IMS Lecture Notes–Monograph Series
Optimality: The Third Erich L. Lehmann Symposium
Vol. 57 (2009) 252–265

DOI: 10.1214/09-LNMS5715

Recovery of Distributions via Moments

Robert M. Mnatsakanov[1,*] and Artak S. Hakobyan[2,*]

West Virginia University

Abstract: The problem of recovering a cumulative distribution function (cdf) and corresponding density function from its moments is studied. This problem is a special case of the classical moment problem. The results obtained within the moment problem can be applied in many indirect models, e.g., those based on convolutions, mixtures, multiplicative censoring, and right-censoring, where the moments of unobserved distribution of actual interest can be easily estimated from the transformed moments of the observed distributions. Nonparametric estimation of a quantile function via moments of a target distribution represents another very interesting area where the moment problem arises. In all such models one can apply the present results to recover a function via its moments. In this article some properties of the proposed constructions are derived. The uniform rates of convergence of the approximation of cdf, its density function, quantile and quantile density function are obtained as well.

Contents

1. Introduction

The probabilistic Stielties moment problem can be described as follows: let a sequence $\nu = \{\mu_j, j = 0, 1, \dots\}$ of real numbers be given. Find a probability distribution on the non-negative real line $\mathbb{R}_+ = [0, \infty)$, such that $\mu_j = \int t^j d\,F(t)$ for $j \in \mathbb{N} = \{0, 1, \dots\}$. The classical Stieltjes moment problem was introduced first by Stieltjes [21]. When the support of the distribution F is compact, say, $supp\{F\} = [0, T]$ with $T < \infty$, then the corresponding problem is known as a Hausdorff moment problem.

Consider two important questions related to the Stieltjes (or Hausdorff) moment problem:

(i) If the distribution F exists, is it uniquely determined by the moments $\{\mu_j\}$?

(ii) How is this uniquely defined distribution F reconstructed?

[1]Department of Statistics, West Virginia University, Morgantown, WV 26506, e-mail: rmnatsak@stat.wvu.edu

*Work supported by the National Institute for Occupational Safety and Health, contract no. 212-2005-M-12857.

[2]Department of Industrial and Management Systems Engineering, West Virginia University, Morgantown, WV 26506, e-mail: artakhak@yahoo.com

AMS 2000 subject classifications: Primary 62G05; secondary 62G20.

Keywords and phrases: probabilistic moment problem, M-determinate distribution, moment-recovered distribution, uniform rate of convergence.

If there is a positive answer to question (*i*) we say that a distribution F is moment-determinate (M-determinate), otherwise it is M-indeterminate.

In this paper we mainly address the question of recovering the M-determinate distribution (density and quantile functions) via its moments in the Hausdorff moment problem, i.e., we study question (*ii*). Another question we focus on here is the estimation of an unknown distribution and its quantile function, given the estimated moments of the target distribution.

It is known from the probabilistic moment problem that under suitable conditions an M-determinate distribution is uniquely defined by its moments. There are many articles that investigated the conditions (for example, the Carlemann's and the Krein's conditions), under which the distributions are either M-determinate or M-indeterminate. See, e.g., Akhiezer [2], Feller [6], Lin [10, 11], and Stoyanov [22–24] among others. However, there are very few works dealing with the reconstruction of distributions via their moments. Several inversion formulas were obtained by inverting the moment generating function and Laplace transform (Shohat and Tamarkin [20], Widder [27], Feller [6], Chauveau *et al.* [4], and Tagliani and Velasquez [25]). These methods are too restrictive, since there are many distributions for which the moment generating function does not exist even though all the moments are finite.

The reconstruction of an M-determinate cdf by means of mixtures having the same assigned moments as the target distribution have been proposed in Lindsay *et al.* [12]. Note that this procedure requires calculations of high-order Hankel determinants, and due to ill-conditioning of the Hankel matrices this method is not useful when the number of assigned moments is large. The reconstruction of an unknown density function using the Maximum Entropy principle with the specified ordinary and fractional moments has been studied in Kevasan and Kapur [9] and Novi Inverardi *et al.* [18], among others.

In Mnatsakanov and Ruymgaart [17] the constructions (2.2) and (3.13) (see Sections 2 and 3 below) have been introduced, and only their convergence has been established.

Different types of convergence of maximum entropy approximation have been studied by Borwein and Lewis [3], Frontini and Tagliani [7], and Novi Inverardi *et al.* [18], but the rates of approximations have not been established yet. Our construction enables us to derive the uniform rate of convergence for moment-recovered cdfs $F_{\alpha,\nu}$, corresponding quantile function Q_α, and the uniform convergence of the moment-recovered density approximation $f_{\alpha,\nu}$, as the parameter $\alpha \to \infty$. Other constructions of moment-recovered cdfs and pdfs (see, (3.13) and (3.14) in Remark 3.2) were proposed in Mnatsakanov [13, 14], where the uniform and L_1-rates of the approximations were established.

The paper is organized as follows: in Section 2 we introduce the notation and assumptions, while in Section 3 we study the properties of $F_{\alpha,\nu}$ and $f_{\alpha,\nu}$. Note that our construction also gives a possibility to recover different distributions through the simple transformations of moment sequences of given distributions (see Theorem 3.1 in Section 3 and similar properties derived in Mnatsakanov [13]: Theorem 1 and Corollary 1). In Theorem 3.3 we state the uniform rate of convergence for moment-recovered cdfs. In Theorem 3.4 as well as in Corollaries 3.2 and 3.5 we apply the constructions (2.2) and (3.11) to recover the pdf f, the quantile function Q, and the corresponding quantile density function q of F given the moments of F. In Section 4 some other applications of the constructions (2.2) and (3.11) are discussed: the uniform convergence of the empirical counterpart of (2.2), the rate of approximation of moment-recovered quantile function (see (4.4) in Section 4) along with the demixing and deconvolution problems in several particular models.

Note that our approach is particularly applicable in situations where other estimators cannot be used, e.g., in situations where only moments (empirical) are available. The results obtained in this paper will not be compared with similar results derived by other methods. We only carry out the calculations of moment-recovered cdfs, pdfs, and quantile functions, and compare them with the target distributions via graphs in several simple examples. We also compare the performances of $F_{\alpha,\nu}$ and $f_{\alpha,\nu}$ with the similar constructions studied in Mnatsakanov [13, 14] (see, Figures 1 (b) and 3 (b)). The moment-estimated quantile function $\hat{Q}_\alpha$ and well known Harrell-Davis quantile function estimator $\hat{Q}_{HD}$ (Sheather and Marron [19]) defined in (4.6) and (4.7), respectively, are compared as well (see Figure 2 (b)).

