On the moment-recovered approximations of regression and derivative functions with applications

Robert M. Mnatsakanov ∗, Broti Garai

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

A R T I C L E   I N F O
Article history:
Received 7 April 2016
Received in revised form 25 October 2016

Keywords:
Moment-determinate functions
Laplace transform inversion
Nonparametric estimation

A B S T R A C T
In this paper three formulas for recovering the conditional mean and conditional variance based on product moments are proposed. The upper bounds for the uniform rate of approximations of regression and derivatives of some moment-determinate function are derived. Two cases where the support of underlying functions is bounded and unbounded from above are studied. Based on the proposed approximations, novel nonparametric estimates of the distribution function and its density in multiplicative-censoring model are constructed. Simulation study justifies the consistency of the estimates.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

In many statistical problems the goal is to express the response Y variable as a function of the predictive variable X, i.e., Y = r(X). Suppose instead of observing the variables (X, Y), we only have information about the averages/moments of the form: m(k, j) = E(Y | X = x), with k = 0, 1, 2 and j = 0, 1, . . . , α. The objective of this paper is to recover regression function r(x) = E(Y | X = x) and conditional variance σ²(x) = Var(Y | X = x) given the information contained in the sequence of the product moments m(k, j), k = 0, 1, 2 and j = 0, 1, . . . , α. Also we address the problem of recovering the derivative of a density function of some random variable, say Y ∼ g, given only its moments E(Y j) for j = 0, 1, . . . , α. Application in the multiplicative-censoring model is outlined as well. In particular, we derived the consistent nonparametric estimates of associated distribution and its density function.

In this paper it is assumed that the regression function that is defined on the positive half line is moment-determinate (M-determinate). The conditions under which a function is M-determinate or M-indeterminate is a classical mathematical problem and has been studied in many works, see, for example, [1–4] among others. Besides, there are several techniques that provide the reconstruction of probability density function from the sequence of its moments. Let us mention here only three of them. The maximum entropy method was used in [5–7], to reconstruct the unknown density from its moments. The article [8] showed that under certain conditions the density function can be approximated by f̃(x) = ψ(x) ∑d k=0 ξk xk. Here ψ(x) is the initial density approximation of f(x), the ξk’s are determined by equating the first d moments obtained from f̃(x) to those of X, i.e. via

\[
\begin{pmatrix}
ξ_0 \\
ξ_1 \\
\vdots \\
ξ_d
\end{pmatrix} = \begin{pmatrix}
m(0) & m(1) & \cdots & m(d) \\
m(1) & m(2) & \cdots & m(d+1) \\
\vdots & \vdots & \ddots & \vdots \\
m(d) & m(d+1) & \cdots & m(2d)
\end{pmatrix}^{-1} \begin{pmatrix}
1 \\
\mu_X(1) \\
\vdots \\
\mu_X(d)
\end{pmatrix}
\]

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

http://dx.doi.org/10.1016/j.cam.2016.10.023
0377-0427/© 2016 Elsevier B.V. All rights reserved.
In the previous equation $\mu_{X}(j) = E(X^j)$ and $m(j) = \int_{0}^{1} x^j \psi(x) dx$.

In the framework of moment problems (Hausdorff, Stieltjes and Hamburger), (see [9–12]) a method for recovering the distribution function (df) as well as corresponding density is proposed. These reconstructions are based on the sequence of $\phi$-transformed moments $[m_{\phi}(j)]_{j \geq 0}^{\nu_{\phi}}$ of $F$ up to some order $\alpha$.

The approximated df and density can be written in a uniform form: for each $\alpha \in \mathbb{N}$, $F_{\alpha,\phi} := x^{-1}m_{\phi}$ and $f_{\alpha,\phi} := \phi^{-1}m_{\phi}$ where

$$
(x^{-1}m_{\phi})(x) = \sum_{k=0}^{\lfloor \alpha \phi(x) \rfloor} \sum_{j=k}^{\alpha} \binom{\alpha}{j} (-1)^{j-k} m_{\phi}(j) \tag{1}
$$

and

$$
(\phi^{-1}m_{\phi})(x) = \frac{\Gamma(\alpha + 2)(\phi(x))}{\Gamma([\alpha\phi(x)] + 1)} \sum_{j=0}^{\lfloor \alpha \phi(x) \rfloor} \frac{(-1)^{j} m_{\phi}(j + \lfloor \alpha \phi(x) \rfloor)}{j!(\alpha - \lfloor \alpha \phi(x) \rfloor - j)!}, \quad x \in S. \tag{2}
$$

