Hausdorff moment problem: Reconstruction of distributions

Robert M. Mnatsakanov*

Department of Statistics, West Virginia University, P.O. Box 6330, Morgantown, WV 26506, USA
Health Effects Laboratory Division, National Institute for Occupational Safety and Health, Morgantown, WV 26505, USA

Received 14 September 2006; received in revised form 10 September 2007; accepted 3 January 2008
Available online 16 January 2008

Abstract

The problem of approximation of the moment-determinate cumulative distribution function (cdf) from its moments is studied. This method of recovering an unknown distribution is natural in certain incomplete models like multiplicative-censoring or biased sampling when the moments of unobserved distributions are related in a simple way to the moments of an observed distribution. In this article some properties of the proposed construction are derived. The uniform and $L_1$-rates of convergence of the approximated cdf to the target distribution are obtained.

Keywords: Hausdorff moment problem; Moment-recovered distribution; $L_1$-rate of approximation; Uniform rate of approximation

1. Introduction

Let us consider the so-called probabilistic (Hamburger) moment problem: For a given sequence $\mu = \{1 = \mu_0, \mu_1, \ldots\}$ of real numbers, define a probability distribution $F$ on the real line $(-\infty, \infty)$, such that $\mu_j$ represents the $j$th moment of $F$. When the support of $F$ is the positive real line, then the problem is known as a Stieltjes moment problem. The case when distribution $F$ has a finite support, say, supp{$F$} = [0, $T$] with $T < \infty$, is known as a Hausdorff moment problem. There are two important questions related to the moment problem:

(a) Is $F$ uniquely determined by the sequence of moments $\mu$?
(b) How should this uniquely defined distribution $F$ be reconstructed from $\mu$?

If there is a positive answer to the question (a), we say that the distribution $F$ is moment-determinate ($M$-determinate), otherwise it is $M$-indeterminate. 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), Feller (1971), Lin (1997), and Stoyanov (2000) among others. We will study question (b) in the Hausdorff case.

The reconstruction of a distribution from its moments is very helpful, for instance, when the target distribution has no closed form although the corresponding moments are easily calculated. One such example represents the distribution of a finite weighted sum of independent chi-squared r.v.'s, each with degree of freedom, say 1.

---

* This research was supported by the National Institute for Occupational Safety and Health, contract no. 212-2005-M-12857.
* 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.
0167-7152/ - see front matter © 2008 Elsevier B.V. All rights reserved.
The estimation of distributions via moments is natural in certain incomplete models, e.g., mixtures, multiplicative-censoring or biased sampling, when the moments of unobserved distributions are related in a simple way to the moments of an observed distribution. In all such cases one can estimate the target distribution via the estimation of the transformed moments of observed distributions. See, for example, Mnatsakanov and Klaassen (2003) and Mnatsakanov and Ruymgaart (2005), where the estimation problem of an unknown mixing cdf in the Poisson mixture model and, the estimation of a cdf in the biased sampling model have been studied, respectively.

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) in the case of the Stieltjes moment problem. Lindsay et al. (2000) proposed to approximate the $M$-determinate cdf $F$ by a finite mixture of gamma distributions. The number of components in a mixture is chosen to match several assigned moments of the transformed moments of observed distributions. See, for example, Tagliani and Velasquez (2003) proposed to use the fractional moments in the Stieltjes case as well, assuming

$$
\mu_{\alpha} = \mathbb{E}(X^{\alpha}) = \int_0^\infty t^\alpha dF(t), \quad \alpha \in \mathbb{N},
$$

and assume that the moment sequence $\mu_F = (\mu_{0,F}, \mu_{1,F}, \ldots)$ determines $F$ uniquely, i.e., $F$ is $M$-determinate.

