Journal of Computational and Applied Mathematics 315 (2017) 354–364

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

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

Recovery of a quantile function from moments

Robert M. Mnatsakanov [a,*], Aleksandre Sborshchikovi [b]

[a] Department of Statistics, P.O. Box 6330, West Virginia University, Morgantown, WV 26506, USA

[b] Faculty of Exact and Natural Sciences, I. Javakhishvili Tbilisi State University, 1 Il. Chavchavadze Ave., Tbilisi 0179, Georgia

ARTICLE INFO

Article history:
Received 29 August 2016
Received in revised form 17 November 2016

Keywords:
Quantile function
Moment-recovered approximation
Laplace transform inversion
The Lorenz curve

ABSTRACT

The problem of recovering a quantile function of a positive random variable via the values of moments or given the values of its Laplace transform is studied. Two new approximations as well as two new estimates of a quantile function given the sample from underlying distribution are proposed. The uniform and L_1 upper bounds of proposed estimates are derived. The plots illustrate the behavior of the recovered approximants for the moderate and large sample sizes.

1. Introduction

The main aim of this work is to construct accurate approximations of a quantile function Q under incomplete information upon the underlying distribution function (df) F. Namely, we assume that the only available data represent the values of a finite number of moments ("cumulative frequency" moments, the Laplace transform) of F or Q.

We propose two methods for approximation and estimation of a quantile function Q. The first method (method I) is based on the knowledge of the so-called "cumulative frequency" moments of F. The second one (method II) enables us to recover Q when only the integer order moments of F are known (or estimated) up to some order $\alpha \in \mathbb{N}$.

There are many different nonparametric methods already proposed for estimating the quantile function Q. Let us mention only some of them. For example, the estimator suggested by Harrell and Davis [1] is based on a weighted sum of order statistics. Sheather and Marron [2] studied properties of kernel based approaches that require choosing the smoothing parameter. Swanepoel and Van Graan [3] suggested the kernel type estimate of a quantile function using the nonparametric transformation technique (see also [4] for details). The method based on the cross-validation has been proposed in [5]. Asymptotic equivalence between various estimators based on weighted order statistics were investigated by Sheather and Marron [2]. Brewer [6] constructed quantile estimate based on Bernstein polynomials.

The advantage of our approach is in its applicability to situations where only ordinary, frequency moments, or the finite number of values of the Laplace transform of underlying distribution F are available. In such cases the traditional estimates proposed, say, in [1,7,8], could not be applied. On the other hand, when observations from F are available, i.e., F is observed directly, the moment-recovered (MR)-constructions (see (2) and (11), as well as (28)) provide the estimator similar to the one proposed in [1]. It is worth noting that the weak convergence of properly normalized empirical quantile function based on the empirical distribution has been proved in Shorack and Wellner [9]. The question about weak convergence of the estimate based on MR-construction will be the subject of our study in a separate article.

The rest of the paper is organized as follows. In Section 2 two MR-approximants of the quantile function Q are introduced, and the rate of approximation in sup- and L_1-norms are derived. Two cases when the support of underlying df F is finite and

* Corresponding author.
E-mail address: Robert.Mnatsakanov@mail.wvu.edu (R.M. Mnatsakanov).

http://dx.doi.org/10.1016/j.cam.2016.11.028
0377-0427/

infinite are considered. The weak upper bounds for corresponding estimates of Q are established as well. In Section 3 we applied the MR-construction for estimating the Lorenz curve L. The weak upper bound in sup-norm of the plug-in estimator of L is derived. Simulation study is conducted and the graphical illustrations for several examples are presented.

2. Moment-recovered constructions

Assume F is an absolute continuous with respect to the Lebesgue measure on $\mathbb{R}_+$. Denote corresponding density function by f. In the sequel we always assume that the moment sequence $\{\int t^j\, dF(t), j \in \mathbb{N}\}$ as well as the sequence of "cumulative frequency" moments $\{\int [F(t)]^j\, dt, j \in \mathbb{N}\}$ of F determines F uniquely. Denote by $\|f\| = \sup_{0\le x\le T} |f(x)| < \infty$ and $\|f\|_{L_1} = \int |f(t)|dt$ the sup- and L_1-norm, respectively.