2. Notation and Assumptions

Suppose that the M-determinate cdf F is absolute continuous with respect to the Lebesgue measure and has support $[0, T]$, $T < \infty$. Denote the corresponding density function by f. Our method of recovering the cdf $F(x)$, $0 \le x \le T$, is based on an inverse transformation that yields a solution of the Hausdorff moment problem.

Let us denote the moments of F by

$$\mu_{j,F} = \int t^j \, dF(t) = (\mathcal{K}F)(j), j \in \mathbb{N}, \tag{2.1}$$

and assume that the moment sequence $\nu = (\mu_{0,F}, \mu_{1,F}, \dots)$ determines F uniquely.

An approximate inverse of the operator $\mathcal{K}$ from (2.1) constructed according to

$$\left(\mathcal{K}_\alpha^{-1}\nu\right)(x) = \sum_{k=0}^{[\alpha x]} \sum_{j=k}^{\infty} \frac{(-\alpha)^{j-k}}{(j-k)!} \frac{\alpha^k}{k!} \mu_{j,F}, \ 0 \le x \le T\,, \ \alpha \in \mathbb{R}_+, \tag{2.2}$$

is such that $\mathcal{K}_\alpha^{-1} \mathcal{K}F \to_w F$, as $\alpha \to \infty$ (see, Mnatsakanov and Ruymgaart [17]). Here $\to_w$ denotes the weak convergence of cdfs, i.e. convergence at each continuity point of the limiting cdf. The success of the inversion formula (2.2) hinges on the convergence

$$P_\alpha(t,x) = \sum_{k=0}^{[\alpha x]} \frac{(\alpha t)^k}{k!} e^{-\alpha t} \to \begin{cases} 1, & t < x, \\ 0, & t > x, \end{cases} \tag{2.3}$$

as $\alpha \to \infty$. This result is immediate from a suitable interpretation of the left hand side as a sum of Poisson probabilities.

For any moment sequence $\nu = \{\nu_j, j \in \mathbb{N}\}$, let us denote by F_ν the cdf recovered via $F_{\alpha,\nu} = \mathcal{K}_\alpha^{-1}\nu$ according to (2.2), when $\alpha \to \infty$, i.e.

$$F_{\alpha,\nu} \to_w F_\nu\,, \text{ as } \alpha \to \infty\,. \tag{2.4}$$

Note that if $\nu = \{\mu_{j,F}, j \in \mathbb{N}\}$ is the moment sequence of F, the statement (2.4) with $F_\nu = F$ is proved in Mnatsakanov and Ruymgaart [17].

To recover a pdf f via its moment sequence $\{\mu_{j,F}, j \in \mathbb{N}\}$, consider the ratio:

$$f_{\alpha,\nu}(x) = \frac{\Delta F_{\alpha,\nu}(x)}{\Delta}, \ \ \Delta = \frac{1}{\alpha}\,, \tag{2.5}$$

where $\Delta F_{\alpha,\nu}(x) = F_{\alpha,\nu}(x+\Delta) - F_{\alpha,\nu}(x)$ and $\alpha \to \infty$.

In the sequel the uniform convergence on any bounded interval in $\mathbb{R}_+$ will be denoted by $\longrightarrow_u$, while the sup-norm between two functions f_1 and f_2 by $\| f_1 - f_2 \|$. Note also that the statements from Sections 3 and 4 are valid for distributions defined on any compact $[0,T], T < \infty$. Without loss of generality we assume that F has support $[0,1]$.

3. Asymptotic Properties of $F_{\alpha,\nu}$ and $f_{\alpha,\nu}$

In this Section we present asymptotic properties of the moment-recovered cdf $F_{\alpha,\nu}$ and pdf $f_{\alpha,\nu}$ functions based on the transformation $\mathcal{K}_\alpha^{-1}\nu$ (2.2). The uniform approximation rate of $F_{\alpha,\nu}$ and the uniform convergence of $f_{\alpha,\nu}$ are derived as well.

Denote the family of all cdfs defined on $[0,1]$ by $\mathbb{F}$. The construction (2.2) gives us the possibility to recover also two non-linear operators $\mathcal{A}_k : \mathbb{F} \times \mathbb{F} \to \mathbb{F}$, $k = 1,2$, defined as follows: denote the convolution with respect to the multiplication operation on $\mathbb{R}_+$ by

$$F_1 \otimes F_2(x) = \int F_1(x/\tau)\, dF_2(\tau) := \mathcal{A}_1(F_1,F_2)(x), \quad 0 \le x \le 1, \tag{3.1}$$

while the convolution with respect to the addition operation is denoted by

$$F_1 \star F_2(x) = \int F_1(x-\tau)\, dF_2(\tau) := \mathcal{A}_2(F_1,F_2)(x), \quad 0 \le x \le 2\,.$$

For any two moment sequences $\nu_1 = \{\mu_{j,F_1}, j \in \mathbb{N}\}$ and $\nu_2 = \{\mu_{j,F_2}, j \in \mathbb{N}\}$, define $\nu_1 \odot \nu_2 = \{\mu_{j,F_1} \times \mu_{j,F_2}, j \in \mathbb{N}\}$ and $\nu_1 \oplus \nu_2 = \{\bar{\nu}_j, j \in \mathbb{N}\}$, where

$$\bar{\nu}_j = \sum_{m=0}^{j} \binom{j}{m} \mu_{m,F_1} \times \mu_{j-m,F_2}, \tag{3.2}$$

while $\mu_F^{\odot\, k} = \{\mu_{j,F}^k, j \in \mathbb{N}\}$ and $F^{\otimes\, k} = F \otimes \cdots \otimes F$ for the corresponding k-fold convolution (cf. (3.1)). Also denote by $F \circ \phi^{-1}$ the composition $F(\phi^{-1}(x)), x \in [0,1]$, with ϕ - continuous and increasing function $\phi : [0,1] \to [0,1]$.

Since cdfs $\mathcal{A}_1(F_1,F_2) = F_1 \otimes F_2$, $\mathcal{A}_2(F_1,F_2) = F_1 \star F_2$, and $F \circ \phi^{-1}$ have compact support, they all are M-determinate and have the moment sequences $\nu_1 \odot \nu_2$, $\nu_1 \oplus \nu_2$, and $\nu = \{\bar{\mu}_j\,, j \in \mathbb{N}\}$, with

$$\bar{\mu}_j = \int [\phi(t)]^j\, dF(t), \tag{3.3}$$

respectively. Hence, applying Theorem 3.1 from Mnatsakanov and Ruymgaart [17] a statement similar to the one in Mnatsakanov [13] (see Theorem 1 and Corollary 1, where $T = T' = 1$) is obtained. Besides, the following statement is true:

Theorem 3.1. *If* $\nu = \sum_{k=1}^m \beta_k\, \mu_F^{\odot\, k}$, *where* $\sum_{k=1}^m \beta_k = 1, \beta_k > 0$, *then (2.4) holds with*

$$F_\nu = \sum_{k=1}^{m} \beta_k F^{\otimes k}\,. \tag{3.4}$$

Proof. The equation (3.4) follows from Theorem 1 (i) (Mnatsakanov [13]) and the linearity of $\mathcal{K}_\alpha^{-1}\nu$. □

The construction (2.2) is also useful when recovering the quantile function $Q(t) = \inf\{x : F(x) \geq t\}$ via moments (see (4.5) in Section 4). Define $Q_\alpha = F_{\alpha,\nu_Q}$, where

$$\nu_Q = \left\{ \int_0^1 [F(u)]^j \, du, j \in \mathbb{N} \right\}. \tag{3.5}$$

The following statement is true:

Corollary 3.2. *If F is continuous, then $Q_\alpha \to_w Q$, as $\alpha \to \infty$.*

Proof. Replacing the functions ϕ and F in (3.3) by F and the uniform cdf on $[0,1]$, respectively, we obtain from Theorem 1 (iv) (Mnatsakanov [13]) that $Q_\alpha = F_{\alpha,\nu_Q} \to_w F_\nu = F^{-1}$ as $\alpha \to \infty$. □

Under additional conditions on the smoothness of F one can obtain the uniform rate of convergence in (2.4) and, hence, in Theorem 3.1 too. Consider the following condition

$$F'' = f' \text{ is bounded on } [0,1]. \tag{3.6}$$

Theorem 3.3. *If $\nu = \{\mu_{j,F}, j \in \mathbb{N}\}$, and (3.6) holds, we have*

$$\sup_{0 \leq x \leq 1} \left| F_{\alpha,\nu}(x) - F(x) \right| = O\left(\frac{1}{\alpha}\right), \quad \text{as } \alpha \to \infty. \tag{3.7}$$

Proof. Let us use the following representation

$$P_\alpha(t,x) = \mathbf{P}\{N_{\alpha t} \leq \alpha x\} = \mathbf{P}\{S_{[\alpha x]} \geq \alpha t\}.$$

Here $\{N_{\alpha t}, t \in [0,1]\}$ is a Poisson process with intensity αt, $S_m = \sum_{k=0}^m \xi_k, S_0 = 0$, with ξ_k being *iid* $Exp(1)$ random variables. Integration by parts gives

$$\begin{aligned} F_{\alpha,\nu}(x) = (\mathcal{K}_\alpha^{-1}\,\nu)(x) &= \int_0^1 \sum_{k=0}^{[\alpha x]} \frac{(\alpha t)^k}{k!} \sum_{j=k}^{\infty} \frac{(-\alpha t)^{j-k}}{(j-k)!} \, dF(t) \\ &= \int_0^1 P_\alpha(t,x) d\,F(t) = \int_0^1 \mathbf{P}\{S_{[\alpha x]} \geq \alpha t\} dF(t) \\ &= F(t)\,\mathbf{P}\{S_{[\alpha x]} \geq \alpha t\}\Big|_0^1 - \int_0^1 F(t)\, d\,\mathbf{P}\{S_{[\alpha x]} \geq \alpha t\} \\ &= \mathbf{P}\{S_{[\alpha x]} \geq \alpha\} + \int_0^1 F(t)\, d\,\mathbf{P}\{S_{[\alpha x]} \leq \alpha t\} = \int_0^\infty F(t)\, d\,\mathbf{P}\{S_{[\alpha x]} \leq \alpha t\}. \end{aligned} \tag{3.8}$$

Thus, (3.6) and the argument used in Adell and de la Cal [1] yield (3.7). □

Remark 3.1. When $supp\{F\} = \mathbb{R}_+$, $F_{\alpha,\nu}(x) = \int_0^\infty P_\alpha(t,x) d\,F(t)$ (cf. with (3.8)). According to Mnatsakanov and Klaassen [16] (see the proof of Theorem 3.1), one can derive the exact rate of approximation of $F_{\alpha,\nu}$ in the space $L_2(\mathbb{R}_+, dF)$. Namely, if the pdf f is bounded, say by $C > 0$, then

$$\int_0^\infty \Big(F_{\alpha,\nu}(x) - F(x)\Big)^2 dF(x) \leq \frac{2\,C}{\alpha}.$$

Now let us consider the moment-recovered density function $f_{\alpha,\nu}$ defined in (2.5) and denote by $\Delta(f,\delta) = \sup_{|t-s| \leq \delta} |f(t) - f(s)|$ the modulus of continuity of f, where $0 < \delta < 1$.

Theorem 3.4. *If the pdf f is continuous on $[0,1]$, then $f_{\alpha,\nu} \longrightarrow_u f$ and*

$$|| f_{\alpha,\nu} - f || \leq \Delta(f,\delta) + \frac{2\,||\,f\,||}{\alpha\delta^2} + o\Big(\frac{1}{\alpha}\Big), \quad as \quad \alpha \to \infty. \tag{3.9}$$

Proof. Since $[\alpha(x+1/\alpha)] = [\alpha x] + 1$, for any $x \in [0,1]$, we have

$$f_{\alpha,\nu}(x) = \alpha \left[\sum_{k=0}^{[\alpha x]+1} \sum_{j=k}^{\infty} \frac{(-\alpha)^{j-k}}{(j-k)!} \frac{\alpha^k}{k!} \mu_{j,F} - \sum_{k=0}^{[\alpha x]} \sum_{j=k}^{\infty} \frac{(-\alpha)^{j-k}}{(j-k)!} \frac{\alpha^k}{k!} \mu_{j,F} \right], \tag{3.10}$$

and, after some algebra (3.10) yields

$$f_{\alpha,\nu}(x) = \frac{\alpha^{[\alpha x]+2}}{\Gamma([\alpha x]+2)} \cdot \sum_{m=0}^{\infty} \frac{(-\alpha)^m}{m!} \mu_{m+[\alpha x]+1,F}\,. \tag{3.11}$$

Let $g(t,a,b)$ denote a gamma pdf with shape and rate parameters a and b, respectively. Substitution of (2.1) into the right hand side of (3.11) gives

$$\begin{aligned} f_{\alpha,\nu}(x) &= \frac{\alpha^{[\alpha x]+2}}{\Gamma([\alpha x]+2)} \int_0^1 \sum_{m=0}^{\infty} \frac{(-\alpha t)^m}{m!} t^{[\alpha x]+1}\, dF(t) \\ &= \int_0^1 g(t,[\alpha x]+2,\alpha)\, f(t) dt. \end{aligned} \tag{3.12}$$

To show (3.9), note that the pdf g in (3.12) has mean $([\alpha x]+2)/\alpha$ and variance $([\alpha x]+2)/\alpha^2$, respectively. The rest of the proof is similar to the lines of Theorem 1 (i) (Mnatsakanov [14]). □

