Statistics and Probability Letters xx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Statistics and Probability Letters

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

Recovery of quantile and quantile density function using the frequency moments

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

[a] Department of Statistics, P.O. Box, WV 26506, 6330, West Virginia University, Morgantown, 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 3 November 2017
Received in revised form 13 February 2018
Accepted 21 April 2018
Available online xxxx

Keywords:
Quantile function
Quantile density function
Frequency moments
Expected shortfall

ABSTRACT

The problem of recovering quantiles and quantile density functions of a positive random variable via the values of frequency moments is studied. The uniform upper bounds of the proposed approximations are derived. Several simple examples and corresponding plots illustrate the behavior of the recovered approximations. Some applications of the constructions are discussed as well. Namely, using the empirical counterparts of the constructions yield the estimates of the quantiles, and the quantile density functions. By means of simulations, the average errors in terms of L_2-norm are evaluated to justify the consistency of the estimate of the quantile density function. As an application of the constructions, the question of estimating the so-called expected shortfall measure in risk models is also studied.

1. Introduction and preliminaries

Recovery of the quantiles and the quantile density functions have become a popular area of investigation since 1980s. See, for instance, the papers of Parzen (1979), Harrell and Davis (1982), Jones (1992), Cheng and Parzen (1997), and Yang (1985) among others. Most of approaches are based on using the kernel smoothing technique applied to the empirical quantile function. Wei et al. (2015) proposed the so-called semi-parametric, tail-extrapolated quantile estimators and studied their performances for small sample sizes.

Recently, in Mnatsakanov and Sborshchikovi (2017) two new approximations of a quantile function $Q(x) = \inf\{t : F(t) \geq x\}$ were suggested and some asymptotic properties of the corresponding approximants were investigated in two different cases: when the support of the distribution function (df) F is finite, say, $supp\{F\} = (0, 1)$, and when it is unbounded from above with $supp\{F\} = (0, \infty)$. In particular, in the former case the approximation Q_α^- of Q was constructed using the "information" contained in the cumulative frequency moments of the underlying df F:

$$m^-(j, F) = \int_0^1 [F(t)]^j \, dt, \quad j = 0, 1, \ldots, \alpha.$$

Namely,

$$Q_\alpha^-(x) = \sum_{k=0}^{\lfloor \alpha x \rfloor} \sum_{j=k}^{\alpha} \binom{\alpha}{j} \binom{j}{k} (-1)^{j-k} m^-(j, F), \quad x \in (0, 1). \tag{1}$$

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

https://doi.org/10.1016/j.spl.2018.04.022

Here, the parameter $\alpha \in \mathbb{N}_+$ is associated with the number of used frequency moments, and determines the order of approximation. In (1) and in the sequel, we will use the symbol $\lfloor a \rfloor$ to denote the integer part of a, while by $\lceil a \rceil$ we denote the rounding part of a.

When the support of F is $(0, \infty)$, another approximation $Q_{\alpha,b}$ of Q, based on a two-step procedure, was proposed (see Mnatsakanov and Sborshchikovi (2017)). In the first step, we used the values of the Laplace transform $\mathcal{L}_F$ evaluated at points $s \in \{0, \ln b, 2\ln b, \ldots, \alpha \ln b\}$:

$$\mathcal{L}_F(j \ln b) := \int_{\mathbb{R}_+} e^{-j\,\ln(b)\,\tau}\, dF(\tau), \quad \text{for} \quad j = 0, 1, \ldots, \alpha, \quad b > 1,$$

in the following approximation of F:

$$F_{\alpha,b}(x) = 1 - \sum_{k=0}^{\lfloor \alpha b^{-x} \rfloor} \sum_{j=k}^{\alpha} \binom{\alpha}{j}\binom{j}{k} (-1)^{j-k}\, \mathcal{L}_F(j \ln b), \quad x \in \mathbb{R}_+.$$

In the second step, the quantile function Q was recovered by means of the approximation:

$$Q_{\alpha,b}(x) = \int_0^\infty B_\alpha(F_{\alpha,b}(u), x)\, du, \quad x \in (0, 1), \tag{2}$$

where

$$B_\alpha(t, x) = \sum_{k=0}^{\lfloor \alpha x \rfloor} \binom{\alpha}{k} t^k (1-t)^{\alpha-k}, \quad t \in (0, 1). \tag{3}$$

This approximation $Q_{\alpha,b}$ depends upon values of two parameters: $\alpha \in \mathbb{N}_+$ and some real number $b > 1$.

