Bahadur representation of sample quantiles for functional of Gaussian dependent sequences under a minimal assumption

Jean-François Coeurjolly

SAGAG, Department of Statistics, Grenoble, 1251 Av. Centrale BP 47, 38040 GRENOBLE Cedex 09, France

A R T I C L E   I N F O

Article history:
Received 9 March 2007
Accepted 14 February 2008
Available online 20 March 2008

A B S T R A C T

We obtain a Bahadur representation for sample quantiles of a nonlinear functional of Gaussian sequences with correlation function decreasing as $k^{-\alpha}$ for some $\alpha > 0$. This representation is derived under a minimal assumption.

© 2008 Elsevier B.V. All rights reserved.

1. Introduction

We consider the problem of obtaining a Bahadur representation of sample quantiles in a certain dependence context. Before stating in what a Bahadur representation consists, let us specify some general notation. Given some random variable $Y$, $F(\cdot) = F_Y(\cdot)$ is referred to as the cumulative distribution function of $Y$, $\xi(p) = \xi_Y(p)$ for some $0 < p < 1$ as the quantile of order $p$. If $F(\cdot)$ is absolutely continuous with respect to Lebesgue measure, the probability density function is denoted by $f(\cdot) = f_Y(\cdot)$. On the basis of the observation of a vector $Y = (Y(1), \ldots, Y(n))$ of $n$ random variables distributed as $Y$, the sample cumulative distribution function and the sample quantile of order $p$ are respectively denoted by $\hat{F}_Y(\cdot; Y)$ and $\hat{\xi}_Y(p; Y)$ or simply by $\hat{F}(\cdot)$ and $\hat{\xi}(p)$.

Let $Y = (Y(1), \ldots, Y(n))$ be a vector of $n$ i.i.d. random variables such that $F''(\xi(p))$ exists and is bounded in a neighborhood of $\xi(p)$ and such that $F'(\xi(p)) > 0$. Bahadur (1966) proved that as $n \to +\infty$,

$$\hat{\xi}(p) - \xi(p) = \frac{p - \hat{F}(p)}{f(\xi(p))} + r_n,$$

with $r_n = O_{a.s.}(n^{-3/4} \log(n)^{3/4})$ where a sequence of random variables $U_n$ is said to be $O_{a.s.}(v_n)$ if $U_n/v_n$ is almost surely bounded. Kiefer (1967) obtained the exact rate $n^{-3/4} \log \log(n)^{3/4}$. Under an assumption on $F(\cdot)$ which is quite similar to the one made by Bahadur, extensions of above results to dependent random variables have been pursued in Sen (1972) for $\phi$-mixing variables, in Yoshihara (1995) for strongly mixing variables, and recently in Wu (2005) for short-range and long-range dependent linear processes, following works of Hesse (1990) and Ho and Hsing (1996). Finally, such a representation has been obtained by Coeurjolly (2008) for a nonlinear functional of Gaussian sequences with correlation function decreasing as $k^{-\alpha}$ for some $\alpha > 0$.

Ghosh (1971) proposed in the i.i.d. case a much simpler proof of Bahadur’s result which suffices for many statistical applications. He established under a weaker assumption on $F(\cdot)$ ($F'(\cdot)$ exists and is bounded in a neighborhood of $\xi(p)$ and $f(\xi(p)) > 0$) that the remainder term satisfies $r_n = o_{a.s.}(n^{-1/2})$, which means that $n^{1/2} r_n$ tends to 0 in probability. This result is sufficient for example to establish a central limit theorem for the sample quantile. Our goal is to extend Ghosh’s result to a nonlinear functional of Gaussian sequences with correlation function decreasing as $k^{-\alpha}$. The Bahadur representation is presented in Section 2 and is applied to a central limit theorem for the sample quantile. Proofs are deferred to Section 3.

E-mail address: Jean-François.Coeurjolly@upmf-grenoble.fr.

0167-7152/$-$ see front matter © 2008 Elsevier B.V. All rights reserved.

doi:10.1016/j.spl.2008.02.037
2. Main result

Let \( \{Y(i)\}_{i=1}^{+\infty} \) be a stationary (centered) Gaussian process with variance 1, and correlation function \( \rho(\cdot) \) such that, as \( i \to +\infty \)

\[
|\rho(i)| \sim i^{-\alpha}
\]

for some \( \alpha > 0 \).