An approximate inverse of the operator $K$ from (1), constructed according to

$$
(K_{\alpha,T}^{-1} \mu_F)(x) = \sum_{k=0}^{[\alpha/2]} \sum_{j=0}^{\alpha-j} \binom{\alpha}{j} \left( \begin{array}{c} j \cr k \end{array} \right) \frac{(-1)^{j-k}}{T^j} \mu_{j,F}, \quad 0 \leq x \leq T, \alpha \in \mathbb{N},
$$

is such that $K_{\alpha,T}^{-1} K F \to w F$, as $\alpha \to \infty$ (see, Mnatsakanov and Ruymgaart (2003), where $T = 1$). 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) hinges on the convergence

$$
B_\alpha(u, v) = \sum_{k=0}^{\lceil \alpha \rceil} \binom{\alpha}{k} u^k (1-u)^{\alpha-k} \rightarrow \begin{cases} 1, & u < v \\ 0, & u > v, \end{cases}
$$

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

For any moment sequence $\nu = \{\nu_j, j \in \mathbb{N}\}$, let us denote by $F_\nu$ the corresponding moment-recovered function via operator $K_{\alpha,T}^{-1}$ as $\alpha \to \infty$ and $\mu_F = \nu$ in (2), i.e.
Proof of Theorem 1. Let us use the following notations:

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

be a pdf of the Beta \((a, b)\) distribution with the parameters \(a = [\alpha_1^\gamma] + 1\) and \(b = \alpha - [\alpha_2^\gamma]\).

3. Asymptotic properties of \(F_{\alpha,v}\)

In this Section we present some asymptotic properties of moment-recovered cdf \(F_{\alpha,v} = \mathcal{K}_{\alpha,T}^{-1} v\) defined in Section 2. The uniform and \(L_1\)-rates of convergence of \(F_{\alpha,v}\) to \(F\) are derived as well.

Let us introduce two different convolutions of cdfs \(F_1\) and \(F_2\). Namely, denote the convolution with respect to the multiplication operation by

\[ F_1 \otimes F_2(x) = \int F_1(x/t) dF_2(t), \quad 0 \leq x \leq T^2, \]

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

\[ F_1 \ast F_2(x) = \int F_1(x - \tau) dF_2(\tau), \quad 0 \leq x \leq 2T. \]

The following notations will be useful as well: if \(v_1 = \{\mu_j, F_1, j \in \mathbb{N}\}\) and \(v_2 = \{\mu_j, F_2, j \in \mathbb{N}\}\) represent two moment sequences, then let us use the following notations: \(v_1 \otimes v_2 = \{\mu_j, F_1 \times \mu_j, F_2, j \in \mathbb{N}\}\) and \(v_1 \oplus v_2 = \{\bar{\nu}_j, j \in \mathbb{N}\}\) where

\[ \bar{\nu}_j = \sum_{m=0}^{j} \binom{j}{m} \mu_{m, F_1} \times \mu_{j-m, F_2}. \]

One can easily verify that the moment sequences of \(F_1 \otimes F_2\) and \(F_1 \ast F_2\) are \(v_1 \otimes v_2\) and \(v_1 \oplus v_2\), respectively. Also the reverse statements are true, see, (i) and (ii) in the following statement.

**Theorem 1.**

(i) If \(v = v_1 \otimes v_2\), then

\[ F_{\alpha,v} \rightarrow_{w} F_v, \quad \text{as } \alpha \rightarrow \infty, \]

with \(F_v = F_1 \otimes F_2\).

(ii) If \(v = v_1 \oplus v_2\), then (6) holds with \(F_v = F_1 \ast F_2\).

(iii) If for some \(a > 0\) and \(b \geq 0\), \(v = \{\nu_j = \mu_{aj+b,F}/\mu_{b,F}, j \in \mathbb{N}\}\), then (6) holds with

\[ F_v(x) = \frac{1}{\mu_{b,F}} \int_0^{x^{1/a}} t^b dF(t), \quad x \in [0, T^a]. \]

(iv) If \(v = \{\bar{\nu}_j, j \in \mathbb{N}\}\), with

\[ \bar{\nu}_j = \int [\phi(t)]^j dF(t), \]

for some continuous and increasing function \(\phi : [0, T] \rightarrow [0, T']\), then (6) holds with \(F_v(x) = F(\phi^{-1}(x)), x \in [0, T']\).