Remark 3.2. In Mnatsakanov [13, 14] the uniform and L_1-rates of moment-recovered approximations of F and f defined by

$$F^*_{\alpha,\nu}(x) = \sum_{k=0}^{[\alpha x]} \sum_{j=k}^{\alpha} \binom{\alpha}{j} \binom{j}{k} (-1)^{j-k} \mu_{j,F} \tag{3.13}$$

and

$$f^*{}_{\alpha,\nu}(x) = \frac{\Gamma(\alpha+2)}{\Gamma([\alpha x]+1)} \sum_{m=0}^{\alpha-[\alpha x]} \frac{(-1)^m\, \mu_{m+[\alpha x],F}}{m!\,(\alpha-[\alpha x]-m)!}, \quad x \in [0,1]\,, \ \alpha \in \mathbb{N}, \tag{3.14}$$

are established. In Section 4, see Example 4.2, the cdf $F(x) = x^3 - 3\,x^3 \ln x$ and its density function $f(x) = -9x^2 \ln x, 0 \leq t \leq 1$, are recovered using $F_{\alpha,\nu}$ and $F^*{}_{\alpha,\nu}$, and $f_{\alpha,\nu}$ and $f^*{}_{\alpha,\nu}$ constructions, (see Figures 1 (b) and 3 (b), respectively).

The formulas (3.11) and (3.14) with $\nu = \nu_Q$ defined according to (3.5) can be used to recover a quantile density function

$$q(x) = Q'(x) = \frac{1}{f(F^{-1}(x))}\,, \quad x \in [0,1].$$

For example, consider $f_{\alpha,\nu_Q} := q_\alpha$: the application of the first line in (3.12) with F^{-1} instead of F yields

$$q_\alpha(x) = \int_0^1 g(F(u),[\alpha x]+2,\alpha)\, du \tag{3.15}$$

and corresponding moment-recovered quantile density function

$$q_{\alpha,\beta}(x) = \int_0^1 g(F_{\beta,\nu}(u), [\alpha x] + 2, \alpha)\, du\,, \quad \alpha\,, \beta \in \mathbb{N}.$$

Here $F_{\beta,\nu}$ is a moment-recovered cdf of F. As a consequence of Theorem 3.4 we have the following

Corollary 3.5. *If $q_\alpha = f_{\alpha,\nu_Q}$, with ν_Q defined in (3.5), and f is continuous on $[0,1]$ with $\inf_{0\le x\le 1} f(x) > \gamma > 0$, then $q_\alpha \longrightarrow_u q$ and*

$$\|\, q_\alpha - q \,\| \le \frac{\Delta(f,\delta)}{\gamma^2} + \frac{2\,\|\, f \,\|}{\alpha\,\delta^2\,\gamma^2} + o\Big(\frac{1}{\alpha}\Big)\,, \quad \textit{as} \;\; \alpha \to \infty.$$

Finally, note that taking $\nu = \nu_Q$ in (3.14), we derive another approximation $q^*_\alpha = f^*_{\alpha,\nu_Q}$ of q based on Beta densities $\beta(\cdot, a, b)$ with the shape parameters $a = [\alpha x] + 1$ and $b = \alpha - [\alpha x] + 1$:

$$q^*_\alpha(x) = \int_0^1 \beta(F(u), [\alpha x] + 1, \alpha - [\alpha x] + 1)\, du\,. \tag{3.16}$$

4. Some Applications and Examples

In this Section the construction of the moment-recovered cdf $F_{\alpha,\nu}$ is applied to the problem of nonparametric estimation of a cdf, its density and a quantile functions as well as to the problem of demixing in exponential, binomial and negative binomial mixtures, and deconvolution in error-in-variable model. In Theorems 4.1 we derive the uniform rate of convergence for the empirical counterpart of $F_{\alpha,\nu}$ denoted by $\widetilde{F}_\alpha$, i.e. for $\widetilde{F}_\alpha = F_{\alpha,\hat{\nu}}$, where $\hat{\nu}$ is the sequence of all empirical moments of the sample from F. In Theorem 4.2 the uniform rate of approximation for moment-recovered quantile function of F is obtained. Finally, the graphs of moment-recovered cdfs, pdfs, and quantile functions are presented in Figures 1-3.

Direct model

Let $X_1, \ldots, X_n$ be a random sample from F defined on $[0,1]$. Denote by $\hat{F}_n$ the empirical cdf (ecdf) of the sample $X_1, \ldots, X_n$:

$$\hat{F}_n(t) = \frac{1}{n}\sum_{i=1}^n I_{[0,t]}(X_i)\,, \quad 0 \le t \le 1\,.$$

Substitution of the empirical moments

$$\hat{\nu}_j = \frac{1}{n}\sum_{i=1}^n X_i^j\,,\; j \in \mathbb{N}\,,$$

instead of $\mu_{j,F}$ into (2.2) yields

$$\widetilde{F}_\alpha(x) = F_{\alpha,\hat{\nu}}(x) = \int_0^1 P_\alpha(t,x) d\,\hat{F}_n(t) = \int_0^1 \mathbf{P}\{S_{[\alpha x]} \ge \alpha t\} d\,\hat{F}_n(t)\,.$$

Furthermore, the empirical analogue of (3.8) admits a similar representation

$$\widetilde{F}_\alpha(x) = \int_0^\infty \hat{F}_n(t)\, d\,\mathbf{P}\{S_{[\alpha x]} \le \alpha t\}.$$

The application of the Theorem 3.3 and the asymptotic properties of $\hat{F}_n$ yield

Theorem 4.1. *If $\nu = \{\mu_{j,F}, j \in \mathbb{N}\}$, then under condition (3.6) we have*

$$\sup_{0\le x\le 1} \left|\widetilde{F}_\alpha(x) - F(x)\right| = O\left(\frac{1}{\sqrt{n}}\right) + O\left(\frac{1}{\alpha}\right) \quad a.s., \quad as \quad \alpha\,, n \to \infty. \tag{4.1}$$

Remark 4.1. In Mnatsakanov and Ruymgaart [17] the weak convergence of the moment-empirical processes $\{\sqrt{n}\{\widetilde{F}_n(t) - F(t)\},\, t \in [0,1]\}$ to the Brownian bridge is obtained.

Of course, when the sample is directly drawn from the cdf F of actual interest, one might use the ecdf $\hat{F}_n$ and empirical process $U_n = \sqrt{n}\,(\hat{F}_n - F)$. The result mentioned in Remark 4.1 yields, that even if the only information available is the empirical moments, we still can construct different test statistics based on the moment-empirical processes $\widetilde{U}_n = \sqrt{n}\,(\widetilde{F}_n - F)$.

