

Hausdorff moment problem: Reconstruction of probability density functions[☆]

Robert M. Mnatsakanov^{*}

Department of Statistics, West Virginia University, Morgantown, WV 26506, USA

Health Effects Laboratory Division, National Institute for Occupational Safety and Health, Morgantown, WV 26505, USA

Received 9 January 2007; received in revised form 27 September 2007; accepted 9 January 2008

Available online 1 February 2008

Abstract

The problem of recovering a moment-determinate probability density function (pdf) from its moments is studied. The proposed construction provides a method for recovery of different pdfs via simple transformations of the moment sequences. Uniform and L_1 -rates of convergence of moment-recovered pdfs are obtained. Finally, some applications and examples are briefly discussed.
© 2008 Elsevier B.V. All rights reserved.

1. Introduction

The problem of non-parametric estimation of indirectly observed probability density functions frequently arises in survival analysis, industry, and biological sciences. In many indirect models, e.g., those based on convolutions, mixtures, multiplicative censoring, and biased sampling, the moments of an unobserved distribution of actual interest can be easily estimated from the transformed moments of an observed distribution. Reconstruction of the domain of a function, or equivalently, the characteristic function of this domain, given its moments, represents another interesting area where the moment problem arises (see, Goldenshluger and Spokoiny (2004)). In all such models one can use the procedure which recovers a function via assigned moments.

Recovering a density function from its moments can be considered as a special case of the classical moment problem (Akhiezer, 1965). There are several formulations of the moment problem. First, let us consider the Hamburger moment problem: for a given sequence $\nu = \{1 = \mu_0, \mu_1, \dots\}$ of real numbers, define a distribution F or its density function f on the real line $(-\infty, \infty)$, such that μ_j represents the j th moment of F , i.e.,

$$\mu_j = \int t^j dF(t) := \mu_{j,F} \quad j = 1, 2, \dots$$

We say that the moment problem has a unique solution if for two distributions F and G , defined on $(-\infty, \infty)$, the equations $\mu_{j,F} = \mu_{j,G}$, $j = 1, 2, \dots$, imply that $F = G$. In the latter case we say that the distribution F is

[☆] This research was supported by the National Institute for Occupational Safety and Health, contract no. 212-2006-M-16936.

^{*} Corresponding address: Department of Statistics, West Virginia University, P.O. Box 6330, Morgantown, WV 26506, USA. Fax: +1 304 293 2272.

E-mail address: rmnatsak@stat.wvu.edu.

moment-determinate (M -determinate), otherwise, F is M -indeterminate. If the support of F , $\text{supp}\{F\} = [0, \infty)$, the problem is known as the Stieltjes moment problem. In this paper we consider the Hausdorff moment problem, where $\text{supp}\{F\} = [0, T]$, $0 < T < \infty$.

There are many articles investigating the conditions (for example, Carleman's and Krein's conditions) under which the distributions are M -determinate or M -indeterminate, see Akhiezer (1965), Lin (1992, 1997), and Stoyanov (2000) among others. Several inversion formulas were obtained by inverting the moment-generating function and Laplace transform (Shohat and Tamarkin, 1943; Feller, 1971; Chauveau et al., 1994; Tagliani and Velasquez, 2003). But there are very few approaches recovering the density functions via their moments. This can be explained by unstable behavior of the approximants when the higher-order moments are involved.

Let us briefly discuss the Maximum Entropy (ME) method (Kevasan and Kapur, 1992). In this approach it is proved that the approximant of f has the form of $f_M(x) = \exp\{-\sum_{j=1}^M \lambda_j x^j\}$, and the maximum value of the entropy of f under the constraints

$$\int t^j f_M(t) dt = \mu_{j,F}, \quad j = 0, 1 \dots M,$$

is equal to the minimum value of the potential function $\Gamma(\lambda_1, \dots, \lambda_M)$:

$$\min_{\lambda_1, \dots, \lambda_M} \Gamma(\lambda_1, \dots, \lambda_M) = \min_{\lambda_1, \dots, \lambda_M} \left\{ \ln \left[\int_0^1 \exp \left\{ - \sum_{j=1}^M \lambda_j x^j \right\} dx \right] + \sum_{j=1}^M \lambda_j \mu_{j,F} \right\}.$$

It is known that the evaluation of coefficients λ_j , $j = 1, \dots, M$, becomes unstable as the number of assigned moments M increases. To reduce the instability in ME approach the fractional moments were recommended to use instead of ordinary ones (see Novi Inverardi et al. (2003) and Frontini and Tagliani (1997) among others). Note also that the uniform, the L_2 , and the entropy convergence of f_M to f , as $M \rightarrow \infty$, have been derived (see, for instance, Borwein and Lewis (1991, 1993), Novi Inverardi et al. (2003) and the references therein), although the rates of convergence were not established.

The aim of this paper is to construct a stable approximant $f_{\alpha,v}$ of f via its moment sequence v , to investigate the asymptotic properties of $f_{\alpha,v}$, and to derive the uniform and L_1 -rates of approximation for $f_{\alpha,v}$, as $\alpha \rightarrow \infty$.