**Proof of Theorem 1.** Note that cdf \(F_1 \otimes F_2\) is defined on \([0, T_0]\), with \(T_0 = T^2\). Therefore, to prove (i), one can use (3) with the arguments of \(B_\alpha(\cdot, \cdot)\) equal to \(tu/T_0\) and \(x/T_0\), respectively. Indeed, with \(v = v_1 \otimes v_2\), we have...
\( F_{\alpha,v}(x) = (K_{\alpha,T_0}^{-1} v)(x) = \int \int \sum_{k=0}^{[\frac{\alpha}{T_0}]} \sum_{j=k}^{\alpha} \binom{\alpha}{j} \left( \frac{(-1)^{j-k}}{T_0^j} \right) t^j u^j dF_1(t) dF_2(u) \)

\[ = \int \int \sum_{k=0}^{[\frac{\alpha}{T_0}]} (\alpha) \binom{\alpha}{k} \left( \frac{tu}{T_0} \right)^k \sum_{m=0}^{\alpha-k} \binom{\alpha-k}{m} \left( \frac{-tu}{T_0} \right)^m dF_1(t) dF_2(u) \]

\[ = \int B_{\alpha} \left( \frac{tu}{T_0}, \frac{x}{T_0} \right) dF_1(t) dF_2(u) \rightarrow \int \int 1_{[0,x]}(tu) dF_1(t) dF_2(u) \]  \hfill (9)

\[ = \int F_1(x/u) dF_2(u) = F_1 \otimes F_2(x). \]

By a similar argument used in (9), we can show (ii)–(iv) very easily. For example, to show (ii), note that replacing in (5) the moments \( \mu_m, F_1 \) and \( \mu_{j-m}, F_2 \) by corresponding integrals with respect to \( F_1 \) and \( F_2 \) we obtain

\[ \bar{v}_j = \int \int (t+u)^j dF_1(t) dF_2(u). \]

Since \( F_1 \otimes F_2 \) is defined on \([0, T_1] \), \( T_1 = 2T \), again using (3) with the arguments of \( B_{\alpha}(-, -) \) equal to \((t+u)/T_1 \) and \( x/T_1 \), respectively, we derive (ii). \( \Box \)

**Corollary 1.**

(i) If \( v = \{\bar{v}_j, j \in \mathbb{N}\} \) with \( \bar{v}_j \) defined according to (8) and \( F_{\alpha,v}^*(x) = F_{\alpha,v}(\phi(x)), x \in [0, T] \), then 

\[ F_{\alpha,v}^* \rightarrow_{w} F, \text{ as } \alpha \rightarrow \infty. \]  \hfill (10)

(ii) If \( v = \{\mu_{aj,F}, j \in \mathbb{N}\} \) for some positive \( a \), then (10) holds with \( F_{\alpha,v}^*(x) = F_{\alpha,v}(x^a), x \in [0, T] \).

(iii) If \( v = \{\beta_1 v_1 + \beta_2 v_2 \} \) with \( \beta_1 + \beta_2 = 1, \beta_1, \beta_2 > 0 \) and \( v_1 = \{a_1^j, j \in \mathbb{N}\} \) and \( v_2 = \{a_2^j, j \in \mathbb{N}\} \) for some \( 0 < a_1 < a_2 < T \), then (6) holds with \( F_v(x) = \beta_1 1_{[a_1,T]}(x) + \beta_2 1_{[a_2,T]}(x) \).

(iv) If \( v = \{a^j \mu_{aj,F}, j \in \mathbb{N}\} \) for some positive \( a \), then (6) holds with \( F_v(x) = F(x/a), x \in [0,aT] \).

**Proof of Corollary 1.** The transformation \( x \mapsto \phi(x) \) and the statement (iv) from Theorem 1 yield Corollary 1(i).