On the other hand, using the construction (3.11), one can estimate the density function f given only the estimated or empirical moments in:

$$f_{\alpha,\hat{\nu}}(x) = \frac{\alpha^{[\alpha x]+2}}{\Gamma([\alpha x]+2)} \sum_{m=0}^{\infty} \frac{(-\alpha)^m}{m!}\, \hat{\nu}_{m+[\alpha x]+1}\,,\ x \in [0,1]\,. \tag{4.2}$$

Remark 4.2. In practice, the parameter α as well as the number of summands in (4.2) (and the number of summands in the inner summation of $F_{\alpha,\hat{\nu}}$) can be chosen as the functions of n: $\alpha = \alpha(n) \to \infty$ and $M = M(n) \to \infty$ as $n \to \infty$, that optimize the accuracy of corresponding estimates. Further analysis is required to derive the asymptotic forms of $\alpha(n)$ and $M(n)$ as $n \to \infty$. This question is currently under investigation and is beyond the scope of the present article.

Note that the construction (4.2) yields the estimator $\hat{f}_\alpha(x) = f_{\alpha,\hat{\nu}}$ with $\hat{\nu} = \{\hat{\nu}_j, j \in \mathbb{N}\}$:

$$\hat{f}_\alpha(x) = \frac{\alpha}{n}\sum_{i=1}^{n} \frac{(\alpha X_i)^{[\alpha x]+1}}{([\alpha x]+1)!}\, e^{-\alpha X_i} = \frac{1}{n}\sum_{i=1}^{n} g(X_i, [\alpha x]+2, \alpha),\ \ x \in [0,1]\,.$$

Here $g(\cdot, [\alpha x]+2, \alpha)$ is defined in (3.12). The estimator $\hat{f}_\alpha$ does not represent a traditional kernel density estimator of f. It is defined by a δ-sequence, which consists of the gamma density functions of varying shapes (the shape and the rate parameters are equal to $[\alpha x] + 2$ and α, respectively). It is natural to use this estimator when $supp\{F\} = [0, \infty)$, since, in this case, the supports of f and gamma kernel densities coincide and one avoids the boundary effect of $\hat{f}_\alpha$ (cf. Chen [5]).

Some asymptotic properties such as the convergence in probability of $\hat{f}_\alpha$ uniformly on any bounded interval and the Integrated Mean Squared Error (*IMSE*) of $\hat{f}_\alpha$ have been studied in Mnatsakanov and Ruymgaart [17] and Chen [5], respectively.

Applying the results from Mnatsakanov and Khmaladze [15], where the necessary and sufficient conditions for L_1-consistency of general kernel density estimates are established, one can prove in a similar way (cf. Mnatsakanov [14], Theorem 3) that if f is continuous on $[0,1]$, then $E\,||\,\hat{f}_\alpha - f\,||_{L_1} \to 0$, as $\sqrt{\alpha}/n \to 0$ and $\alpha\,, n \to \infty$.

Exponential mixture model

Let $Y_1, \ldots, Y_n$ be a random sample from the mixture of exponentials

$$G(x) = \int_0^T (1 - e^{-x/\tau})\, dF(\tau)\,,\ x \ge 0\,.$$

The unknown cdf F can be recovered according to the construction $F_{\alpha,\nu} = \mathcal{K}_\alpha^{-1}\,\nu$ with $\nu = \{\mu_{j,G}/j!, j \in \mathbb{N}\}$. Similarly, given the sample $Y_1, \ldots, Y_n$ from G and taking $F_{\alpha,\hat\nu} = \mathcal{K}_\alpha^{-1}\,\hat\nu$, where $\hat\nu = \{\hat\mu_{j,G}/j!, j \in \mathbb{N}\}$, we obtain the estimate of F. Here $\{\hat\mu_{j,G}, j \in \mathbb{N}\}$ are the empirical moments of the sample $Y_1, \ldots, Y_n$. The regularized inversion of the noisy Laplace transform and the L_2-rate of convergence were obtained in Chauveau *et al.* [4].

Binomial and negative binomial mixture models

When $Y_1, \ldots, Y_n$ is a random sample from the binomial or negative binomial mixture distributions, respectively:

$$p(x) := P(Y = x) = \int_0^1 \binom{m}{x} \tau^x\,(1-\tau)^{m-x} dF(\tau)\,,\ \ x = 0, \ldots, m\,,$$

$$p(x) := P(Y = x) = \int_0^1 \frac{\Gamma(r+x)}{\Gamma(r)\,x!} \Big(\frac{1}{1+\tau}\Big)^r \Big(\frac{\tau}{1+\tau}\Big)^x dG(\tau)\,,\ \ x = 0, 1, \ldots\,,$$

where m and r are given positive integers. Assume that the unknown mixing cdfs F and G are such that F has at most $\frac{m+1}{2}$ support points in $(0,1)$, while G is a right continuous cdf on $(0,1)$. In both models the mixing distributions are identifiable (see, for example, Teicher [26] for binomial mixture model). Note also that the jth moments of F and G are related to the jth factorial moments of corresponding Y_i's in the following ways:

$$\mu_{j,F} = \frac{1}{m^{[j]}}\,E(Y_1^{[j]}) \ \text{ and } \ \mu_{j,G} = \frac{1}{r_{(j)}}\,E(Y_1^{[j]})\,.$$

Here $y^{[j]} = y(y-1)\cdots(y-j+1)$ and $r_{(j)} = r(r+1)\cdots(r+j-1)$. To estimate F and G one can use the moment-recovered formulas (2.2) or (3.13) with $\mu_{j,F}$ and $\mu_{j,G}$ defined in previous two equations where the theoretical factorial moments are replaced by corresponding empirical counterparts. The asymptotic properties of the derived estimators of F and G will be studied in a separate work.

Deconvolution problem: error-in-variable model

Consider the random variable $Y = X + U$, with cdf G, where U (the error) has some known symmetric distribution F_2, X has cdf F_1 with a support $[0, T]$, and U and X are independent. This model, known as an error-in-variable model, corresponds to the convolution $G = F_1 \star F_2$. Assuming that all moments of X and U exist, the moments $\{\bar\nu_j, j \in \mathbb{N}\}$ of Y are described by (3.2). Hence, given the moments of U (with $E(U) = 0$), we can recalculate the moments of F_1 as follows: $\mu_{1,F_1} = \bar\nu_1$, $\mu_{2,F_1} = \bar\nu_2 - \mu_{2,F_2}$, and so on. So that, assuming that we already calculated μ_{k,F_1}, or estimated them by μ^*_{k,F_1} for $1 \le k \le j-2$, we will have, for any $j \ge 1$:

$$\mu_{j,F_1} = \bar\nu_j - \sum_{m=2}^{j} \binom{j}{m} \mu_{m,F_2} \times \mu_{j-m,F_1}$$

or, respectively,

$$\mu^*_{j,F_1} = \hat\mu_{j,G} - \sum_{m=2}^{j} \binom{j}{m} \mu_{m,F_2} \times \mu^*_{j-m,F_1}$$

given the sample $Y_1, \ldots, Y_n$ from cdf G. Now the moment-recovered estimate of F_1 will have the form $F_{\alpha,\hat{\nu}} = \mathcal{K}_\alpha^{-1}\,\hat{\nu}$, where $\hat{\nu} = \{\mu^*_{j,F_1}, j \in \mathbb{N}\}$. The alternative construction of the kernel type estimate of F_1 based on the Fourier transforms is studied in Hall and Lahiri [8], where the $\sqrt{n}$-consistency and other properties of the estimated moments $\mu^*_{j,F_1}, j \in \mathbb{N}$, are derived as well.