The paper is organized as follows. In Section 2 the construction of moment-recovered function $f_{\alpha,v}$ is given. In Section 3 we study the asymptotic behavior of functions $f_{\alpha,v}$ corresponding to different choices of the moment sequence v and its transformations. In Theorem 1, we derive the uniform and L_1 -rates of convergence for $f_{\alpha,v}$. The question of recovering the convolutions $f_1 \otimes f_2$ and $f_1 \star f_2$ via transformations of corresponding moment sequences $v_1 = \{\mu_{j,F_1}, j \in \mathbb{N}\}$ and $v_2 = \{\mu_{j,F_2}, j \in \mathbb{N}\}$ is addressed in Theorem 2. In Section 4 some applications and examples are discussed. In particular, L_1 -consistency of moment-density estimator \hat{f}_α derived from (5) is proved.

2. Notations and preliminaries

Suppose that the cumulative distribution function (cdf) F is absolutely continuous with respect to the Lebesgue measure and has a finite support, say, $\text{supp}\{F\} = [0, 1]$. Denote the pdf and the moments of F by f and

$$\mu_{j,F} = \int_0^1 t^j dF(t) = (\mathcal{K}F)(j), \quad j \in \mathbb{N} = \{0, 1, \dots\}, \quad \mu_{0,F} = 1, \quad (1)$$

respectively. Assume that the moment sequence $v = (\mu_{0,F}, \mu_{1,F}, \dots)$ determines F uniquely, i.e., F is M -determinate.

Our construction in (5) is based on the inverse of the operator \mathcal{K} from (1). It is given by

$$(\mathcal{K}_\alpha^{-1}v)(x) = \sum_{k=0}^{[\alpha x]} \sum_{j=k}^{\alpha} \binom{\alpha}{j} \binom{j}{k} (-1)^{j-k} \mu_{j,F}, \quad 0 \leq x \leq 1, \quad \alpha \in \mathbb{N}, \quad (2)$$

and is such that $\mathcal{K}_\alpha^{-1} \mathcal{K}F \rightarrow_w F$, as $\alpha \rightarrow \infty$ (see, Mnatsakanov and Ruymgaart (2003)). Here \rightarrow_w denotes the weak convergence of cdfs, i.e., convergence at each continuity point of the limiting cdf F . In what follows we assume,

for simplicity, that the parameter α is a positive integer. The following notations will be useful as well: denote the supremum norm of a function $\phi : [0, 1] \rightarrow \mathbb{R}$ by

$$\|\phi\| = \sup_{0 \leq x \leq 1} |\phi(x)|,$$

and the L_1 -norm by $\|\phi\|_{L_1} = \int |\phi(x)| dx$, respectively. For any moment sequence $v = \{v_j, j \in \mathbb{N}\}$, let $F_{\alpha,v} = \mathcal{K}_{\alpha}^{-1} v$ be the moment-approximant function of F :

$$F_{\alpha,v} \rightarrow_w F, \quad \text{as } \alpha \rightarrow \infty. \quad (3)$$

When F is a smooth function, we can derive the uniform and L_1 -rate of approximations in (3) (see, Mnatsakanov (2008)):

Theorem. Let $v = \{\mu_{j,F}, j \in \mathbb{N}\}$. If f' is bounded on $[0, 1]$ and $\alpha \rightarrow \infty$, then

$$\begin{aligned} \|F_{\alpha,v} - F\| &\leq \frac{1}{\alpha + 1} \left\{ \|f\| + \frac{1}{2} \|f'\| \right\} + o\left(\frac{1}{\alpha}\right), \\ \|F_{\alpha,v} - F\|_{L_1} &\leq \frac{1}{\alpha + 1} \left\{ 1 + \frac{1}{12} \|f'\| \right\} + o\left(\frac{1}{\alpha}\right). \end{aligned}$$

To approximate pdf f let us consider the ratio:

$$\frac{\Delta F_{\alpha,v}(x)}{\Delta}, \quad \text{as } \Delta \rightarrow 0, \alpha \rightarrow \infty. \quad (4)$$

Here $\Delta F_{\alpha,v}(x) = F_{\alpha,v}(x) - F_{\alpha,v}(x - \Delta)$ and $\Delta = 1/\alpha$. Scaling the left-hand side of (4) by $(\alpha + 1)/\alpha$ and denoting it by $f_{\alpha,v}(x)$, we obtain the very simple form:

$$f_{\alpha,v}(x) = \frac{\Gamma(\alpha + 2)}{\Gamma([\alpha x] + 1)} \sum_{m=0}^{\alpha - [\alpha x]} \frac{(-1)^m \mu_{m+[\alpha x],F}}{m! (\alpha - [\alpha x] - m)!}, \quad x \in [0, 1]. \quad (5)$$

If f is defined on $[0, T]$, $0 < T < \infty$, and not on $[0, 1]$, the previous display takes the form

$$f_{\alpha,v}(x) = \frac{\Gamma(\alpha + 2)}{\Gamma([\alpha x/T] + 1)} \cdot \frac{1}{T^{[\alpha x/T] + 1}} \sum_{m=0}^{\alpha - [\alpha x/T]} \frac{(-1/T)^m \mu_{m+[\alpha x/T],F}}{m! (\alpha - [\alpha x/T] - m)!}, \quad x \in [0, T], \quad (6)$$

(cf. Mnatsakanov (2008), formula (2)). In Section 3 the uniform convergence of $f_{\alpha,v}$ to f_v , as $\alpha \rightarrow \infty$, is denoted by

$$f_{\alpha,v} \longrightarrow_u f_v, \quad (7)$$

and the modulus of continuity of f as $\Delta(f, \delta) = \sup_{|t-s| \leq \delta} |f(t) - f(s)|$, $0 < \delta < 1$.

