Smooth nonparametric quantiles

Dana Draghicescu and Sucharita Ghosh

Abstract. For short memory processes nonparametric estimation of quantile functions was considered in Abberger (1996, 1997). Ghosh et. al (1997) consider the kernel estimaton of these functions for suitable transformations of an underlying Gaussian process with long memory. In this paper we consider similar processes, however the underlying Gaussian process is assumed to have serial correlations which decay fast so that the sum of all correlations is finite. We prove the consistency of the kernel estimates of distribution functions which are inverted to obtain consistent estimates of the quantiles. We also discuss a smoother estimate of the quantiles and illustrate the procedure by an application to a precipitation time series from Switzerland.

Mathematics Subject Classification 2000: 62M10, 62G05, 62G20.

Key words: Kernel smoothing, quantile estimation, short memory, time series.

1 Introduction

Consider a stochastic process $W_i,\ i = 1, 2, ...$ that is generated by a time-dependent transformation (see (1) below) of a zero mean stationary Gaussian process Z_i. Processes of this type were considered in Ghosh et al. (1997). Note that the probability distribution function of W_i may change with time so that in this sense, the process is not stationary. Ghosh et al. (1997) consider estimation of the conditional quantiles of such processes by kernel smoothing and derive asymptotic properties of the estimator in the presence of long-memory serial correlations in W_i (for detailed information on long-memory processes, see Beran (1994) and Cox (1984)). Specifically,

$$W_i \,=\, G(Z_i, t_i), \quad i = 1, 2, \ldots, \tag{1.1}$$

where $t_i = \frac{i}{n},\ i = 1, 2, \ldots$ are rescaled time points, $\{Z_i\}$ is a stationary Gaussian process with $E(Z_i) = 0,\ Var(Z_i) = 1$ and covariances $\gamma_Z(l) = cov(Z_i, Z_{i+l}),\ l = 0, \pm1, \pm2, \ldots$. Let the conditional probability distribution function of the process W at the rescaled time t be denoted by

$$F_t(w) \,=\, P(W(t) \le w|t). \tag{1.2}$$

The unknown function $G : \mathbb{R} \times [0, 1] \to \mathbb{R}$ is Lebesgue measurable with

$$1_{\{G(Z_i,t_i)\le w\}} - F_{t_i}(w) \,=\, \sum_{k=m}^{\infty} c_k(t_i, w)H_k(Z_i). \tag{1.3}$$

Proceedings of The 2-nd International Colloquium of Mathematics in Engineering and Numerical Physics (MENP-2), April 22-27, 2002, University Politehnica of Bucharest, Romania.
BSG Proceedings 8, pp. 45-52, Geometry Balkan Press, 2003.

Here $H_k(\cdot)$ is the Hermite polynomial of degree k and the coefficients $c_k(t,w)$ are twice continuously differentiable with respect to t and continuous with respect to w. In this article we assume that the process $\{Z_i\}$ has short memory, so that the infinite sum of the covariances $\gamma_Z(l),\ l = 0, \pm 1, \pm 2, \ldots$ is finite. By inverting (3), one obtains the quantile function

$$\theta_\alpha(t) \,=\, inf\{w \in \mathbb{R} | F_t(w) \geq \alpha\},\ 0 < \alpha < 1. \tag{1.4}$$

2 Smooth quantiles

Consider the kernel estimator of $F_t(w)$

$$\hat{F}_t(w) \,=\, \frac{1}{nb_1} \sum_{i=1}^{n} K_1\Big(\frac{t_i - t}{b_1}\Big) \cdot 1_{\{W_i \leq w\}},\quad t \in [0,1],\ w \in \mathbb{R}. \tag{2.5}$$

The corresponding estimate of the α - quantile may be obtained as

$$\hat{\theta}_\alpha(t) = inf\{w \in \mathbb{R} \big| \hat{F}_t(w) \geq \alpha, \}\ t \in [0,1],\ 0 < \alpha < 1. \tag{2.6}$$