Let us recall some background on Hermite polynomials: the Hermite polynomials form an orthogonal system for the Gaussian measure and are in particular such that \( \mathbb{E}(H_i(Y)H_k(Y)) = \delta_{i,k} \), where \( Y \) is referred to a standard Gaussian variable. For some measurable function \( g(\cdot) \) defined on \( \mathbb{R} \) such that \( \mathbb{E}(g(Y)^2) < +\infty \), the following expansion holds:

\[
g(t) = \sum_{j \geq \tau} c_j H_j(t) \quad \text{with} \quad c_j = \mathbb{E}(g(Y)H_j(Y)),
\]

where the integer \( \tau \) defined by \( \tau = \inf \{ j \geq 0, c_j \neq 0 \} \), is called the Hermite rank of the function \( g \). Note that this integer plays an important role. For example, it is related to the correlation of \( g(Y) \) where the integer \( g \) is the minimal Hermite rank of \( A \) satisfying Assumption \( \tau \).

We denote by \( \hat{\rho} \) the restriction of \( \rho(\cdot) \) on \( U \) and where \( \hat{\rho}(\cdot) \) is referred to the probability density function of a standard Gaussian variable.

Now, define, for some real \( u \), the function \( h_u(\cdot) \) by

\[
h_u(t) = \mathbf{1}_{g(U) \subseteq (t)}(t) - f_{g(Y)}(u).
\]

We denote by \( \tau(u) \) the Hermite rank of \( h_u(\cdot) \). For the sake of simplicity, we set \( \tau_p = \tau(\xi_{g(Y)}(p)) \) for some function \( g(\cdot) \) satisfying Assumption \( A(\xi(\cdot)) \), we define

\[
\tau_p = \inf_{y \in U, l \neq l} \tau(Y),
\]

that is the minimal Hermite rank of \( h_u(\cdot) \) for \( u \) in a neighborhood of \( \xi_{g(Y)}(p) \). Denote also by \( c_j(u) \) the \( j \)-th Hermite coefficient of the function \( h_u(\cdot) \).

**Theorem 1.** Under Assumption \( A(\xi(\cdot)) \), the following result holds as \( n \to +\infty \):

\[
\tilde{\xi}(p; g(Y)) - \xi_{g(Y)}(p) = \frac{p - \hat{\xi}(\xi_{g(Y)}(p); g(Y))}{f_{g(Y)}(\xi_{g(Y)}(p))} + o_p\left( r_n(\alpha, \tau_p) \right).
\]

where \( g(Y) = (g(Y(1)), \ldots, g(Y(n))) \), for \( i = 1, \ldots, n \), and where the sequence \( (r_n(\alpha, \tau_p))_{n \geq 1} \) is defined by

\[
r_n(\alpha, \tau_p) = \begin{cases} 
    n^{-1/2} & \text{if } \alpha \tau_p > 1, \\
    n^{-1/2} \log(n)^{1/2} & \text{if } \alpha \tau_p = 1, \\
    n^{-\alpha\tau_p/2} & \text{if } \alpha \tau_p < 1.
\end{cases}
\]

**Remark 1.** The sequence \( r_n(\alpha, \tau_p) \) is related to the short-range or long-range dependent behaviour of the sequence \( h_u(Y(1)), \ldots, h_u(Y(n)) \) for \( u \) in a neighborhood of \( \xi(p) \). More precisely, it corresponds to the asymptotic behaviour of the sequence

\[
\left( \frac{1}{n} \sum_{|i| < n} \rho(i) Y_i \right)^{1/2}.
\]

**Corollary 2.** Under Assumption \( A(\xi(p)) \), then the following convergences in distribution hold as \( n \to +\infty \):

(i) if \( \alpha \tau_p > 1 \)

\[
\sqrt{n} \left( \tilde{\xi}(p; g(Y)) - \xi_{g(Y)}(p) \right) \overset{d}{\rightarrow} \mathcal{N}(0, \sigma_p^2),
\]

(ii) if \( \alpha \tau_p = 1 \)

\[
\sqrt{n} \left( \tilde{\xi}(p; g(Y)) - \xi_{g(Y)}(p) \right) \overset{d}{\rightarrow} \mathcal{N}(0, \sigma_p^2),
\]