3. Some properties of moment-recovered pdfs

In this Section we study some asymptotic properties of the approximants $f_{\alpha,v}$ in (5) and derive the uniform and L_1 -rates of convergence for $f_{\alpha,v}$, as $\alpha \rightarrow \infty$.

Theorem 1. Let $v = \{\mu_{j,F}, j \in \mathbb{N}\}$.

(i) If pdf f is continuous on $[0, 1]$, then $f_{\alpha,v} \longrightarrow_u f$ and for some $0 < \delta < 1$,

$$\|f_{\alpha,v} - f\| \leq \Delta(f, \delta) + \frac{2\|f\|}{\delta^2 (\alpha + 2)};$$

(ii) If f'' is continuous on $[0, 1]$, then

$$f_{\alpha,v}(x) - f(x) = \frac{1}{\alpha + 2} \left[f'(x) (1 - 2x) + \frac{1}{2} f''(x) x (1 - x) \right] + o\left(\frac{1}{\alpha}\right), \quad \text{as } \alpha \rightarrow \infty; \quad (8)$$

and

$$\|f_{\alpha,v} - f\|_{L_1} \leq \frac{1}{\alpha+2} \int_0^1 |f'(x)(1-2x) + \frac{1}{2}f''(x)x(1-x)| dx + o\left(\frac{1}{\alpha}\right), \quad \text{as } \alpha \rightarrow \infty; \quad (9)$$

(iii) If f'' is bounded on $[0, 1]$, then

$$\|f_{\alpha,v} - f\| \leq \frac{1}{\alpha+2} \left\{ 2\|f'\| + \frac{1}{2}\|f''\| + \frac{2}{\alpha+1}\|f''\| \right\}. \quad (10)$$

Proof. (i): Let us denote by

$$\beta(u, a, b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} u^{a-1} (1-u)^{b-1}, \quad 0 < u < 1, \quad (11)$$

a pdf of Beta (a, b) distribution with the shape parameters a and b . Assume here that $a = [\alpha x] + 1$ and $b = \alpha - [\alpha x] + 1$, respectively. Changing the order of summation and integration on the right-hand side of (5), we obtain

$$\begin{aligned} f_{\alpha,v}(x) &= \frac{\Gamma(\alpha+2)}{\Gamma([\alpha x]+1)} \cdot \sum_{m=0}^{\alpha-[\alpha x]} \frac{(-1)^m \mu_{m+[\alpha x], F}}{m! (\alpha - [\alpha x] - m)!} \\ &= \frac{\Gamma(\alpha+2)}{\Gamma([\alpha x]+1)} \cdot \int_0^1 t^{[\alpha x]} \sum_{m=0}^{\alpha-[\alpha x]} \frac{(-t)^m}{m! (\alpha - [\alpha x] - m)!} f(t) dt \\ &= \int_0^1 \beta(t, [\alpha x] + 1, \alpha - [\alpha x] + 1) f(t) dt. \end{aligned} \quad (12)$$

Now one can derive $f_{\alpha,v} \rightarrow_u f$ from the last line in (12) (see, Bouezmarni and Rolin (2003), or Feller (1971), Ch. VII), since $\{\beta(\cdot, [\alpha x] + 1, \alpha - [\alpha x] + 1), \alpha \in \mathbb{N}\}$ forms a δ -sequence at x , as $\alpha \rightarrow \infty$, its mean $\eta_\alpha = ([\alpha x] + 1)/(\alpha + 2)$ and variance

$$\sigma_\alpha^2 = \frac{([\alpha x] + 1)(\alpha - [\alpha x] + 1)}{(\alpha + 2)^2 (\alpha + 3)} \leq \frac{1}{\alpha + 2}, \quad (13)$$

are such that

$$\eta_\alpha - x = \frac{1-2x}{\alpha+2} + \Delta_1(x) \quad \text{and} \quad \sigma_\alpha^2 = \frac{1}{\alpha+2} x(1-x) + \Delta_2(x), \quad (14)$$

with $\Delta_k(x) \leq 2(\alpha+2)^{-2}$, $k = 1, 2$ (see, Johnson et al. (1994) and Chen (1999)).