The statement (ii) is a special case of (i) since \( \mu_{aj,F} = \int \phi(t)^j dF(t) \) with \( \phi(t) = t^a \). Since the distribution with the moments \( \mu_j = a^j, j \in \mathbb{N} \) is degenerated at \( a \), the linearity of the inverse transformation \( K_{\alpha,T}^{-1} \) defined in (2) yields Corollary 1(iii), while Theorem 1(i) yields Corollary 1(iv), respectively. \( \Box \)

**Remark 1.** If \( a = 1 \) in Theorem 1(iii), then cdf \( F_v \) from (7) represents the biased sampling model with the weight function \( w(t) = t^b \).

**Remark 2.** The construction \( F_{\alpha,v}^* \) from Corollary 1(ii) can be very helpful when only the finite number of integer moments up to, say, \( m^* < \infty \) are available. In this case one can use the fractional moments \( \mu_j^* = \mu_{aj,F} \) with \( a = m^*/\alpha \) and \( j = 0, 1, \ldots, \alpha \), and reconstruct \( F \) by means of \( F_{\alpha,v}^* \) where \( v = \{\mu_j^*, j = 0, 1, \ldots, \alpha\} \). The application of the fractional moments in the problem of recovering pdf \( f \) based on the ME approach has been considered by Novi Inverardi et al. (2003) (see also the bibliography in the later article).

From Corollary 1(iii) it follows that one can reconstruct (or estimate via the empirical moments) a discrete cdf. Actually, for smooth cdf \( F \) the convergence in (6) is uniform. Since the proofs of statements in Theorem 1(i)–(iv) are similar to each other, let us consider here only the case with \( v = \mu_F \).

**Theorem 2.** Let \( v = \{\mu_{j,F}, j \in \mathbb{N}\} \). If \( f' \) is bounded on \([0, T]\) and \( \alpha \rightarrow \infty \), then

\[ \| F_{\alpha,v} - F \| \leq \frac{T}{\alpha + 1} \left\{ \| f \| + \frac{T}{2} \| f' \| \right\} + o \left( \frac{1}{\alpha} \right). \]  \hfill (11)

\[ \| F_{\alpha,v} - F \|_{L_1} \leq \frac{T}{\alpha + 1} \left\{ 1 + \frac{T^2}{12} \| f' \| \right\} + o \left( \frac{1}{\alpha} \right). \]  \hfill (12)
Proof of Theorem 2. From (1) and (2) we have

\[ F_{\alpha,\nu}(x) = (K_{\alpha,\nu}^{-1} K F)(x) = \int \sum_{k=0}^{\lfloor \alpha^{-1} \rfloor} \sum_{j=0}^{\lfloor \alpha^{1/2} \rfloor} \left( \frac{\alpha}{k} \right) \left( \frac{j}{k} \right) \left( \frac{\mu}{T} \right)^k \left( \frac{-\mu}{T} \right)^{j-k} \ dF(t) \]

\[ = \int \sum_{k=0}^{\lfloor \alpha^{1/2} \rfloor} \left( \frac{\alpha}{k} \right) \left( \frac{\mu}{T} \right)^k \sum_{m=0}^{\lfloor \alpha^{-1} \rfloor} \left( \frac{\alpha}{m} \right) \left( \frac{-\mu}{T} \right)^m \ dF(t) \]

(13)

Now, integration by parts in the last equation of (13) enable us to write

\[ F_{\alpha,\nu}(x) - F(x) = \int_0^1 [F(uT) - F(x)] \beta(u, a, b) du, \]

(14)

where

\[ F(uT) - F(x) = f(x) T \left( u - \frac{x}{T} \right) + T^2 \int_0^u \int_{u}^{T} f'(yT) dy. \]

(15)

Note also that the mean and variance of Beta \((a, b)\) distribution defined in (4) are such that

\[ \eta_a := \int_0^1 u \beta(u, a, b) du = \frac{[\alpha^{-1} + 1]}{\alpha + 1}, \]

(16)

\[ \sigma_a^2 := \int_0^1 (u - \eta_a)^2 \beta(u, a, b) du = \frac{([\alpha^{-1} + 1](\alpha - [\alpha^{-1}]))}{(\alpha + 1)^2(\alpha + 2)}, \]