In the above definitions, $b_1 = b_1(n)$ is a sequence of bandwidths such that, as $n \to \infty$, $b_1 \to 0$ and $nb_1^3 \to \infty$. Also K_1 in (2.5) is a kernel that is assumed to be a symmetric, Lipschitz-continuous and twice continuously differentiable density function with support $[-1,1]$ that integrates to 1. Ghosh et. al (1997) prove the consistency of the estimators (2.5) and (2.6) when the underlying Gaussian process $\{Z_i\}$ in (1.1) has long memory. The procedure was also used for nonparametric prediction of distribution functions and quantiles under long-memory in Ghosh and Draghicescu (2002). Applications of this estimator to the precipitation time series of Switzerland can be found in Draghicescu and Ghosh (2000, 2001) and Ghosh and Draghicescu (2001). For short-memory processes, kernel estimation of the quantile functions was considered in Abberger (1996, 1997). For further references to kernel estimation in general, see Wand and Jones (1995). In this article we prove the consistency of the same estimators for processes of the type $\{W_i\}$ when the underlying Gaussian process $\{Z_i\}$ has short memory (infinite sum of auto-covariances is finite). We also discuss a smoother estimating procedure for the quantiles and illustrate the method by an application to a Swiss precipitation series. In what follows, we assume that $\frac{\partial}{\partial w}\Big[F_t(w)\Big] = f_t(w)$ and $\frac{\partial^2}{\partial t^2}\Big[F_t(w)\Big]$ exist. Moreover, $K_1(0) \geq K_1(u)$, for all $u \in [-1,1]$. The proof of the following result is given in the appendix.

Theorem 1 *If* $\sum_{u=-\infty}^{\infty} |\gamma_Z^m(u)| < \infty$ *then, as* $n \to \infty$,

(a) Bias of $\hat{F}_t(w)$:

$$E\Big(\hat{F}_t(w)|t\Big) - F_t(w) = \frac{b_1^2}{2} \int_{-1}^{1} u^2 K_1(u) du \frac{\partial^2}{\partial t^2}\Big[F_t(w)\Big] + o\big(b_1^2\big). \tag{2.7}$$

(b) Variance: An upper bound of the variance of $\hat{F}_t(w)$ is given by

$$V_n(t,w) = \frac{K_1^2(0)}{(nb_1)^2} \sum_{k=m}^{\infty} c_k^2(t,w)k! \sum_{i,j=n(t-b_1)}^{n(t+b_1)} \gamma_Z^k(i-j) = \; O\left(\frac{1}{nb_1}\right).$$

Remark 1 When $W_1, W_2, \ldots, W_n$ are independent, the dominating term in the variance of $\hat{F}_t(w)$ is

$$\frac{1}{nb_1} F_t(w)[1 - F_t(w)] \int_{-1}^{1} K_1^2(u)du.$$

Since the expression for the bias does not depend on the underlying correlation structure, it remains the same as in Theorem 1.

Remark 2 Theorem 1 immediately implies that the mean squared error of the estimator $\hat{F}_t$ converges to zero as $n \to \infty$ implying consistency of $\hat{F}_t$. Finally, since the quantile function $\theta_\alpha(t)$ is obtained by inverting $\hat{F}_t$ (equation (4)), consistency of $\hat{\theta}_\alpha(t)$ also follows.

3 Smoother quantiles

Tukey (1977) describes many smoothing methods, their advantages and disadvantages being analyzed via numerous applications. He illustrates the idea to "make the smooth still smoother" (p. 534) via running medians, repeated medians, hanning and combinations of these, the basic idea being that "anything we can do once, we can do twice" (p. 234). Similar to these procedures, Wu and Chu (1992) consider "double smoothing" to estimate the mean function in the classic nonparametric regression model with independent identically distributed data. They apply the Gasser-Müller kernel estimator to the smooth estimator of the regression curve obtained by kernel smoothing via the Nadaraya-Watson kernel and study the asymptotic properties of this double smoothed estimator. In this paper we discuss a second kernel estimate of the quantile functions that also essentially follows this principle. The method is similar to Tukey's twicing (p.526), because in our procedure kernel smoothing is done twice, but, as in Wu and Chu (1992), the kernels need not be the same. The method is illustrated in Figure **??**.

We thus define the "smoother" estimator of the probability distribution function of $W_i,\ i = 1, 2, \ldots$ (see (1.1))

$$\tilde{F}_t(w) \;=\; \frac{1}{nb_2} \sum_{i=1}^{n} K_2\Big(\frac{t_i - t}{b_2}\Big) \cdot \hat{F}_{t_i}(w), \quad t \in [0,1],\ w \in \mathbb{R}, \tag{3.8}$$

where $\hat{F}_{t_i}(w)$ are constructed as in (2.5), K_1 and K_2 are kernels and the bandwidth b_2 is defined as b_1.