To derive the inequality in (i), let us split the integration on the right-hand side of (12) into two parts. We obtain

$$\sup_{0 \leq x \leq 1} \left(\int_{|t-x| \leq \delta} + \int_{|t-x| > \delta} \right) |f(t) - f(x)| \beta(t, [\alpha x] + 1, \alpha - [\alpha x] + 1) dt \leq \Delta(f, \delta) + R_\alpha,$$

where $0 < \delta < 1$. Furthermore, since the beta distribution $\beta(\cdot, [\alpha x] + 1, \alpha - [\alpha x] + 1)$ concentrates its masses around the point x , we can estimate its tail using (13) and Tchebishev's inequality:

$$R_\alpha \leq 2\|f\| \sup_{0 \leq x \leq 1} \int_{|t-x| > \delta} \beta(t, [\alpha x] + 1, \alpha - [\alpha x] + 1) dt \leq \frac{2\|f\|}{\delta^2(\alpha+2)}.$$

Combining the last two inequalities we complete the proof of (i).

To prove (ii), we will use the Taylor expansion of f at x . Application of the continuity property of f'' and (14) leads to (8) (cf. Chen (1999)). In (8) the term $o\left(\frac{1}{\alpha}\right)$ is uniform in x . Therefore, (9) follows from (8).

To prove (iii) let us note that

$$f(t) - f(x) = f'(x)(t-x) + \int_x^t d\tau \int_x^\tau f''(y) dy, \quad (15)$$

and

$$|\eta_\alpha - x| \leq \frac{2}{\alpha + 2}. \quad (16)$$

Thus the combination of (13), (15) and (16) yields (10), since

$$\begin{aligned} \|f_{\alpha,v} - f\| &\leq \sup_{0 \leq x \leq 1} \left| f'(x) \int_0^1 (t-x) \beta(t, a, b) \, du \right| + \sup_{0 \leq x \leq 1} \left| \int_0^1 \beta(t, a, b) \, dt \left[\int_x^t d\tau \int_x^\tau f''(y) \, dy \right] \right| \\ &\leq \|f'\| \sup_{0 \leq x \leq 1} |\eta_\alpha - x| + \|f''\| \sup_{0 \leq x \leq 1} \int_0^1 \frac{1}{2} (t-x)^2 \beta(t, a, b) \, dt \\ &\leq \frac{2}{\alpha + 2} \|f'\| + \frac{1}{2(\alpha + 2)} \|f''\| + \frac{2}{(\alpha + 2)^2} \|f''\|. \quad \square \end{aligned}$$

For any two M -determinate pdfs f_1 and f_2 with moment sequences $\nu_1 = \{\mu_{j,F_1}, j \in \mathbb{N}\}$ and $\nu_2 = \{\mu_{j,F_2}, j \in \mathbb{N}\}$, respectively, let us use the following notations for the convolutions

$$\begin{aligned} f_1 \otimes f_2(x) &= \int f_1(x/\tau) f_2(\tau) \frac{1}{\tau} \, d\tau, \quad 0 \leq x \leq 1, \\ f_1 \star f_2(x) &= \int f_1(x - \tau) f_2(\tau) \, d\tau, \quad 0 \leq x \leq 2. \end{aligned}$$

The following notations will be useful as well: define $\nu_1 \odot \nu_2 = \{\mu_{j,F_1} \times \mu_{j,F_2}, j \in \mathbb{N}\}$ and $\nu_1 \oplus \nu_2 = \{\bar{\nu}_j, j \in \mathbb{N}\}$, with

$$\bar{\nu}_j = \sum_{m=0}^j \binom{j}{m} \mu_{m,F_1} \times \mu_{j-m,F_2}. \quad (17)$$

One can easily verify that the moment sequences of $f_1 \otimes f_2$ and $f_1 \star f_2$ are $\nu_1 \odot \nu_2$ and $\nu_1 \oplus \nu_2$, respectively. Also, the reverse statements are true, see (i) and (ii) in Theorem 2.

Let us consider the following conditions:

$$\int_0^1 f_k(t) \frac{1}{t} \, dt < \infty, \quad \text{for } k = 1, 2. \quad (18)$$

Theorem 2. Assume that the densities $f, f_k, k = 1, 2$, are continuous on $[0, 1]$.

- (i) If $\nu = \nu_1 \odot \nu_2$ and the conditions (18) are satisfied, then (7) holds with $f_\nu = f_1 \otimes f_2$;
- (ii) If $\nu = \nu_1 \oplus \nu_2$, then (7) holds with $f_\nu = f_1 \star f_2$;
- (iii) If $\nu = \{\mu_j, j \in \mathbb{N}\}$ with

$$\mu_j = \int [\phi(x)]^j \, dF(x), \quad (19)$$

for some continuous and increasing function $\phi : [0, 1] \rightarrow [0, 1]$, then (7) holds with $f_\nu(x) = f(\phi^{-1}(x)) [\phi'(\phi^{-1}(x))]^{-1}$.

- (iv) If $\nu = \{\bar{\mu}_j, j \in \mathbb{N}\}$ with $\bar{\mu}_j = \beta_1 \mu_{j,F_1} + \beta_2 \mu_{j,F_2}$, $\beta_1, \beta_2 > 0$, then $f_{\alpha,v} \rightarrow_u \beta_1 f_1 + \beta_2 f_2$.