(iii) if \( \alpha \tau_p < 1 \)

\[
\sqrt{n} \left( \tilde{\xi}(p; g(Y)) - \xi_{g(Y)}(p) \right) \overset{d}{\rightarrow} \mathcal{N}(0, \sigma_p^2).
where
\[
\sigma_p^2 = \frac{1}{f(p)^2} \sum_{i=1}^{\kappa} \sum_{j \geq \kappa_p} \frac{c_j(p)^2}{j!} \rho(i) \quad \text{with } f(p) = f_{g(Y)}(\xi_{g(Y)}(p)) \text{ and } c_j(p) = c_j(\xi_{g(Y)}(p)).
\]

\[\text{(ii) if } \alpha \tau_p < 1 \]
\[
n^{\tau_p/2} \left( \tilde{\xi}(p; g(Y)) - \xi_{g(Y)}(p) \right) \xrightarrow{d} \frac{c_{\tau_p}(p)}{\tau_p f(p)} Z_{\tau_p},
\]
where
\[
Z_{\tau_p} = K(\tau_p, \alpha) \int_{\mathcal{L}^n} \exp((\lambda_1 + \cdots + \lambda_{\tau_p})) - 1 \prod_{j=1}^{\tau_p} |\lambda_j|^{(\alpha-1)/2} \bar{B}(d\lambda_j)
\]
and
\[
K(\tau_p, \alpha) = \left( \frac{(1 - \alpha \tau_p/2)(1 - \alpha \tau_p)}{\tau_p (2 \Gamma(\alpha) \sin(\pi(1 - \alpha)/2))^{\tau_p}} \right)^{1/2}.
\]

The measure \( \bar{B} \) is a Gaussian complex measure and the symbol \( f' \) means that the domain of integration excludes the hyperdiagonals \( \{ \lambda_i = \pm \lambda_j, i \neq j \} \).

The proof of this result is omitted since it is a direct application of Theorem 1 and general limit theorems adapted to nonlinear functionals of Gaussian sequences; e.g. Breuer and Major (1983) and Dehling and Taqqu (1989).

3. Proofs

3.1. Auxiliary lemma

Lemma 3. For every \( j \geq 1 \) and for all positive sequence \( (u_n)_{n \geq 1} \) such that \( u_n \to 0 \), as \( n \to +\infty \), we have, under Assumption A(\( \xi(p) \)),
\[
I = \int_{U} H_{j}(t) \phi(t) 1_{[\xi(t)-\xi_{g(Y)}(p) \in [0, u_n)]} dt \sim u_n \kappa_j,
\]
where \( \kappa_j \) is defined, for every \( j \geq 1 \), by
\[
\kappa_j = \begin{cases} 
-2 \sum_{i=1}^{j-1} \frac{\phi'(g_i^{-1}(\xi(p)))}{g'(g_i^{-1}(\xi(p)))} & \text{if } j = 1, \\
2(-1)^{j-1} \sum_{i=1}^{j-1} \frac{\phi'(g_i^{-1}(\xi(p)))}{g'(g_i^{-1}(\xi(p)))} & \text{if } j > 1.
\end{cases}
\]

Proof. Under Assumption A(\( \xi(p) \)), there exists \( n_0 \in \mathbb{N} \) such that for all \( n \geq n_0 \),
\[
I = \sum_{i=1}^{r} I_i \quad \text{with } I_i = \int_{U_i} H_{j}(t) \phi(t) 1_{[\xi(t)-u_n \leq \xi(p) \leq \xi(t) + u_n]} dt.
\]
Assume without loss of generality that the restriction of \( g(\cdot) \) on \( U_i \) (denoted by \( g_i(\cdot) \)) is an increasing function; we have
\[
I_i = \int_{U_i} H_{j}(t) \phi(t) 1_{[\xi(t)-u_n \leq \xi(p) \leq \xi(t) + u_n]} dt
\]
\[
= \int_{\xi(p)-u_n}^{\xi(p)+u_n} H_{j}(t) \phi(t) dt
\]
\[
= \left\{ \begin{array}{ll}
\phi(m_{i,n}) - \phi(M_{i,n}) = (m_{i,n} - M_{i,n}) & \text{if } j = 1 \\
(-1)^{j} \left( \phi^{(j-1)}(M_{i,n}) - \phi^{(j-1)}(m_{i,n}) \right) & \text{if } j > 1,
\end{array} \right.
\]
where \( M_{i,n} = g_i^{-1}(\xi(p) + u_n) \) and \( m_{i,n} = g_i^{-1}(\xi(p) - u_n) \). Then, there exists \( \omega_{n,i,j} \in [m_{i,n}, M_{i,n}] \) for every \( j \geq 1 \) such that
\[
I_i = \left\{ \begin{array}{ll}
(m_{i,n} - M_{i,n}) \phi^{(j)}(\omega_{n,i,j}) & \text{if } j = 1 \\
(-1)^{j} (M_{i,n} - m_{i,n}) \phi^{(j-1)}(\omega_{n,i,j}) & \text{if } j > 1.
\end{array} \right.
\]
Under Assumption $A(\xi(p))$, we have, as $n \to +\infty$,
\[
\alpha_{n,i,j} \sim g_i^{-1}(\xi(p)) \quad \text{and} \quad M_{n,i} - m_{i,n} \sim 2u_i \frac{1}{g'(g_i^{-1}(\xi(p)))},
\]
which ends the proof.  