To keep notations simple, "hat" will denote the kernel estimator defined in (5) and "tilde" will be used to denote the smoother estimator. The quantiles can then be estimated as

$$\tilde{\theta}_\alpha(t) = inf\{w \in \mathbb{R} \big| \tilde{F}_t(w) \geq \alpha\}, \quad t \in [0,1],\ 0 < \alpha < 1. \tag{3.9}$$

In what follows we discuss the consistency of this estimator when $W_1, W_2, \ldots, W_n$ are independently distributed. Note that this is the classical assumption for many statistical applications including standard regression models. Under the same conditions of Theorem 1 and as $n \to \infty$, we have

Theorem 2 *(a) Bias:*

$$E\Big(\tilde{F}_t(w)|t\Big) - F_t(w) = \frac{\partial^2}{\partial t^2}\Big[F_t(w)\Big] \cdot \Big(\frac{b_1^2}{2}\int_{-1}^{1} u^2 K_1(u)du + \frac{b_2^2}{2}\int_{-1}^{1} u^2 K_2(u)du\Big) + \\ + o\Big(max\big(b_1^2, b_2^2\big)\Big). \tag{3.10}$$

(b) Variance:

$$Var\Big(\tilde{F}_t(w)|t\Big) = \frac{\int_{-1}^{1} K_1^2(u)du}{nb_1} \cdot \frac{\int_{-1}^{1} K_2^2(u)du}{nb_2} F_t(w)\Big(1 - F_t(w)\Big) + \\ + o\Big(max\left(\frac{1}{nb_1}, \frac{1}{nb_2}\right)\Big). \tag{3.11}$$

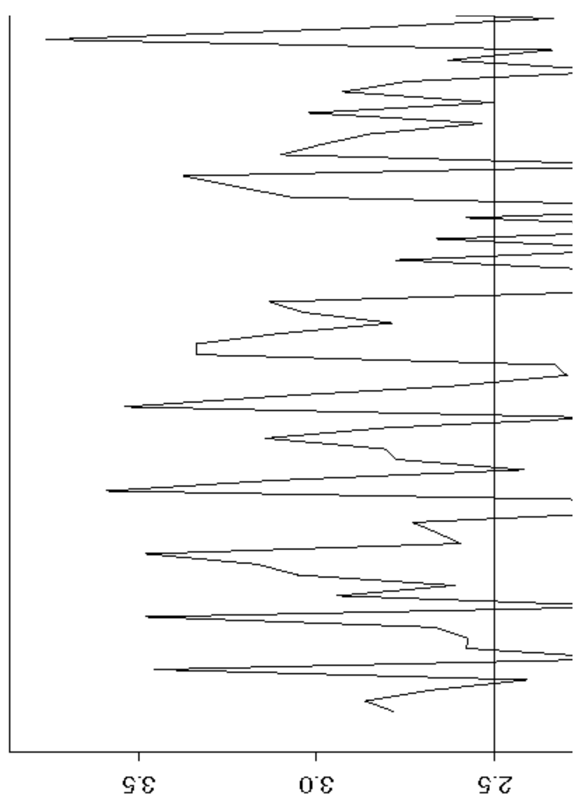
Remark 3 While the bias of the "smoother" estimator will be larger or smaller than the bias of the first kernel estimator depending on the curvature of the distribution function $F_t(w)$ at t, an appropriate choice of b_2 can lead to smaller variance in the "smoother" method. In particular, a proper choice of the bandwidth b_2 can lead to a more efficient "smoother" estimate of the distribution function and hence of the quantiles. For related discussions in the context of nonparametric trend estimation from replicated time series, see Ghosh (2001).

4 A data example

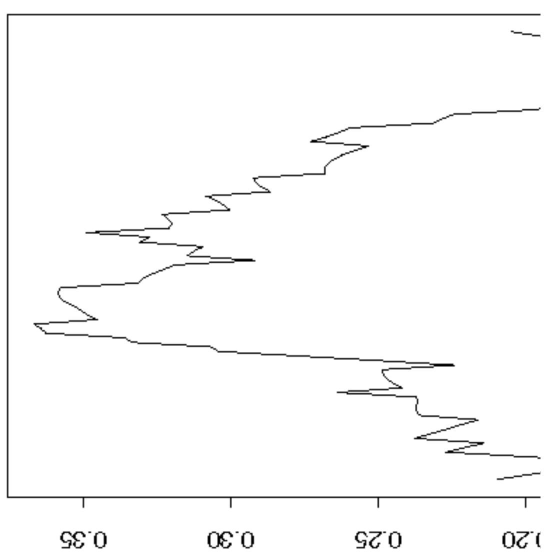
Figure 1 illustrates the time series of yearly means of daily precipitation (in mm) in Bern during 1901-1999. The horizontal line corresponds to the value of the yearly mean of daily precipitation equal to 2.5 mm. We estimated

$$P(Yearly\ mean\ of\ daily\ precipitation\ \leq 2.5\ mm|\ year)$$

by using both (2.5) and (3.8) - see Figure 2. We used the truncated Gaussian density kernel as K_1 and the box kernel as K_2.