Here, $m_{\phi}(j) = \int [\phi(t)]^j dF(t)$, for $j = 0, 1, \ldots, \alpha$, and $S := \text{supp}(F)$ is a support of $F$. It is assumed that the map $\phi: S \mapsto [0, 1]$ is specified according to the form of $S$. For example, $\phi(x) = x/T$, if $S = [0, T]$, and $\phi(x) = b^{-x}$ if $S = \mathbb{R}_{+}$, for some $0 < T < \infty$ and $b > 1$. The recovered constructions $F_{\alpha,\phi}$ and $f_{\alpha,\phi}$ provide stable approximates of $F$ and $f$, respectively (see [12]).

It is worth mentioning that the first two methods were applied only for reconstruction of a density $f$ while the last one can be used for recovering arbitrary $M$-determinate function (e.g., the derivative) as well (see [12] and Example 5 in Section 3). This fact enables us to apply the construction (2) to approximate the regression function, the conditional variance as well as the derivative function.

Under very smooth conditions on underlying regression function $r$, [13] approximated the conditional expectation $E(Y|X = x)$ using the sequence of moments $m(1, j)$ for $j = 0, 1, \ldots, 2n$. Namely, when the support of $r$ is finite, say $[0,1]$, the following approximation was introduced:

$$
r_n(x) = \frac{1}{f(x)} \sum_{j=0}^{2n} v_{n(M_n+h)/M_n} \frac{M_{n+1}^{j+1}}{M_n^{j+1}} m_{1,j}, \quad 0 \leq x \leq 1. \tag{3}
$$

Here $v_{ij}$ is the $(i,j)$th entry of the inverted Vandermonde matrix $V(-n, -n + 1, \ldots, n) [14]. f(x) = F'(x)$ is the marginal density of $X$ and $M_n$ is a sequence of positive real numbers such that $M_n/n \to 0$ as $n \to \infty$. The author showed that the reconstruction of $r(x)$ based on (3) provides an improved approximation if compared to normal approximation.

The article is organized as follows. In Section 2 two cases of recovering regression function are investigated: (a) the distribution of a predictive variable $X$ is known and (b) the distribution of $X$ is unknown. In Section 3 approximation of the derivative function $g'$ for some continuous density function $g$ is provided utilizing the proposed construction. The upper bounds for the uniform rate of convergence for the approximated regression and the derivative functions are derived and asymptotic behavior of the distances between the approximated and the true regression functions is investigated. Furthermore, in Section 3 we address the problem of estimating the distribution function $F$ and corresponding density $f$ in the framework of multiplicative-censoring model. In Sections 2 and 3 the simulation study and a comparison with other constructions are conducted. Graphical illustrations and tables with the values of errors of recovered functions are provided. The mean squared error and integrated mean squared error of associated nonparametric estimators will be the subject of investigation in a separate work.

2. Approximations of $r(x)$

2.1. Distribution of $X$ is known

In this subsection we will approximate regression function $r(x)$ when the distribution of $X$ is known. In addition, we assume that the distribution function $F$ of $X$ possesses a finite support, i.e., $\text{supp}(F) = [0, T]; 0 < T < \infty$. For the sake of simplicity we assume $T = 1$. Two cases are considered in this subsection: (1) $X$ follows the uniform distribution on $(0,1)$, i.e., $X \sim U(0, 1)$ (see Example 1), and (2) $X \sim \text{Beta}(a, b)$ with known parameters $a$ and $b$ (see Example 2). The density of $X$ is denoted by $\beta(x, a, b) = \frac{1}{B(a,b)} x^{a-1}(1-x)^{b-1}$ for $x \in (0, 1), a > 0, b > 0$. Here $B(a,b)$ is the classical Euler-beta function.