Proof. (i): Assume that $\nu = \nu_1 \odot \nu_2$. By substitution of the product $\mu_{m+[\alpha x], F_1} \cdot \mu_{m+[\alpha x], F_2}$ instead of $\mu_{m+[\alpha x], F}$ on the right-hand side of (5) we obtain

$$\begin{aligned}
f_{\alpha, \nu}(x) &= \frac{\Gamma(\alpha + 2)}{\Gamma([\alpha x] + 1)} \sum_{m=0}^{\alpha - [\alpha x]} \frac{(-1)^m \mu_{m+[\alpha x], F_1} \cdot \mu_{m+[\alpha x], F_2}}{m! (\alpha - [\alpha x] - m)!} \\
&= \frac{\Gamma(\alpha + 2)}{\Gamma([\alpha x] + 1)} \int_0^1 \int_0^1 (t u)^{[\alpha x]} \sum_{m=0}^{\alpha - [\alpha x]} \frac{(-t u)^m}{m! (\alpha - [\alpha x] - m)!} f_1(u) f_2(t) du dt \\
&= \int_0^1 \left[\int_0^1 \beta(\tau, [\alpha x] + 1, \alpha - [\alpha x] + 1) f_1\left(\frac{\tau}{t}\right) d\tau \right] f_2(t) \frac{1}{t} dt \longrightarrow f_1 \otimes f_2(x)
\end{aligned} \tag{20}$$

uniformly on $[0, 1]$, as $\alpha \rightarrow \infty$. In the last line of (20) the Lebesgue dominant convergence theorem is used: by taking into account (14) and the conditions of Theorem 2, we have for each fixed t

$$\phi_\alpha(x, t) := \int_0^1 \beta(\tau, [\alpha x] + 1, \alpha - [\alpha x] + 1) f_1\left(\frac{\tau}{t}\right) d\tau \longrightarrow f_1\left(\frac{x}{t}\right) \quad \text{uniformly on } [0, 1], \text{ as } \alpha \rightarrow \infty,$$

(see, Feller (1971), Ch. VII); on the other hand, the sequence of functions $\{\phi_\alpha(\cdot, \cdot), \alpha \in \mathbb{N}\}$ is uniformly bounded on $[0, 1]^2$.

Now, in a similar way, as we did in (20), we can derive (ii) and (iii). Let us prove only (ii). Replacing the moments μ_{m, F_1} and μ_{j-m, F_2} in (17) by corresponding integrals taken with respect to dF_1 and dF_2 we obtain

$$\bar{v}_j = \int \int (t + u)^j f_1(u) f_2(t) du dt.$$

Since $f_1 \star f_2$ is defined on $[0, 2]$, the substitution of \bar{v}_j into (6), where $T = 2$, yields

$$\begin{aligned}
f_{\alpha, \nu}(x) &= \frac{\Gamma(\alpha + 2)}{\Gamma([\alpha \frac{x}{2}] + 1)} \int_0^1 \int_0^1 \left(\frac{t + u}{2}\right)^{[\alpha \frac{x}{2}]} \sum_{m=0}^{\alpha - [\alpha \frac{x}{2}]} \frac{\left(-\frac{t+u}{2}\right)^m}{m! (\alpha - [\alpha \frac{x}{2}] - m)!} f_1(u) f_2(t) \frac{1}{2} du dt \\
&= \int_0^1 \left[\int_0^1 \beta\left(\tau, \left[\alpha \frac{x}{2}\right] + 1, \alpha - \left[\alpha \frac{x}{2}\right] + 1\right) f_1(2\tau - t) d\tau \right] f_2(t) dt \longrightarrow_u f_1 \star f_2(x),
\end{aligned}$$

as $\alpha \rightarrow \infty$.

The proof of statement (iv) follows from Theorem 1 (i) and from the linearity of $f_{\alpha, \nu}$ with respect to ν . \square

Corollary 1. (i) If f is continuous and for some $a > 0$ and $b \geq 0$, $\nu = \{\mu_{aj+b, F}, j \in \mathbb{N}\}$, then (7) holds with

$$f_\nu(x) = \frac{1}{a \mu_{b, F}} f\left(x^{\frac{1}{a}}\right) x^{\frac{b+1}{a}-1}; \tag{21}$$

(ii) If f is continuous and $\nu = \{a^j \mu_{j, F}, j \in \mathbb{N}\}$ for some positive a , then (7) holds with $f_\nu(x) = f(x/a)/a$, $x \in [0, a]$.

(iii) If f is continuous, $f_{\alpha, \nu}^*(x) = f_{\alpha, \nu}(\phi(x)) \phi'(x)$, and $\nu = \{\mu_j, j \in \mathbb{N}\}$ with μ_j defined according to (19), then $f_{\alpha, \nu}^* \longrightarrow_u f$;

(iv) If f is continuous, $f_{\alpha, \nu}^*(x) = f_{\alpha, \nu}(x^a) a x^{a-1}$, and $\nu = \{\mu_{aj, F}, j \in \mathbb{N}\}$ for some $a > 0$, then $f_{\alpha, \nu}^* \longrightarrow_u f$.