(17)

and

\[ |\eta_a - \frac{x}{T}| \leq \frac{1}{\alpha + 1} \max \left( \frac{x}{T}, 1 - \frac{x}{T} \right). \]

(18)

Hence, the combination of (14)--(18) yields (11):

\[ \|F_{\alpha,\nu} - F\| \leq \sup_{0 \leq x \leq T} \left| f(x) T \int_0^1 \left( u - \frac{x}{T} \right) \beta(u, a, b) du \right| \]

\[ + \sup_{0 \leq x \leq T} \left| \int_0^1 \beta(u, a, b) du \left[ T^2 \int_0^u \int_{u}^{T} f'(yT) dy \right] \right| \]

\[ \leq T \|f\| \sup_{0 \leq x \leq T} \left| \eta_a - \frac{x}{T} \right| + \|f'\| T^2 \sup_{0 \leq x \leq T} \int_0^1 \frac{1}{2} \left( u - \frac{x}{T} \right)^2 \beta(u, a, b) du \]

\[ \leq \frac{T}{\alpha + 1} \|f\| + \frac{T^2}{2(\alpha + 1)} \|f'\| + \frac{T^2}{2(\alpha + 1)^2} \|f''\|. \]

The proof of (12) is based on Eqs. (14)--(18) again and is omitted here. \[\square\]

4. Examples

Example 1. Let us recover the distribution \( F \) with assigned moments \( v_j = 9/(j + 3)^2, j \in \mathbb{N} \). Note that \( v_j = 1/(j + 3)^2 = v_{ajG} \), where \( a = 1/3, v_G = \mu \odot \mu, \) and \( \mu = (1/(j + 1), j \in \mathbb{N}) \). Since the uniform cdf \( U[0, 1] \) is \( M \)-determinate with the moment sequence \( \mu = \{1/(j + 1), j \in \mathbb{N}\} \), we conclude from Theorem 1, see (i), that \( G(x) = U[0, 1] \otimes U[0, 1](x) = x - x \log(x), 0 \leq x \leq 1 \). Now taking \( a = 1/3, \) and \( b = 0 \) in Theorem 1(iii), we derive \( F(x) = G(x^3) = x^3 - x^3 \log(x^3), 0 \leq x \leq 1 \). We conducted the computations of moment-recovered cdf \( F_{\alpha,\nu} \) when \( \alpha = 30 \). See Fig. 1 with \( F \), the dashed curve, and \( F_{\alpha,\nu} \), the step function, respectively.
Example 2. Let us approximate the cdf $A_2$ of the sum of two independent Beta$(1/2, 1)$ random variables having the sequence of moments: $\mu_{j,F_k} = 1/(2j + 1)$, $j \in \mathbb{N}$, $k = 1, 2$. Based on Theorem 1(ii), we can use the approximation $F_{u,v}$ with $v = \{\tilde{v}_j, j \in \mathbb{N}\}$ defined according to (5) and $\mu_{j,F_k} = 1/(2j + 1)$. On Fig. 2 we plotted

$$A_2(x) = \begin{cases} \frac{\pi x}{4}, & 0 \leq x \leq 1 \\ \sqrt{x-1} + \frac{x}{2} \left(2 \sin^{-1} \sqrt{\frac{1}{x} - \frac{\pi}{2}} \right), & 1 < x \leq 2, \end{cases}$$

the dashed curve, and its moment-recovered cdf $F_{u,v}$, the step function, respectively. Here we took again $\alpha = 30$.

Remark 3. By taking parameter $\alpha = 100$, one can verify that in the Examples 1 and 2 the approximants $F_{u,v}$ almost coincide with the corresponding cdfs $F$ and $A_2$, respectively.

Fig. 1. Approx. of $F$ by $F_{u,v}$.

Fig. 2. Approx. of $A_2$ by $F_{u,v}$.
Acknowledgments

The author is thankful to Estate Khmaladze and Cecil Burchfiel for helpful discussions and S. Li for conducting the calculations in Examples 1 and 2. The findings and conclusions in this paper are those of the author’s and do not necessarily represent the views of the National Institute for Occupational Safety and Health.

References


