^k). \end{aligned} \quad \dots (7)$$

Now by (5) we get a.e. that

$$\begin{aligned} T(n, k, t) &= n^{2\delta} \int_0^{3n^{-\delta} \log n} (F_n^{-1}(p+x) - F_n^{-1}(p)) h^*(xn^{\delta}) dx + O(n^{2\delta - 1} (\log n)^k) \\ &= n^{\delta} \int_0^{3 \log n} (F_n^{-1}(p + vn^{-\delta}) - F_n^{-1}(p)) h^*(v) dv + O(n^{2\delta - 1} (\log n)^k). \end{aligned}$$

As

$$\int_{3 \log n}^{\infty} y^j e^{-y} dy = O(n^{-3} (\log n)^j),$$

we have from (7) and Lemma 1, that a.e.,

$$T(n, k, t) = (1 + I(n, k, t)) a_1 + O(n^{\delta - 3/4} \log n) + O(d_n) + o(n^{-(k-1)\delta}),$$

where

$$I(n, k, t) = -n^{\delta} \int_0^{3 \log n} (V_n(t + vn^{-\delta}) - V_n(t) - vn^{-\delta}) h^*(v) dv.$$

Since $|f(F^{-1}(t)) a_1 - 1| \ll |V_n^{-1}(t) - (t)| \ll d_n$, we obtain a.e.

$$A(n, k, t)\sigma = n^{\delta} I(n, k, t) + o(1). \quad \dots (8)$$

By Theorem 4.4.1 of Csörgo and Révész (1981), there exists a Brownian motion W and a process Z_n such that

$$\{\sqrt{n}(V_n(t) - t) : 0 \leqslant t \leqslant 1\} \stackrel{D}{=} \{W(t) - tW(1) + Z_n(t) : 0 \leqslant t \leqslant 1\}$$

and

$$\sup_{0 \leqslant t \leqslant 1} |Z_n(t)| = O_p(n^{-1/2} \log n)$$

where $O_p(g_n)$ is a sequence of random variables H_n such that H_n/g_n is bounded in probability and $X \overset{D}{=} Y$ denotes that the random variables X and Y have the same distribution. As W and $-W$ have the same distribution we have

$$n^{\beta} I(n, k, t) \overset{D}{=} n^{\beta/2} \int_0^{3 \log n} (W(t+vn^{-\delta}) - W(t))h^*(v)dv$$

$$+O_p(\log n)^2 n^{-\beta}) + W(1)n^{\beta-(1/2)}. \qquad \ldots (9)$$

Now (8) and (9) give the asymptotic independence of $\{A(n, k, p_1), \ldots, A(n, k, p_m)\}$ and for each j,

$$A(n, k, p_j) \overset{D}{=} \sigma^{-1} \int_0^{\infty} W(v)h(v)e^{-v}dv + o_p(1),$$

where $o_p(1)$ denotes a sequence of random variables tending to zero in probability. But $\int_0^{\infty} W(v)h(v)e^{-v}\,dv$ is a normal variable with mean zero and variance σ. This completes the proof.

Remark 1 : Suppose for some $k \geqslant 2$, the distribution function F satisfies condition $A(k)$ at $F^{-1}(t)$, for $t \in [a, b]$, $0 < a < b < 1$. First note that $\inf\{f(F^{-1}(t)) : t \in [a, b]\} > 0$. From Theorem 1, it is clear that

$$A(n, k) = \sup\{|A(n, k, t)| : a < t < b\} \qquad \ldots (10)$$

tends to infinity in probability. The next theorem gives bounds for $A(n, k)$ in probability.

Theorem 2 : *Under the conditions of Remark* 1, *there exists a* $b(k) > 1$ *such that as* $n \to \infty$,

$$P((2\delta(k)\log n - 4 \log\log n)^{1/2} \leqslant A(n, k) \leqslant b(k)(\log n)^{1/2}) \to 1.$$

Proof : Let σ_k and $\delta(k)$ be as before. We shall now partition the interval $[a, b]$ into smaller intervals of length $6(\log n)n^{-\delta(k)}$ and leave a gap at each of the end points a and b. To do this let $s = s_n = [(b-a)n^{\delta(k)}/6 \log n] - 5$ and let $a < p_0 < p_1 < \ldots < p_{s+1} < b$ be such that $p_{i+1} - p_i = 6(\log n)n^{-\delta(k)}$ for $1 \leqslant i \leqslant s$. Put

$$K(n, k, t) = n^{\delta(k)/2}\sigma_k^{-1} \int_0^{3 \log n} (W(t+vn^{-\delta(k)}) - W(t))h_k(v)e^{-v}dv$$