Proof of Corollary 1. The statement (i) with $b = 0$ is a special case of Theorem 2 (iii), where $\phi(x) = t^a$. When $b \neq 0$, the proof of (i) is reduced to the case with $b = 0$ by replacing pdf f with $f_b(x) = x^b f(x)/\mu_{b, F}$. The case (ii) can be proved using (6), with $T = a$, and with the argument similar to the one used in the proof of Theorem 1(i): in last line of (12) we have an extra factor $1/a$, and x/a instead of x . The transformation $x \rightarrow \phi(x)$ combined with the statement of Theorem 2(iii) yields Corollary 1(iii), while (iv) is a special case of (iii), where $\phi(x) = x^a$. \square

Remark 1. The construction $f_{\alpha, \nu}^*$ from Corollary 1 (iv) can be very helpful when the pdf f has only a finite number of moments up to, say, $m^* < \infty$. In this case one can use the fractional moments $\mu_j^* = \mu_{aj, F}$, $j = 0, 1, \dots, \alpha$, with $a = m^*/\alpha$, and reconstruct f by means of $f_{\alpha, \nu}^*$ where $\nu = \{\mu_j^*, j = 0, 1, \dots, \alpha\}$ (cf., for example, Novi Inverardi et al. (2003)).

4. Some applications and examples

The constructions (2) and (5) can be useful in the estimation problem of a cdf F or its density function f in indirect models (cf. Mnatsakanov and Ruymgaart (2003, 2005)). Apart from being an alternative to the traditional estimation technique in a direct model, this approach is applicable to situations where we do not have the observations from a target distribution. For example, assume that we have n copies Y_1, \dots, Y_n from a cdf G which is related in some specific way to the target distribution F with $\text{supp}\{F\} = [0, 1]$. Suppose that we can estimate the moments of F from the given data. Denote these estimated moments by $\hat{v} = \{\hat{\mu}_{j,F}, j \in \mathbb{N}\}$. Estimation of f may now be carried out by replacing the moments $\mu_{j,F}$ by their estimators $\hat{\mu}_{j,F}$ in (5). This yields a moment-density estimator (mde) of f :

$$f_{\alpha, \hat{v}}(x) = \frac{\Gamma(\alpha + 2)}{\Gamma([\alpha x] + 1)} \sum_{m=0}^{\alpha - [\alpha x]} \frac{(-1)^m \hat{\mu}_{m+[\alpha x], F}}{m! (\alpha - [\alpha x] - m)!}, \quad x \in [0, 1], \quad (22)$$

which can be helpful when only estimated moments of the target distribution F are available. These types of situations occur in many indirect models like length-biased, multiplicative censoring, and mixture models where the moments of an unobserved distribution F can be consistently estimated from the moments of an observed distribution G . Let us briefly discuss the direct case (when $G = F$) and the length-biased model.

Direct model. Suppose that the observations X_1, \dots, X_n from the target cdf F are available. Consider estimator (22) with the empirical moments:

$$\hat{\mu}_{j,F} = \int_0^1 t^j d\hat{F}_n(t), \quad j \in \mathbb{N}, \quad \hat{F}_n(t) = \frac{1}{n} \sum_{i=1}^n \{X_i \leq t\}.$$

This estimator looks extremely unstable, because the higher-order empirical moments are unstable. But, somewhat surprisingly, substitution of the empirical moments $\hat{\mu}_{j,F}$ into (22) justifies that the estimator $f_{\alpha, \hat{v}}$ is actually stable:

$$\hat{f}_\alpha(x) = f_{\alpha, \hat{v}}(x) = \frac{1}{n} \sum_{i=1}^n \beta(X_i, [\alpha x] + 1, \alpha - [\alpha x] + 1), \quad x \in [0, 1]. \quad (23)$$

Here $\beta(\cdot, a, b)$ is the beta pdf defined in (11). The estimator \hat{f}_α does not represent a traditional kernel density estimator of f : although the sequence of beta kernels in (23) represents a δ -sequence, the shape of each kernel varies according to the position of a point x , where f is estimated. Some asymptotic properties of \hat{f}_α such as the Integrated Mean Squared Error, uniform convergence in probability, and convergence in L_1 -norm were studied in Chen (1999) and Bouezmarni and Rolin (2003). Note that \hat{f}_α is free from the edge effects as well. By checking the necessary and sufficient condition for L_1 -consistency of general density estimators (see, Mnatsakanov and Khmaladze (1981)), we derive very easily the L_1 -consistency of mde \hat{f}_α :

Theorem 3. *If f is continuous on $[0, 1]$, then*

$$E\|\hat{f}_\alpha - f\|_{L_1} \rightarrow 0, \quad \text{as } \frac{\sqrt{\alpha}}{n} \rightarrow 0 \text{ and } \alpha, n \rightarrow \infty. \quad (24)$$

Proof. From (12) and (23) we conclude $E\hat{f}_\alpha = f_{\alpha, v}$, while according to Theorem 1(i), we have