In the current work, two new approximants $Q_{\alpha,S}$ and $q_{\alpha,S}$ (see Eqs. (5) and (7)) of the quantile Q and quantile density function $q = Q'$ are proposed when F is supported by $\mathbb{R}_+ = (0, \infty)$.

The main goal of this paper is to introduce a one-step procedure for recovering the quantiles and the quantile density functions, and investigate their asymptotic properties as $\alpha \to \infty$. Note that the proposed construction $Q_{\alpha,S}$ (see (5)) depends only on one parameter α, and its evaluation time is shorter then that of $Q_{\alpha,b}$. Besides, in some models the latter one cannot be used, while the former one produces considerably better approximations of Q (see Fig. 7(b) and Remark 3 on p. 12). That is why we recommend to use $Q_{\alpha,S}$ instead of $Q_{\alpha,b}$.

In proposed constructions we are assuming that the sequence of frequency moments $m_S^- = \{m^-(j, S), j = 1, \ldots, \alpha\}$ is given, with

$$m^-(j, S) = \int_0^\infty [S(t)]^j\, dt. \tag{4}$$

Here $S = 1 - F$ is the survival function of F. We assume also that $m^-(1, S) = E(X)$ is finite; that is why the values of $m^-(j, S), j = 1, 2, \ldots, \alpha$, are finite as well. Consider the following approximation of Q:

$$Q_{\alpha,S}(x) = \left(\bar{\mathcal{K}}_\alpha^{-1} m_S^-\right)(x), \quad x \in (0, 1), \tag{5}$$

where

$$\left(\bar{\mathcal{K}}_\alpha^{-1} m_S^-\right)(x) = \sum_{k=\alpha-\lfloor \alpha x \rfloor}^{\alpha} \sum_{j=k}^{\alpha} \binom{\alpha}{j}\binom{j}{k} (-1)^{j-k}\, m^-(j, S). \tag{6}$$

The reader is referred to Remark 1 in Mnatsakanov et al. (2015), where the rationale of the formula (1) is explained when the ordinary moments of F are used instead of frequency moments $m^-(j, S), j = 1, \ldots, \alpha$. Also it is worth mentioning that to have a better performance of approximation $Q_{\alpha,S}(x)$ and to evaluate its value at $x = 1$, we need to exclude undefined expression $m^-(0, S)$ from construction (6). Now, we can display the curve of $Q_{\alpha,S}$ on the entire range $[0, 1]$; we recommend using $\lfloor (\alpha - 1)x \rfloor$ instead of $\lfloor \alpha x \rfloor$ in (6). See, for example, Fig. 1(a).

To recover the quantile density function q let us consider the approximation from Mnatsakanov (2017), (see Eq. (5) therein) with $m_\phi(j) = m^-(j, S)$ and $\phi(x) = 1 - x$ for $x \in (0, 1)$. We obtain

$$q_{\alpha,S}(x) := \left(\mathcal{B}_\alpha^{-1} m_S^-\right)(x) = \frac{|\phi'(x)|\, \Gamma(\alpha+2)}{\Gamma(\lfloor \alpha\phi(x) \rfloor + 1)} \sum_{j=0}^{\alpha - \lfloor \alpha\phi(x) \rfloor} \frac{(-1)^j\, m^-(j + \lfloor \alpha\phi(x) \rfloor, S)}{j!\, (\alpha - \lfloor \alpha\phi(x) \rfloor - j)!}. \tag{7}$$

It is worth mentioning that the suggested construction $Q_{\alpha,S}$ has advantages when compared to $Q_{\alpha,b}$ defined in (2), i.e., it depends only on one parameter α, and is constructed by means of one-step procedure.

In the sequel the following notations will be used:

$$\beta(t, c, d), \quad t \in (0, 1),$$

Please cite this article in press as: Mnatsakanov R.M., Sborshchikovi A., Recovery of quantile and quantile density function using the frequency moments. Statistics and Probability Letters (2018), https://doi.org/10.1016/j.spl.2018.04.022.