Journal of Multivariate Analysis 146 (2016) 29–45

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

On the estimation of the functional Weibull tail-coefficient

Laurent Gardes [a], Stéphane Girard [b,*]

[a] Université de Strasbourg & CNRS, IRMA, UMR 7501, 7 rue René Descartes, 67084 Strasbourg Cedex, France

[b] Inria Grenoble Rhône-Alpes & LJK, team Mistis, Inovallée, 655, av. de l'Europe, Montbonnot, 38334 Saint-Ismier cedex, France

ARTICLE INFO

Article history:
Received 12 September 2014
Available online 15 May 2015

AMS subject classifications:
62G32
62G05
62E20

Keywords:
Conditional Weibull tail-coefficient
Extreme quantiles
Nonparametric estimation

ABSTRACT

We present a nonparametric family of estimators for the tail index of a Weibull tail-distribution when functional covariate is available. Our estimators are based on a kernel estimator of extreme conditional quantiles. Asymptotic normality of the estimators is proved under mild regularity conditions. Their finite sample performances are illustrated both on simulated and real data.

© 2015 Elsevier Inc. All rights reserved.

1. Introduction

Weibull tail-distributions encompass a variety of light-tailed distributions, such as Weibull, Gaussian, gamma and logistic distributions. Let us recall that a cumulative distribution function F has a Weibull tail if it satisfies the following property: there exists $\theta > 0$ such that for all $t > 0$,

$$\lim_{y\to\infty} \frac{\log(1-F(ty))}{\log(1-F(y))} = t^{1/\theta}. \quad (1)$$

The parameter θ is referred to as the Weibull tail-coefficient. A general account on Weibull tail-distributions can be found in [6], see also [5] for an application to the modeling of large claims in non-life insurance. Dedicated methods have been proposed to estimate the Weibull tail-coefficient since the relevant information is only contained in the extreme upper part of the sample denoted hereafter by $Y_1, \dots, Y_n$. A first direction was investigated in [8] where an estimator based on the record values is proposed. Another family of approaches [3,4,11,17] consists of using the k_n upper order statistics $Y_{n-k_n+1,n} \le \cdots \le Y_{n,n}$ where $k_n \to \infty$ as $n \to \infty$. Note that, since θ is defined through an asymptotic behavior of the tail, the estimator should only use the extreme-values of the sample and thus the extra condition $k_n/n \to 0$ is required. More specifically, most recent estimators are based on the log-spacings between the k_n upper order statistics [6,16,24–27].

Here, we focus on the situation where some covariate information X is recorded simultaneously with the quantity of interest Y. In the general case, the tail heaviness of Y given X depends on X, and thus the Weibull tail-coefficient is a function $\theta(X)$ of the covariate. When the covariate is finite dimensional, some new tools have been introduced [15,14] to estimate extreme conditional quantiles. We refer to [18] for an application to the risk modeling associated with extreme rainfalls. In this case, the selected covariate is the geographical location but other relevant informations could be included such as

* Corresponding author.
E-mail address: Stephane.Girard@inria.fr (S. Girard).

http://dx.doi.org/10.1016/j.jmva.2015.05.007
0047-259X/© 2015 Elsevier Inc. All rights reserved.

climatic curves. More generally, covariates may be curves (electricity price/demand curves, medical curves,...) in many other situations coming from applied sciences, see [10, Paragraph 1.2.2]. However, the estimation of the Weibull tail-coefficient with functional covariates has not been addressed yet. Our approach relies on the use of $\hat{q}_n$ a functional kernel estimator of conditional quantiles, see [20] for an example. Similarly to the unconditional case, the estimation of $\theta(X)$ is based on the extreme observations of $Y|X$. Therefore, a close study of the asymptotic properties of $\hat{q}_n$ when estimating extreme quantiles is necessary. Two statistical fields are thus involved in this study: nonparametric smoothing techniques adapted to functional data are required in order to deal with the covariate X while extreme-value analysis is used to study the tail behavior of $Y|X$.

The family of nonparametric functional estimators is introduced in Section 2 and its asymptotic normality is established. A particular sub-family of estimators is exhibited in Section 3, their finite sample behavior is illustrated on some simulated data in Section 4 and on a real dataset in Section 5. Proofs are postponed to Section 6.