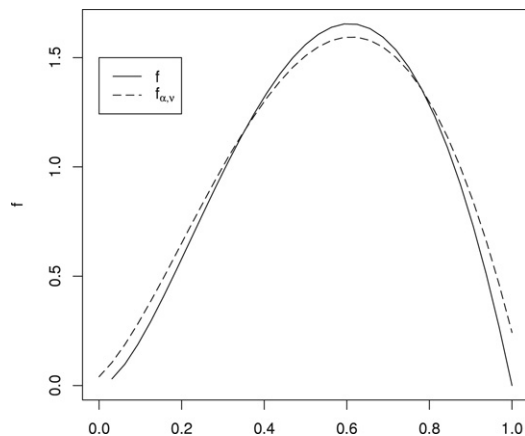
$$\|E\hat{f}_\alpha - f\|_{L_1} \leq \|f_{\alpha, v} - f\| \rightarrow 0, \quad \text{as } \alpha \rightarrow \infty. \quad (25)$$

Now, to prove (24), i.e., $E\|\hat{f}_\alpha - E\hat{f}_\alpha\|_{L_1} \rightarrow 0$, it is sufficient to show that

$$F\{A_{\alpha, n}(\delta)\} = F\left\{x : \int_0^1 \beta^2(\tau, [\alpha x] + 1, \alpha - [\alpha x] + 1) f(\tau) d\tau \geq n\delta\right\} \rightarrow 0,$$

for any $\delta > 0$ and $\alpha, n \rightarrow \infty$ (see, Theorem 1 in Mnatsakanov and Khmaladze (1981)). But

$$\beta(\tau, [\alpha x] + 1, \alpha - [\alpha x] + 1) \leq \frac{c_1 \sqrt{\alpha}}{\sqrt{x(1-x)}}, \quad 0 < x < 1, \quad (26)$$

Fig. 1. Approx. of f by $f_{\alpha, \nu}$.

is valid for some positive constant c_1 (see, for example, [Chen \(2000\)](#)). Hence, application of (25) and (26) yields

$$\begin{aligned} F\{A_{\alpha, n}(\delta)\} &\leq \frac{1}{n\delta} \int_{A_{\alpha, n}(\delta)} f(x) dx \int_0^1 \beta^2(\tau, [\alpha x] + 1, \alpha - [\alpha x] + 1) f(\tau) d\tau \\ &\leq \frac{c_1 \sqrt{\alpha}}{n\delta} \|f_{\alpha, \nu} - f\| \int_0^1 \frac{f(x)}{\sqrt{x(1-x)}} dx + \frac{c_1 \sqrt{\alpha}}{n\delta} \int_0^1 \frac{f^2(x)}{\sqrt{x(1-x)}} dx \rightarrow 0, \end{aligned} \quad (27)$$

as $\sqrt{\alpha}/n \rightarrow 0$, $\alpha, n \rightarrow \infty$. \square

Remark 2. The condition $\sqrt{\alpha}/n \rightarrow 0$ in (24) is equivalent to the condition $nh \rightarrow \infty$, $h \rightarrow 0$, when $\alpha = h^{-2}$. The latter condition represents one of the necessary and sufficient conditions for L_1 -consistency in the traditional kernel density estimation (see, for example, [Devroye \(1984\)](#)).

Length-biased model. In this case we observe a random variable Y with a cdf

$$G(y) = W^{-1} \int_0^y x dF(x), \quad y \in [0, 1],$$

where $W = \int x dF(x)$. It is immediate that the moments $\mu_{j, F}$ of F are related to the moments of G according to $\mu_{j, F} = W \mu_{j-1, G}$. The corresponding mde $f_{\alpha, \hat{\nu}}$ in (22), where $\hat{\nu} = \{\hat{W} \hat{\mu}_{j-1, G}, j \geq 1\}$, will have the form

$$\hat{f}_{\alpha, \hat{\nu}}(x) = \frac{\hat{W}}{n} \sum_{i=1}^n \frac{1}{Y_i} \beta(Y_i, [\alpha x] + 1, \alpha - [\alpha x] + 1), \quad x \in [0, 1], \quad (28)$$

with

$$\hat{W} = \left(\int_0^1 t^{-1} d\hat{G}_n(t) \right)^{-1} \quad \text{and} \quad \hat{G}_n(t) = \frac{1}{n} \sum_{i=1}^n \{Y_i \leq t\}.$$

Applying a similar argument used in the proof of [Theorem 3](#), one can derive the L_1 -consistency of mde $\hat{f}_{\alpha, \hat{\nu}}$ as well. We investigate this model in detail in a later paper.

Example 1. Assume that we want to recover the pdf g of a distribution G whose moments are $\mu_{j, G} = 3/(j+3)$, $j \in \mathbb{N}$. Since $\mu_{j, G} = \mu_{aj, F}$, with the moment sequence $\mu_{j, F} = 1/(j+1)$, $j \in \mathbb{N}$, corresponding to the uniform pdf f on $[0, 1]$, we conclude from (21), where $a = 1/3$ and $b = 0$, that $g(x) = 3x^2$, $0 \leq x \leq 1$. The approximation of $g(x)$ by $f_{\alpha, \nu}(x)$ is very good already for $\alpha = 32$ and $x = k/\alpha$, $k = 0, 1, \dots, \alpha$.