2.5
3.0
3.5


0.30
0.25


0.35
0.30
0.25
0.20

5 Appendix

Proof of Theorem 1

(a) Bias: The proof of the bias follows from standard arguments in particular by symmetry of the kernel K_1 around zero as well as the due to the fact that the function K_1 integrates to unity.

(b) Variance: The dominating term is

$$V_n(t,w) = \frac{K_1^2(0)}{(nb_1)^2} \sum_{k=m}^{\infty} c_k^2(t,w)k! \sum_{u=-nb_1}^{nb_1} \big(2nb_1 + 1 - |u|\big)k!\gamma_Z^k(u).$$

Now, following the line of proof given in Ghosh (2001), note that

$$V_n(t,w) = K_1^2(0)\Big(2A_n + B_n + C_n\Big)$$

where

$$A_n = \frac{1}{nb_1} \sum_{k=m}^{\infty} \Big[c_k^2(t,w)k! \cdot \sum_{u=-nb_1}^{nb_1} \gamma_Z^k(u)\Big],$$

$$B_n = \frac{1}{(nb_1)^2} \sum_{k=m}^{\infty} \Big[c_k^2(t,w)k! \cdot \sum_{u=-nb_1}^{nb_1} \gamma_Z^k(u)\Big],$$

and

$$C_n = \frac{1}{(nb_1)^2} \sum_{k=m}^{\infty} \Big[c_k^2(t,w)k! \cdot \sum_{u=-nb_1}^{nb_1} |u|\gamma_Z^k(u)\Big].$$

The proof is completed by noting that $B_n = o(A_n),\ C_n = o(nb_1)$ and $\gamma_Z^k(l) \leq \gamma_Z^m(l),\ \forall l,\ \forall k \geq m.$

Aknowledgement. The research of Dana Draghicescu is supported by the Swiss Federal Research Institute WSL and a PhD student grant of the Swiss National Science Foundation. The precipitation series (source: SMA, Zürich) was provided by Dr. Christoph Frei, Climate Research Unit, ETH Zürich, whose help is gratefully acknowledged.

References

[1] Abberger, K. (1997) Quantile smoothing in financial time series. *Statistical Papers,* 38, 125-148.

[2] Abberger, K. (1996) Nichtparametrische Schätzung bedingter Quantile in Zeitreihen. Konstanzer Dissertation. Band 530, Hartung-Gorre, Konstanz.

[3] Beran, J. (1994) Statistics for long-memory processes. Chapman and Hall.

[4] Cox, D.R. (1984) Long-range dependence: A review. Statistics: An Appraisal. *Proceedings 50th Anniversary Conference. H.A. David, H.T. David (eds.).* The Iowa State University Press, 55-74.

[5] Draghicescu, D., Ghosh, S. (2000) Different aspects for quantile estimation for transformation of Gaussian processes. *Abstracts,* Annual symposium on Forecasting, Lisbon, Portugal, 21-26 June 2000.

[6] Draghicescu, D., Ghosh, S. (2001) Nonparametric quantile estimation for time series data. *Proceedings,* The 53rd Session of the ISI, Seoul, South Korea, 22-29 August 2001.

[7] Gasser, T., Müller, H. (1979) Kernel estimation of regression functions. Smoothing techniques for curve estimartion, *Proceedings,* Heidelberg 1979, 23-68.

[8] Ghosh, S. (2001) Nonparametric trend estimation in replicated time series. *Journal of Statistical Planning and Inference*, Vol. 97, 263-274.

[9] Ghosh, S., Beran, J. Innes, J. (1997) Nonparametric conditional quantile estimation in the presence of long memory. *Student* Vol. 2, No. 2, 109-117.

[10] Ghosh, S., Draghicescu, D. (2001) Quantile estimation to assess extreme climate events. *Geophysical Research Abstracts,* Vol. 3, European Geophysical Society, Nice, France, 25-30 March 2001.

[11] Ghosh, S., Draghicescu, D. (2002) Predicting the distribution function for long-memory processes. *International Journal of Forecasting* (to appear).

[12] Tukey, J. (1977) Exploratory data analysis. Adison-Wesley.

[13] Wand, M., Jones, M. (1995) Kernel Smoothing. Chapman and Hall.

[14] Wu, J. S., Chu, C. K. (1992) Double smoothing for kernel estimators in nonparametric regression. *Nonparametric Statistics,* Vol. 1, 375-386.

Dana Draghicescu
Swiss Federal Research Institute WSL, 8903 Birmendorf Switzerland,
Email: dana.draghicescu@wsl.ch, Tel. +41-1-739-2549, Fax +41-1-739-2215

Sucharita Ghosh
Swiss Federal Research Institute WSL, 8903 Birmendorf Switzerland,
Email: rita.ghosh@wsl.ch, Tel. +41-1-739-2431, Fax +41-1-739-2215