Example 4.1. Consider the moment sequence $\mu = \{1/(j+1), j \in \mathbb{N}\}$. The corresponding moment-recovered distribution $F_{\alpha,\mu} = \mathcal{K}_\alpha^{-1}\,\mu$ is a good approximation of $F(x) = x$ already with $\alpha = 50$ and $M = 100$.

Assume now that we want to recover the distribution G with corresponding moments $\nu_{j,G} = 1/(j+1)^2, j \in \mathbb{N}$. Since we can represent $\nu_G = \mu \odot \mu$, we conclude from Theorem 1 (i) in Mnatsakanov [13], that $G = F \otimes F$, with $F(x) = x$, and hence $G(x) = x - x\ln x, 0 \le x \le 1$. We plotted the curves of F_{α,ν_G} (the solid line) and G (the dashed line) on Figure 1 (a). We took $\alpha = 50$ and $M = 200$, the number of terms in the inner summation of the formula (2.2). From Figure 1 (a) we can see that the approximation of G by F_{α,ν_G} at $x = 0$ is not as good as inside of the interval $[0, 1]$. This happened because the condition (3.6) from Theorem 3.3 is not valid for $g'(x) = G''(x) = -1/x$.

Example 4.2. To recover the distribution F via moments $\nu_j = 9/(j+3)^2, j \in \mathbb{N}$, note that $\nu_j = \nu_{aj,G}$, with $a = 1/3$. Hence, $F(x) = G(x^3) = x^3 - x^3\ln(x^3)$, $0 \le x \le 1$ (Theorem 1 (iii), Mnatsakanov [13]). We conducted computations of moment-recovered cdf $F_{\alpha,\nu}$ when $\alpha = 50$ and the number of terms in the inner summation of the formula (2.2) is equal to 200. Also, we calculated $F^*_{\alpha,\nu}$ defined in (3.13) with $\alpha = 32$. See Figure 1 (b), where we plotted $F_{\alpha,\nu}$ (the solid blue line), $F^*_{\alpha,\nu}$ (the solid red line), and F (the dashed line), respectively. These two approximations of cdf F justify a good fit already with $\alpha = 50$ and $M = 200$ for the first one and with $\alpha = 32$ for the second one. From Figure 1 (b) we can see that the performance of $F^*_{\alpha,\nu}$ is slightly better compared to $F_{\alpha,\nu}$: $F^*_{\alpha,\nu}$ does not have the "boundary" effect around $x = 1$.

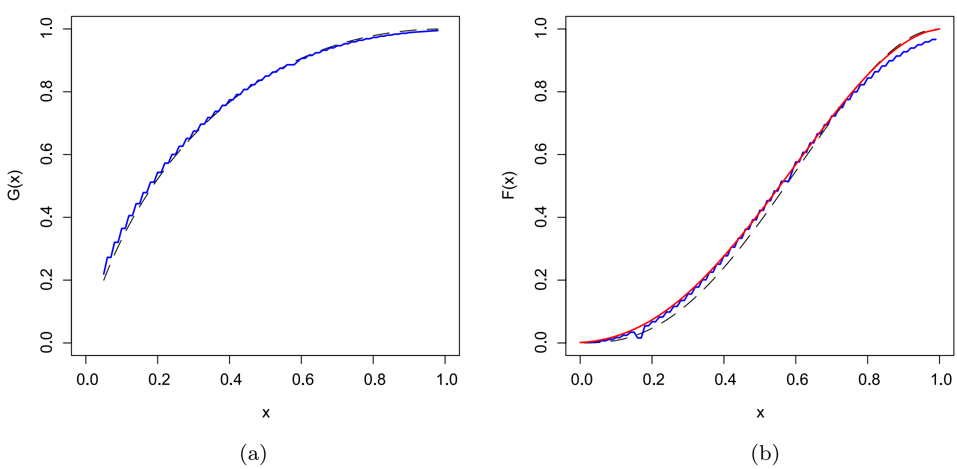

Fig 1. *(a) Approximation of* $G(x) = x - x\ln x$ *by* $F_{\alpha,\nu}$ *and (b) Approximation of* $G(x^3)$ *by* $F_{\alpha,\nu}$ *and by* $F^*_{\alpha,\nu}$.

Estimation of a quantile function Q and quantile density function q

Assume that a random variable X has a continuous cdf F defined on $[0,1]$. To approximate (estimate) the quantile function Q given only the moments (estimated moments) of F, one can use Corollary 3.2. Indeed, after some algebra, we have

$$Q_\alpha(x) = F_{\alpha,\nu_Q}(x) = \int_0^1 P_\alpha\big(F(u), x\big)\, du\,, \quad 0 \le x \le 1\,, \tag{4.3}$$

where ν_Q and $P_\alpha(\cdot,\cdot)$ are defined in (3.5) and in (2.3), respectively. Comparing (4.3) and (3.8) we can prove in a similar way (see, the proof of Theorem 3.3) the following

Theorem 4.2. *If f' is bounded and* $\inf_{0\le x\le 1} f(x) > \gamma > 0$, *then*

$$\sup_{0\le x\le 1} \big|Q_\alpha(x) - Q(x)\big| = O\left(\frac{1}{\alpha}\right), \quad \textit{as } \alpha \to \infty. \tag{4.4}$$

Now, given only the moment sequence ν of F, one can construct the approximation $Q_{\alpha,\beta}$ of Q by substituting the moment-recovered cdf $F_{\beta,\nu}$ (instead of F) in the right hand side of (4.3). Let us denote the corresponding approximation of Q by

$$Q_{\alpha,\beta}(x) = \int_0^1 P_\alpha\big(F_{\beta,\nu}(u), x\big)\, du\,, \quad \alpha,\beta \in \mathbb{N}. \tag{4.5}$$

Figure 2 (a) shows the cdf $F(x) = x^3 - x^3\ln(x^3)$ (the dashed line), introduced in Example 4.2, and its quantile approximation $Q_{\alpha,\beta}$ (the solid line), when $\nu = \{9/(j+3)^2, j \in \mathbb{N}\}$, $\alpha = \beta = 100$, and $M = 200$.