and $r_n = (2\delta(k) \log n - (7/2)\log\log n)^{1/2}$. Clearly, $\{K(n, k, p_i) : 1 < i < s\}$ are independent. Since for any $x > 0$, the standard normal distribution function Φ satisfies

$$(x^{-1} - x^{-3})\exp\left(-\frac{1}{2}x^2\right) \leqslant \sqrt{2\pi}(1 - \Phi(x)),$$

we obtain

$$P\Big(\sup_{a<t<b}|K(n,k,t)|<r_n\Big)\leqslant P\Big(\sup_{1\leqslant i\leqslant s}|K(n,k,p_i)|<r_n\Big)$$

$$\leqslant\Big(P\Big(\Big|\int_0^{3\log n}W(v)h_k(v)e^{-v}dv\Big|\leqslant\sigma_k r_n\Big)\Big)^s$$

$$\leqslant(1-2(1-\Phi(r_n+O(n^{-2}))))^s. \qquad \ldots (11)$$

Note that

$$1-\Phi(r_n+O(n^{-2}))\geqslant\frac{1}{\sqrt{4\pi\delta(k)}}(\log n)^{-1/2}(1+O(\log n^{-1/2}))n^{-\delta(k)}(\log n)^{7/4}.$$

Since $1-x\leqslant e^{-x}$ for $x>0$, the left side of (11) is not more than

$$(1-(1/\sqrt{4\pi\delta(k)}(1+O((\log n)^{-1/2}))n^{-\delta(k)}(\log n)^{5/4})^s.$$

$$\ll\exp(-((b-a)/\sqrt{4\pi\delta(k)})(\log n)^{1/4})\to 0,$$

as $n\to\infty$. Since $o(1)$ estimate in (8) holds uniformly in any closed interval J as long as $f\circ F^{-1}$ is bounded away from zero and $F'(k-1)\circ F^{-1}$ is continuous in J, we get from (9) that

$$A(n,k)\overset{\mathrm{D}}{=}\sup\{|K(n,k,t)|:a<t<b\}+o_p(1). \qquad \ldots (12)$$

So from (12) we obtain that $P(A(n,k)\geqslant(\log n-4\log\log n)^{1/2})\to 1$.

To complete the proof we use (see (1.1.12) of Csörgö and Révész (1981)

$$\limsup_{s\to 0}\sup_{0\leqslant t\leqslant 1-s}\sup_{0<v<s}\frac{|W(t+v)-W(t)|}{\sqrt{2v\log(1/v)}}\leqslant 1 \text{ a.e.}$$

and the integrability of $\sqrt{v\log(1/v)}h_k(v)e^{-v}$, to get

$$\limsup_{n\to\infty}(\log n)^{-1/2}\sup_{a<t<b}|K(n,k,t)|$$

$$\leqslant\limsup_{n\to\infty}\int_0^{\infty}(2\delta(k)v+2v(\log(1/v))/(\log n))^{1/2}|h_k(v)|e^{-v}dv$$

$$\leqslant\int_0^{\infty}(2\delta(k)v)^{1/2}|h_k(v)|e^{-v}dv=b(k)<\infty \text{ a.e.}$$

The theorem now follows from (12).

3. Density quantile estimation in the dependent case

A sequence $\{X_n\}$ is called ϕ-mixing if there exists a monotone sequence $\{\phi(n)\}$ such that $\phi(n)\to 0$ as $n\to\infty$ and

$$\sup\{|P(A\mid B)-P(A)|:A\in M_{m+n},B\in M_1^m,P(B)>0,m\geqslant 1\}\leqslant\phi(n),$$

where M_m^r denotes the σ-field generated by $\{X_i:m\leqslant i\leqslant r\}$.

The following theorem gives a.e. bounds in the dependent case.

Theorem 3 : *Let $\{X_n\}$ be a strictly stationary ϕ-mixing sequence with $\Sigma\phi^{1/2}(j) < \infty$ and X_1 having a continuous distribution function F. Let for some $k \geqslant 2$ and $0 < a < b < 1$, F satisfy condition $A(k)$ at $F^{-1}(t)$ for $t \in [a, b]$. Let $A(n, k)$ be as in (10). Then $A(n, k) \ll \log n$ a.e.*

Let $U_t = F(X_t)$. Clearly $\{U_t\}$ is a stationary ϕ-mixing sequence of $U[0, 1]$ random variables with $\Sigma\phi^{1/2}(j) < \infty$. Let V_n denote the empirical distribution of $U_1, \ldots, U_n$. We recall some results from Babu and Singh (1978).