Example 2. Let us recover the distribution F via moments $\nu = \{\mu_{j,F} = 9/(j+3)^2, j \in \mathbb{N}\}$. In this case $\mu_{j,F} = \mu_{j,G} \cdot \mu_{j,G}, j \in \mathbb{N}$, so that the pdf f of F has the form (see, [Theorem 2\(i\)](#)):

$$f(x) = \int g\left(\frac{x}{\tau}\right) g(\tau) \frac{1}{\tau} d\tau = -9x^2 \log x, \quad 0 \leq x \leq 1.$$

We conducted the computations of the moment-recovered pdf $f_{\alpha,\nu}(x)$ at $x = k/\alpha$, $k = 0, 1, \dots, \alpha$, with $\alpha = 32$. See [Fig. 1](#), where we plotted the pdf f (the solid curve), and its moment-recovered pdf $f_{\alpha,\nu}$ (the dashed curve), respectively. We can see that the approximation is good.

Acknowledgments

The author is thankful to Estate Khmaladze, Vladimir Hnizdo, and Cecil Burchfiel for helpful discussions and to Artak Hakobyan for conducting calculations in [Examples 1](#) and [2](#). The author is also grateful to the anonymous referee for many helpful comments and suggestions. 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

- Akhiezer, N.I., 1965. The Classical Moment Problem and Some Related Questions in Analysis. Oliver and Boyd, Edinburgh.
- Borwein, J.M., Lewis, A.S., 1991. Convergence of best entropy estimates. *SIAM J. Optim.* 1, 191–205.
- Borwein, J.M., Lewis, A.S., 1993. A survey of convergence results for maximum entropy methods. In: Djafari, M., Demoments, G. (Eds.), *Maximum Entropy and Bayesian Methods*. Kluwer Academic Publishers, pp. 39–48.
- Bouezmarni, T., Rolin, J.-M., 2003. Consistency of the beta kernel density function estimator. *Canad. J. Statist.* 31, 89–98.
- Chauveau, D.E., van Rooij, A.C.M., Ruymgaart, F.H., 1994. Regularized inversion of noisy Laplace transforms. *Adv. Appl. Math.* 15, 186–201.
- Chen, S.X., 1999. Beta kernel estimators for density functions. *Comput. Statist. Data Anal.* 31, 131–145.
- Chen, S.X., 2000. Beta kernel for regression curves. *Statist. Sinica* 10, 73–92.
- Devroye, L., 1984. *Nonparametric Density Estimation: The L_1 View*. John Wiley and Sons, New York.
- Feller, W., 1971. *An Introduction to Probability Theory and its Applications*, vol. II. Wiley, New York.
- Frontini, M., Tagliani, A., 1997. Entropy-convergence in Stieltjes and Hamburger moment problem. *Appl. Math. Comput.* 88, 39–51.
- Goldenshluger, A., Spokoiny, V., 2004. On the shape-from-moments problem and recovering edges from noisy Radon data. *Probab. Theory Related Fields* 128, 123–140.
- Johnson, N.L., Kotz, S., Balakrishnan, N., 1994. *Continuous Univariate Distributions*. Wiley, New York.
- Kevasan, H.K., Kapur, J.N., 1992. *Entropy Optimization Principles with Applications*. Academic Press, New York.
- Lin, G.D., 1992. Characterizations of distributions via moments. *Sankhya* 54 (Series A), 128–132.
- Lin, G.D., 1997. On the moment problem. *Statist. Probab. Lett.* 35, 85–90.
- Mnatsakanov, R.M., 2008. Hausdorff moment problem: reconstruction of distributions. *J. Statist. Probab. Lett.* doi:10.1016/j.spl.2008.01.011.
- Mnatsakanov, R.M., Khmaladze, E.V., 1981. On L_1 -convergence of statistical kernel estimators of distribution densities. *Soviet Math. Dokl.* 23, 633–636.
- Mnatsakanov, R., Ruymgaart, F.H., 2003. Some properties of moment-empirical cdf's with application to some inverse estimation problems. *Math. Methods Statist.* 12, 478–495.
- Mnatsakanov, R., Ruymgaart, F.H., 2005. Some results for moment-empirical cumulative distribution functions. *J. Nonparametr. Stat.* 17, 733–744.
- Novi Inverardi, P.L., Petri, A., Pontuale, G., Tagliani, A., 2003. Hausdorff moment problem via fractional moments. *Appl. Math. Comput.* 144, 61–74.
- Shohat, J.A., Tamarkin, J.D., 1943. *The Problem of Moments*. Amer. Math. Soc., New York.
- Stoyanov, J., 2000. Krein condition in probabilistic moment problems. *Bernoulli* 6, 939–949.
- Tagliani, A., Velasquez, Y., 2003. Numerical inversion of the Laplace transform via fractional moments. *Appl. Math. Comput.* 143, 99–107.