Replacing F by the empirical $\hat{F}_n$ in (4.3), (3.15), and in (3.16) yields the following estimators, respectively, based on the spacings $\Delta X_{(i)} = X_{(i)} - X_{(i-1)}$, $i = 1, \ldots, n+1$:

$$\hat{Q}_\alpha(x) = F_{\alpha,\hat{\nu}_Q}(x) = \int_0^1 P_\alpha\big(\hat{F}_n(u), x\big)\, du = \sum_{i=1}^{n+1} \Delta X_{(i)}\, P_\alpha\left(\frac{i-1}{n}, x\right), \tag{4.6}$$

$$\hat{q}_\alpha(x) = \int_0^1 g(\hat{F}_n(u), [\alpha x]+2, \alpha)\, du = \sum_{i=1}^{n+1} \Delta X_{(i)}\, g\left(\frac{i-1}{n}, [\alpha x]+2, \alpha\right),$$

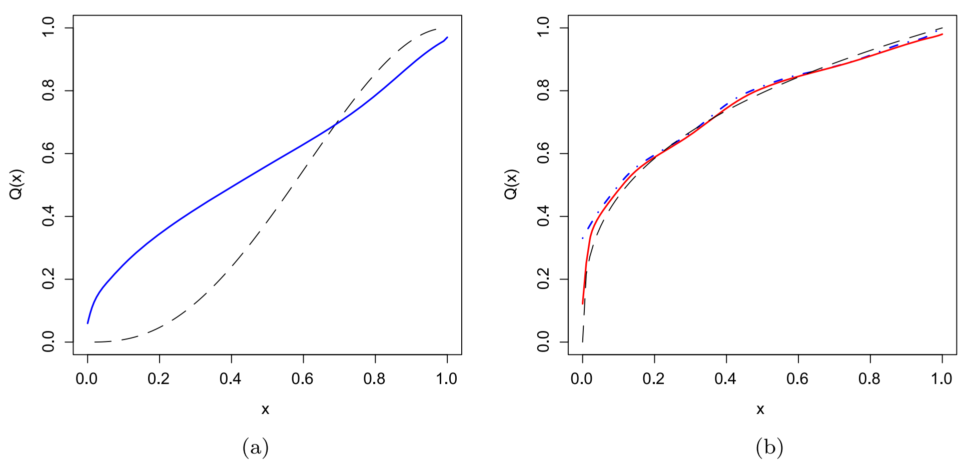

FIG 2. *(a) Approximation of Q by $Q_{\alpha,\beta}$ and (b) Estimation of $Q(x) = x^{1/3}$ by $\hat{Q}_\alpha$ and by $\hat{Q}_{HD}$.*

and

$$\hat{q}^*_\alpha(x) = \sum_{i=1}^{n+1} \Delta X_{(i)}\, \beta\left(\frac{i-1}{n}, [\alpha x]+1, \alpha-[\alpha x]+1\right).$$

Here $\hat{\nu}_Q = \{\int_0^1 [\hat{F}_n(u)]^j du, j \in \mathbb{N}\}$, while $X_{(i)}, i = 1, \ldots, n, X_{(0)} = 0, X_{(n+1)} = 1$, are the order statistics of the sample $X_1, \ldots, X_n$.

Now, let us compare the curves of $\hat{Q}_\alpha$ and the well known Harrell-Davis estimator

$$(4.7) \qquad \hat{Q}_{HD}(x) = \sum_{i=1}^{n} X_{(i)}\, \Delta Beta\left(\frac{i}{n},\ (n+1)x,\ (n+1)(1-x)\right),$$

where $Beta\,(\cdot, a, b)$ denotes the cdf of a $Beta$ distribution with the shape parameters $a > 0$ and $b > 0$. For asymptotic expressions of MSE and the bias term of $\hat{Q}_{HD}$ we refer the reader to Sheather and Marron [19]. Let us generate $n = 100$ independent random variables $X_1, \ldots, X_n$ from $F(x) = x^3, 0 \le x \le 1$. Taking $\alpha = 100$, we estimate (see, Figure 2 (b)) the corresponding quantile function $Q(x) = x^{1/3}, 0 \le x \le 1$, (the dashed line) by means of $\hat{Q}_\alpha$ (the solid line) and by $\hat{Q}_{HD}$ (the dashed-dotted line), defined in (4.6) and (4.7), accordingly. Through simulations we conclude that the asymptotic behavior of the moment-recovered estimator $\hat{Q}_\alpha$ and the Harrell-Davis estimator $\hat{Q}_{HD}$ are similar. The MSE and other properties of $\hat{Q}_\alpha, \hat{q}_\alpha$, and $\hat{q}^*_\alpha$ will be presented in a separate article.

Example 4.1 (continued). Assume now that we want to recover pdf of the distribution G studied in the Example 4.1 via the moments $\nu_{j,G} = 1/(j+1)^2, j \in \mathbb{N}$. On the Figure 3 (a) we plotted the curves of the moment-recovered density $f_{\alpha,\nu}$ (the solid line) defined by (3.11) and $g(x) = G'(x) = -\ln x,\ 0 \le x \le 1$ (the dashed line), respectively. Here we took $\alpha = 50$ and $M = 200$.

Example 4.2 (continued). Now let us recover the pdf $f(x) = -9x^2 \ln x, 0 \le x \le 1$, of distribution F defined in Example 4.2 where $\nu_{j,F} = 9/(j+3)^2, j \in \mathbb{N}$. We applied the approximations $f_{\alpha,\nu}$ and ${f^*}_{\alpha,\nu}$ defined in (3.11) and (3.14), respectively, by calculating the values of $f_{\alpha,\nu}$ and ${f^*}_{\alpha,\nu}$ at the points $x = k/\alpha,\ k = 1, 2, \ldots, \alpha$. Figure 3 (b) shows the curves of $f_{\alpha,\nu}$ (the blue dashed-dotted line), and $f^*_{\alpha,\nu}$ (the