3.2. Proof of Theorem 1

For the sake of simplicity, we set $\hat{\xi} (p) = \hat{\theta} (p; g(Y))$, $\xi(p) = \xi_{gY}(p)$, $\tilde{\theta} (\cdot) = \tilde{\theta} (\cdot; g(Y))$, $F(\cdot) = f_{gY}(\cdot)$, $f(\cdot) = f_gY(\cdot)$ and $r_n = r_n(\alpha, \tau_p)$. Define
\[
V_n = r_n^{-1} \left( \hat{\xi} (p) - \xi(p) \right) \quad \text{and} \quad W_n = r_n^{-1} \left( \frac{p - \tilde{\theta}(p)}{f(p)} \right).
\]
The result is established if $V_n - W_n \to 0$ as $n \to +\infty$. It suffices to prove that $V_n$ and $W_n$ satisfy the conditions of Lemma 1 of Ghosh (1971):

- **Condition (a):** for all $\delta > 0$, there exists $\epsilon = \epsilon(\delta)$ such that $P(|W_n| > \epsilon) < \delta$.
- **Condition (b):** for all $y \in \mathbb{R}$ and for all $\epsilon > 0$.
\[
\lim_{n \to +\infty} P(V_n \leq y, W_n \geq k + \epsilon) \quad \text{and} \quad \lim_{n \to +\infty} P(V_n \geq y + \epsilon, W_n \geq k).
\]

**Condition (a):** From the Bienaymé–Tchebyshev inequality it is sufficient to prove that $E W_n^2 = O(1)$. Rewrite $W_n = n^{-1} \sum_{i=1}^{n} h_{\xi(p)} (Y(i))$. Let $c_j$ (for some $j \geq 0$) denote the $j$-th Hermite coefficient of $h_{\xi(p)} (\cdot)$. Since $h_{\xi(p)} (\cdot)$ has at least Hermite rank $\tau_p$, then
\[
E W_n^2 = \frac{r_n^{-2}}{n^2} \sum_{i_1,i_2=1}^{n} E(h_{\xi(p)} (Y(i_1)) h_{\xi(p)} (Y(i_2)))
= \frac{r_n^{-2}}{n^2} \sum_{i_1,i_2=1}^{n} \sum_{j_1,j_2 \geq \tau_p} c_{j_1} c_{j_2} E(H_{j_1} (Y(i_1)) H_{j_2} (Y(i_2)))
= \frac{r_n^{-2}}{n^2} \sum_{i_1,i_2=1}^{n} \sum_{j \geq \tau_p} (\xi)^2 \rho(i_2 - i_1) j!
= \sigma \left( \frac{r_n^{-2}}{n} \sum_{i \geq -n} \rho(i) \tau_p \right) = O(1),
\]
from Remark 1.