2.1. Method I

Consider the following moment sequence $m_F^- = \{m^-(j, F), j = 0, 1, \ldots, \alpha\}$ with

$$m^-(j, F) = \int_0^T [F(t)]^j\, dt. \tag{1}$$

We refer to [10], where the question of moment-determination based on the so-called frequency moments is studied. Without loss of generality let us assume $T = 1$. For each $x \in [0, 1]$ consider the following approximation of a quantile function $Q(x) = \inf\{t : F(t) \ge x\}$:

$$Q_\alpha^-(x) = \left(\mathcal{K}_\alpha^{-1} m_F^-\right)(x), \tag{2}$$

where

$$\left(\mathcal{K}_\alpha^{-1} m_F^-\right)(x) = \sum_{k=0}^{\lfloor \alpha x\rfloor} \sum_{j=k}^{\alpha} \binom{\alpha}{j}\binom{j}{k} (-1)^{j-k}\, m^-(j, F), \quad x \in [0, 1]. \tag{3}$$

Here $\alpha \in \mathbb{N}$ is a parameter that specifies the number of moments used in the proposed approximation (cf. with [11]). Here, in (3) and in the sequel, we will use notations $\lfloor \alpha x\rfloor$ and $\lceil \alpha x\rceil$ to denote the integer and the rounding part of αx, respectively.

Theorem 1. *If Q', Q'' are bounded on $(0, 1)$, then*

$$Q_\alpha^-(x) - Q(x) = \frac{1}{\alpha+1}\left[\left(1 - x + \lfloor \alpha x\rfloor - \alpha x\right)Q'(x) + \frac{1}{2}x(1-x)\, Q''(x)\right] + o\left(\frac{1}{\alpha}\right) \tag{4}$$

and, if in (3)–(4) we use $\lceil \alpha x\rceil$ instead of $\lfloor \alpha x\rfloor$, then

$$\|Q_\alpha^- - Q\| \le \frac{1}{\alpha+1}\left\{\frac{3}{2}\|Q'\| + \frac{1}{8}\|Q''\|\right\} + o\left(\frac{1}{\alpha}\right), \quad \text{as } \alpha \to \infty.$$

Proof. Substitution of (1) in right hand side of (3) yields another representation of Q_α^-:

$$Q_\alpha^-(x) = \int_0^1 B_\alpha(F(t), x)\, dt = \int_0^1 B_\alpha(t, x)\, dQ(t). \tag{5}$$

Here, the integrant in (5) has very simple form with

$$B_\alpha(t, x) = \sum_{k=0}^{\lfloor \alpha x\rfloor} \binom{\alpha}{k} t^k (1-t)^{\alpha-k}, \quad 0 < t < 1. \tag{6}$$

Now, integration by parts in the right-hand side of (5) gives:

$$Q_\alpha^-(x) - Q(x) = \int_0^1 B_\alpha(t, x)\, dQ(t) - Q(x) = \int_0^1 [Q(t) - Q(x)]\, \beta(t, c, d-1)dt. \tag{7}$$

Here $\beta(\cdot, c, d)$ is a density function of a Beta distribution with the shape parameters $c = \lfloor \alpha x\rfloor + 1$ and $d = \alpha - \lfloor \alpha x\rfloor + 1$. It forms a δ sequence as $\alpha \to \infty$ with mean, $\theta_\alpha = \frac{\lfloor \alpha x\rfloor + 1}{\alpha+1}$ and variance

$$\sigma_\alpha^2 = \frac{(\lfloor \alpha x\rfloor + 1)(\alpha - \lfloor \alpha x\rfloor)}{(\alpha+1)^2(\alpha+2)}.$$

Note also that $\theta_\alpha - x = \frac{1-x}{\alpha+1} + \Delta_{1,\alpha}(x)$ with $\Delta_{1,\alpha}(x) = \frac{\lfloor \alpha x\rfloor - \alpha x}{\alpha+1}$ and $\sigma_\alpha^2 = \frac{x(1-x)}{(\alpha+2)} + \Delta_{2,\alpha}(x)$, where $\Delta_{2,\alpha} \le \frac{2}{(\alpha+2)^2}$ (see [12,13]). Applying the Taylor series expansion of $Q(t)$ around $t = x$ in (7) and using similar steps performed in the proof of Theorem 1(ii) in [14], we derive (4). □