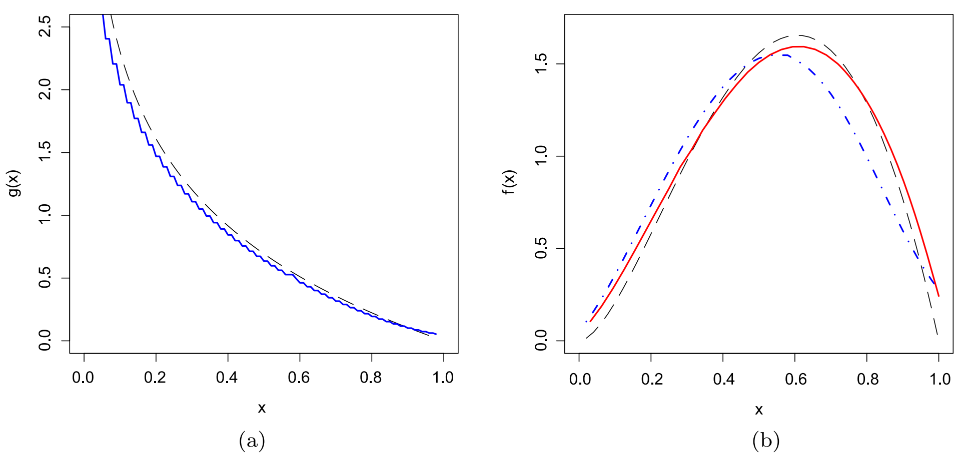

Fig 3. *(a) Approximation of* $g(x) = -\ln x$ *by* $f_{\alpha,\nu}$ *and (b) Approximation of* $f(x) = -9x^2 \ln x$ *by* $f_{\alpha,\nu}$ *and* $f^*_{\alpha,\nu}$.

red solid line), and f (the black dashed line). Here, we took $\alpha = 50$ and $M = 200$ when calculating $f_{\alpha,\nu}$ and $\alpha = 32$ in $f^*_{\alpha,\nu}$. One can see that the performance of $f^*_{\alpha,\nu}$ with $\alpha = 32$ is better than the performance of $f_{\alpha,\nu}$ with $\alpha = 50$ and $M = 200$.

After conducting many calculations of moment-recovered approximants for several models we conclude that the accuracy of the formulas (2.2) and (3.11) are not as good as the ones defined in (3.13) and (3.14) in the Hausdorff case. On the other hand, the constructions (2.2) and (3.11) could be useful in the Sieltjes moment problem as well.

Acknowledgments

The authors are thankful to Estate Khmaladze, E. James Harner, and Cecil Burchfiel for helpful discussions. The work of the first author was supported by the National Institute for Occupational Safety and Health, contract no. 212-2005-M-12857. The findings and conclusions in this paper are those of the authors and do not necessarily represent the views of the National Institute for Occupational Safety and Health.

References

[1] Adell, J. A. and de la Cal, J. (1993). On the uniform convergence of normalized Poisson mixtures to their mixing distribution. *J. Statist. Probab. Lett.* **18** 227–232.

[2] Akhiezer, N. I. (1965). *The Classical Moment Problem and Some Related Questions in Analysis*. Oliver & Boyd, Edinburgh.

[3] Borwein, J. M. and Lewis, A. S. (1993). A survey of convergence results for maximum entropy methods. In *Maximum Entropy and Bayesian Methods* (M. Djafari and G. Demoments, eds.) 39–48. Kluwer Academic Publishers.

[4] Chauveau, D. E., van Rooij, A. C. M. and Ruymgaart, F. H. (1994). Regularized inversion of noisy Laplace transforms. *Adv. in Appl. Math.* **15** 186–201.

[5] Chen, S. X. (2000). Probability density function estimation using gamma kernels. *Ann. Inst. Statist. Math.* **52** 471–480.

[6] Feller, W. (1971). *An Introduction to Probability Theory and Its Applications* **II**. Wiley, New York.

[7] Frontini, M. and Tagliani, A. (1997). Entropy-convergence in Stieltejs and Hamburger moment problem. *Appl. Math. Comput.* **88** 39–51.

[8] Hall, P. and Lahiri, S. N. (2008). Estimation of distributions, moments and quantiles in deconvolution problems. *Ann. Statist.* **36** 2110–2134.

[9] Kevasan, H. K. and Kapur, J. N. (1992). *Entropy Optimization Principles with Applications*. Academic Press, New York.

[10] Lin, G. D. (1992). Characterizations of distributions via moments. *Sankhya Ser. A* **54** 128–132.

[11] Lin, G. D. (1997). On the moment problem. *J. Statist. Probab. Lett.* **35** 85–90.

[12] Lindsay, B. G., Pilla, R. S. and Basak, P. (2000). Moment-based approximations of distributions using mixtures: Theory and applications. *Ann. Inst. Statist. Math.* **52** 215–230.

[13] Mnatsakanov, R. M. (2008). Hausdorff moment problem: Reconstruction of distributions. *J. Statist. Probab. Lett.* **78** 1612–1618.

[14] MNATSAKANOV, R. M. (2008). Hausdorff moment problem: Reconstruction of probability density functions. *J. Statist. Probab. Lett.* **78** 1869–1877.

[15] MNATSAKANOV, R. M. and KHMALADZE, E. V. (1981). On L_1-convergence of statistical kernel estimators of distribution densities. *Soviet Math. Dokl.* **23** 633–636.

[16] MNATSAKANOV, R. M. and KLAASSEN, C. A. J. (2003). Estimation of the mixing distribution in the Poisson mixture models. Uncensored and censored samples. In *Proc. Hawaii International conference on Statistics and Related Fields*, Honolulu, Hawaii, June 4-8, 2003, 1-18, CD: ISSN# 1539–7211.

[17] MNATSAKANOV, R. M. and RUYMGAART, F. H. (2003). Some properties of moment-empirical cdf's with application to some inverse estimation problems. *Math. Meth. Statist.* **12** 478–495.

[18] NOVI INVERARDI, P. L., PETRI, A., PONTUALE, G. and TAGLIANI, A. (2003). Hausdorff moment problem via fractional moments. *Appl. Math. Computat.* **144** 61–74.

[19] SHEATHER, S. J. and MARRON, J. S. (1990). Kernel quantile estimators. *J. Amer. Statist. Assoc.* **85** 410–416.

[20] SHOHAT, J. A. and TAMARKIN, J. D. (1943). *The Problem of Moments.* Amer. Math. Soc., New York.

[21] STIELTJES, T. (1894). Recherches sur les fractions continues. *Anns Sci. Fac. Univ. Toulouse (1894-1895)* **8** J1–J122; A5–A47.

[22] STOYANOV, J. (1997). *Counterexamples in Probability*, 2nd ed. Wiley, Chichester.

[23] STOYANOV, J. (2000). Krein condition in probabilistic moment problems. *Bernoulli* **6** 939–949.

[24] STOYANOV, J. (2004). Stieltjes classes for moment indeterminate probability distributions. *J. Appl. Probab.* **41A** 281–294.

[25] TAGLIANI, A. and VELASQUEZ, Y. (2003). Numerical inversion of the Laplace transform via fractional moments. *Appl. Math. Computat.* **143** 99–107.

[26] TEICHER, H. (1963). Identifiability of finite mixtures. *Ann. Math. Statist.* **34** 1265–1269.

[27] WIDDER, D. V. (1934). The inversion of the Laplace integral and the related moment problem. *Trans. Amer. Math. Soc.* **36** 107–200.