2. Main result

Let $(X_i, Y_i),\ i = 1, \dots, n$, be independent copies of a random pair $(X, Y) \in E \times \mathbb{R}$ where E is an arbitrary space associated with a semi-metric d. Recall that a semi-metric (or pseudometric) may allow the distance between two different points to be zero, see [21, Definition 3.2]. The conditional survival function of Y given $X = x \in E$ is denoted by $\bar{F}(y|x) := \mathbb{P}(Y > y|X = x)$ and is supposed to be continuous and strictly decreasing with respect to y. Discussing the existence of regular versions of $\bar{F}(.|.)$ is beyond the scope of this paper. Let us just note that such an existence is insured when (E, d) is a Polish space [30]. The associated conditional cumulative hazard function is defined by $H(y|x) := -\log \bar{F}(y|x)$ and the conditional quantile is therefore given by $q(\alpha|x) := \bar{F}^{-1}(\alpha|x) = H^{-1}(\log(1/\alpha)|x)$, for all $\alpha \in (0, 1)$. In this paper, we focus on conditional Weibull tail-distributions. In such a case, analogously to (1), $H(.|x)$ is a regularly varying function with index $1/\theta(x)$, *i.e.*

$$\lim_{y\to\infty} \frac{H(ty|x)}{H(y|x)} = t^{1/\theta(x)},$$

for all $t > 0$. In this situation, $\theta(.)$ is an unknown positive function of the covariate $x \in E$ referred to as the functional Weibull tail-coefficient. From [9, Theorem 1.5.12], $H^{-1}(.|x)$ is also a regularly varying function with index $\theta(x)$ and thus, there exists a slowly-varying function $\ell(.|x)$ such that

$$q(e^{-y}|x) = H^{-1}(y|x) = y^{\theta(x)}\ell(y|x). \tag{2}$$

Recall that the slowly-varying function $\ell(.|x)$ is such that

$$\lim_{y\to\infty} \frac{\ell(ty|x)}{\ell(y|x)} = 1, \tag{3}$$

for all $t > 0$, see [9] for a general account on regular variation theory. In view of (2), it appears that the functional Weibull tail-coefficient drives the asymptotic behavior of conditional extreme quantiles. The object of interest is thus $\theta(x)$ where x is in some arbitrary semi-metric space (E, d). The most usual case is $E = \mathbb{R}^p$, but our framework also includes the infinite dimensional case, for instance when x is a curve. We propose a family of estimators of $\theta(x)$ based on some properties of the log-spacings of the conditional quantiles: let $\alpha \in (0, 1)$ be small enough and $\tau \in (0, 1)$,

$$\begin{aligned}\log q(\tau\alpha|x) - \log q(\alpha|x) &= \log H^{-1}(-\log(\tau\alpha)|x) - \log H^{-1}(-\log(\alpha)|x)\\ &= \theta(x)(\log_{-2}(\tau\alpha) - \log_{-2}(\alpha)) + \log\left(\frac{\ell(-\log(\tau\alpha)|x)}{\ell(-\log(\alpha)|x)}\right)\\ &\approx \theta(x)(\log_{-2}(\tau\alpha) - \log_{-2}(\alpha)) \approx \theta(x)\frac{\log(1/\tau)}{\log(1/\alpha)},\end{aligned} \tag{4}$$

where $\log_{-2}(.) := \log\log(1/.)$, see Lemma 2 in Section 6 for a more precise asymptotic expansion. Hence, for a decreasing sequence $0 < \tau_J < \cdots < \tau_1 \le 1$, where J is a positive integer, and for all functions ϕ satisfying the shift and location invariance condition

(A.1) $\phi : \mathbb{R}^J \to \mathbb{R}$ is a twice differentiable function such that $\phi(\eta z) = \eta\phi(z)$, $\phi(\eta u + z) = \phi(z)$ for all $\eta \in \mathbb{R} \setminus \{0\}$, $z \in \mathbb{R}^J$ and where $u = (1, \dots, 1)^t \in \mathbb{R}^J$,

one has:

$$\theta(x) \approx \log(1/\alpha)\frac{\phi(\log q(\tau_1\alpha|x), \dots, \log q(\tau_J\alpha|x))}{\phi(\log(1/\tau_1), \dots, \log(1/\tau_J))}. \tag{5}$$

Thus, the estimation of $\theta(x)$ relies on the estimation of conditional quantiles $q(.|x)$. This problem is addressed using a two-step estimator. First, $\bar{F}(y|x)$ is estimated by the kernel estimator defined for all $(x, y) \in E \times \mathbb{R}$ by

$$\hat{\bar{F}}_n(y|x) = \sum_{i=1}^n K(d(x, X_i)/h_n)\mathbb{I}\{Y_i > y\} \Big/ \sum_{i=1}^n K(d(x, X_i)/h_n), \tag{6}$$

Download English Version:

https://daneshyari.com/en/article/1145209

Download Persian Version:

https://daneshyari.com/article/1145209

Daneshyari.com