Lemma 3 : *For $0 \leqslant u \leqslant v \leqslant 1$, let*

$$x_t(u, v) = I(u \leqslant U_t \leqslant v) - (v - u),$$

where I denotes the indicator function. Then there exists $q, q_1, q_2 > 0$ such that, whenever $0 \leqslant u < 1$, $0 < c < 1-u$, $|v-u| \leqslant c$ and $0 < d < cn^{19/24}$, we have

$$P\left(\left|\sum_{i=1}^{n} x_t(u, v)\right| > 2dq\right) \leqslant q_1 n^{-4} + q_2 \exp(-8d^2 n^{-1} c^{-1}).$$

For a proof see Lemma 2.1 of Babu and Singh (1978).

Lemma 4 : *We have a.e.,*

$$\sup\{|V_n^{-1}(t) - t| : 0 < t < 1\} \ll d_n$$

and

$$\sup_{0 < t < 1} |V_n^{-1}(t) - V_n(t) - 2t| \ll b^{-3/4} \log n.$$

For a proof see Theorem 1 and Lemma 2.3 of Babu and Singh (1978).

Lemma 5 : *For any $0 < \delta \leqslant 1/2$,*

$$D_n = \sup\{|V_n(t+v) - V_n(t-v) - 2v| : 0 \leqslant v \leqslant \lambda_n, 0 < t < 1\} \ll c_n \text{ a.e.,}$$

where $c_n = n^{-(\delta+1)/2} \log n$ and $\lambda_n = 3n^{-\delta} \log n$.

Proof : Let $v_n = [1/c_n] + 1$ and $w_n = [\lambda_n/c_n] + 2$. Note that

$$\begin{aligned} D_n &\leqslant \sup\{|V_n(t) - V_n(u) - t + u| : 0 < t < 1,\ |x-t| \leqslant \lambda_n\} \\ &\leqslant \sup\{|V_n(t + jc_n) - V_n(t) - jc_n| + c_n : 0 < t < 1,\ |j| \leqslant w_n\} \\ &\leqslant \max\{|V_n((i+j)c_n) - V_n(ic_n) - jc_n| : 1 \leqslant i \leqslant v_n,\ |j| \leqslant w_n\} + 2c_n \\ &= R_n + 2c_n \text{ (say).} \end{aligned}$$

It is enough to show that a.e., $R_n \ll c_n$. We apply Lemma 3 with $c = \lambda_n$ and $d = 3nc_n$ to get

$$\begin{aligned} P(nR_n > 2qd) &\leqslant 2n \sup\{P(n|V_n(t+s) - V_n(s) - t| > 2qd) : 0 \leqslant s \leqslant 1,\ |t| \leqslant \lambda_n\} \\ &\ll n^{-2}. \end{aligned}$$

Now the result follows from Borel-Cantelli Lemma.

Proof of Theorem 3 : From Lemma 4 and the proof of Theorem 1, we have (8) uniformly for $t \in (a, b)$. So by lemma 5 we get a.e.

$$A(n,k) \ll n^{\beta(k)}\sup\{|J(n,k,t)| : a < t < b\}+o(1)$$

$$\ll n^{\beta(k)+\delta(k)}\sup\{|V_n(t+v)-V_n(t)-v| : a < t < b,$$

$$0 < vn^{\delta(k)} < \log n\}+o(1)$$

$$\ll n^{\beta(k)+\delta(k)-(\delta(k)+1)/2}\log n \ll \log n.$$

This completes the proof.

Remark 2. A similar result holds for strong-mixing random variables also. In this case one needs to use Theorem 3, Lemmas 3.3 and 3.4 of Babu and Singh (1978).

Acknowledgements. The author would like to thank S. C. Bagchi, J. C. Gupta and B. V. Rao of Stat-Math. Division, for discussions leading to the proofs of Lemmas 1 and 2.

References

Babu, G. J. and Singh, K. (1978) : On deviations between empirical and quantile processes for mixing random variables. *Jour. Multivariate Analysis*, **8**, 532-549.

Csorgo, M. and Revesz, P. (1981) : *Strong Approximations in Probability and Statistics*. Academic Press, New York.

Kiefer, J. (1967) : On Bahadur's representation of sample quantiles. *Annals of Math. Statist.*, **38**, 1323-1342.

Parzen, E. (1979) : Density quantile estimation approach to statistical data modelling. *Smoothing techniques for curve estimation*. (Th. Gasser and M. Rosenblatt, Ed.) Springer Lecture notes in Mathematics No. 757, 155-180.

Paper received : November, 1983.

Revised : March, 1985.