**Condition (b):** Let $y \in \mathbb{R}$; we have
\[
\{ V_n \leq y \} = \{ \hat{\theta} (p) \leq y \times r_n + \xi(p) \}
= \{ p \leq \tilde{\theta} (y \times r_n + \xi(p)) \} = \{ Z_n \leq y_n \},
\]
with
\[
Z_n = \frac{r_n^{-1}}{f(\xi(p))} \left( F(y \times r_n + \xi(p)) - \tilde{\theta} \left( \frac{y}{\sqrt{r_n}} + \xi(p) \right) \right)
\]
and
\[
y_n = \frac{r_n^{-1}}{f(\xi(p))} \left( F(y \times r_n + \xi(p)) - p \right).
\]

Under Assumption $A(\xi(p))$, we have $y_n \to y$, as $n \to +\infty$. Now, prove that $Z_n - W_n \xrightarrow{p} 0$. Without loss of generality, assume $y > 0$. Then, we have
\[
W_n - Z_n = \frac{r_n^{-1}}{f(\xi(p))} \left( \tilde{\theta} (y \times r_n + \xi(p)) - F(y \times r_n + \xi(p)) - \tilde{\theta} (\xi(p)) + F(\xi(p)) \right)
= \frac{r_n^{-1}}{f(\xi(p))} \frac{1}{n} \sum_{i=1}^{n} h_{\xi(p,n)} (Y(i))
\]
where $h_{\xi(p,n)} (\cdot)$ is the function defined for $t \in \mathbb{R}$ by
\[
h_{\xi(p,n)} (t) = 1_{\xi(p) \leq t < \xi(p) + y \times r_n} (t) - P(\xi(p) \leq g(Y) \leq \xi(p) + y \times r_n).
\]
For $n$ sufficiently large, the function $h_{\xi(p),n}(\cdot)$ has Hermite rank $\tau_p$. Denote by $c_{j,n}$ the $j$-th Hermite coefficient of $h_{\xi(p),n}(\cdot)$. From Lemma 3, there exists a sequence $(\kappa_j)_{j \geq \tau_p}$ such that, as $n \to +\infty$,

$$c_{j,n} \sim \kappa_j \times r_n.$$ 

Since, for all $n \geq 1$, $E(h_n(Y)^2) = \sum_{j=\tau_p}^{\infty} (c_{j,n})^2/j! < +\infty$, it is clear that the sequence $(\kappa_j)_{j \geq \tau_p}$ is such that $\sum_{j=\tau_p}^{\infty} (\kappa_j)^2/j! < +\infty$. By denoting as $\lambda$ a positive constant, we get, as $n \to +\infty$,

$$E(W_n - Z_n)^2 = \frac{r_n^2}{n^2} \frac{1}{f(\xi(p))^2} \sum_{i_1,i_2=1}^{n} E(h_{\xi(p),n}(Y(i_1)) h_{\xi(p),n}(Y(i_2)))
= \frac{r_n^2}{n^2} \frac{1}{f(\xi(p))^2} \sum_{i_1,i_2=1}^{n} \sum_{j_1,j_2=\tau_p}^{\infty} c_{j_1,n} c_{j_2,n} E(H_{j_1}(Y(i_1)) H_{j_2}(Y(i_2)))
= \frac{r_n^2}{n^2} \frac{1}{f(\xi(p))^2} \sum_{i_1,i_2=1}^{n} \sum_{j=\tau_p}^{\infty} \frac{\kappa_j^2}{j!} \rho(j_2 - i_1)^j
\leq \lambda \frac{r_n^2}{n} \sum_{j=\tau_p}^{\infty} \frac{(\kappa_j)^2}{j!} \sum_{i=|i|<n} \rho(i)^j = o\left(\frac{1}{n} \sum_{i=|i|<n} \rho(i)^n\right) = o\left(r_n^2\right),$$

from Remark 1. Therefore, $W_n - Z_n$ converges to 0 in probability, as $n \to +\infty$. Thus, for all $\epsilon > 0$, we have, as $n \to +\infty$,

$$P(V_n \leq y, W_n \geq y + \epsilon) = P(Z_n \leq y, W_n \geq y + \epsilon) \to 0.$$ 

Following the sketch of this proof, we also have $P(V_n \geq y + \epsilon, W_n \leq y) \to 0$, ensuring condition (b). Therefore, $W_n - Z_n$ converges to 0 in probability, as $n \to +\infty$. Thus, for all $\epsilon > 0$, we have, as $n \to +\infty$,

$$P(V_n \leq y, W_n \geq y + \epsilon) = P(Z_n \leq y, W_n \geq y + \epsilon) \to 0.$$ 

Following the sketch of this proof, we also have $P(V_n \geq y + \epsilon, W_n \leq y) \to 0$, ensuring condition (b).

**Acknowledgement**

The author would like to thank the referee for comments improving the statement of Corollary 2.

**References**