2.1.1. Model 1: $X \sim U(0, 1)$

Let us assume $X \sim U(0, 1)$ with some joint probability density function of $(X, Y)$ denoted by $h : [0, 1]^2 \to \mathbb{R}_{+}$. Denote $m_k = \{m(k,j), j = 0, 1, \ldots, \alpha\}$ where $k = 0, 1, 2$, and

$$
m(k,j) = E(Y^jX^k) = \int_{0}^{1} \int_{0}^{1} y^j x^k h(y, x) \, dy \, dx.
$$
Our first construction defines the following approximation of a regression function \( r(x) = E(Y|X = x) \) specified by Eq. (2) with \( \phi(x) = x \):

\[
r_{a, 1}(x) = (\beta^{-1}_a m_1)(x),
\]

where

\[
(\beta^{-1}_a m_1)(x) = \frac{\Gamma(a + 2)}{\Gamma([a] + 1)} \sum_{j=0}^{\alpha - [a]} \frac{(-1)^j m(1, j + [a])}{j! (\alpha - [a] - j)!}, \quad 0 \leq x \leq 1.
\]

Convergence in the sup-norm will be denoted by “\( \rightarrow \)”. The following statement regarding the upper bound of the approximated regression function \( r_{a, 1} \) is valid.

**Theorem 1.** (i) If \( r \) is continuous on \((0, 1)\), then \( r_{a, 1} \rightarrow r \) as \( \alpha \rightarrow \infty \), and

\[
\| r_{a, 1} - r \| \leq \Delta(r, \delta) + \frac{2\|r\|}{\delta^2(a + 2)},
\]

where \( 0 < \delta < 1 \), \( \Delta(r, \delta) = \sup_{|u - x| \leq \delta} |r(u) - r(x)| \) represents the modulus of continuity of \( r \), and \( \|r\| \) is the sup-norm of \( r \).

(ii) If \( r'' \) is continuous on \((0, 1)\), then for each \( x \in (0, 1) \)

\[
r_{a, 1}(x) - r(x) = \frac{1}{\alpha + 2} \left[ (1 - 2x + [a\alpha] - \alpha\alpha)r'(x) + \frac{1}{2} x(1 - x)r''(x) \right] + o\left(\frac{1}{\alpha}\right), \quad \text{as } \alpha \rightarrow \infty,
\]

and

\[
\| r_{a, 1} - r \| \leq \frac{1}{\alpha + 2} \left\{ 2\|r''\| + \frac{1}{8}\|r''\| \right\} + o\left(\frac{1}{\alpha}\right), \quad \text{as } \alpha \rightarrow \infty.
\]

**Proof.** (i) The approximated regression function

\[
r_{a, 1}(x) = \frac{\Gamma(a + 2)}{\Gamma([a\alpha] + 1)} \sum_{j=0}^{\alpha - [a]} \frac{(-1)^j E(YX^{j+[a\alpha]})}{j! (\alpha - [a\alpha] - j)!}
\]

can be rewritten as,

\[
r_{a, 1}(x) = \frac{\Gamma(a + 2)}{\Gamma([a\alpha] + 1)} \sum_{j=0}^{\alpha - [a]} \frac{(-1)^j E(YX^{j+[a\alpha]})}{j! (\alpha - [a\alpha] - j)!} \int_0^1 E(Y|X = u)f(u)du
\]

\[
= \frac{\Gamma(a + 2)}{\Gamma([a\alpha] + 1)} \sum_{j=0}^{\alpha - [a]} \frac{(-1)^j}{j! (\alpha - [a\alpha] - j)!} \int_0^1 u^{j+[a\alpha]}E(Y|X = u)f(u)du.
\]

Changing the order of integration and summation and considering \( X \sim U(0, 1) \), we arrive at

\[
r_{a, 1}(x) = \int_0^1 E(Y|X = u) \frac{\Gamma(a + 2)}{\Gamma([a\alpha] + 1)} \sum_{j=0}^{\alpha - [a]} \frac{(-u)^j}{j! (\alpha - [a\alpha] - j)!}u^{j+[a\alpha]}du
\]

\[
= \int_0^1 \beta(u, [a\alpha] + 1, \alpha - [a\alpha] + 1)r(u)du.
\]

From (7) we derive \( r_{a, 1}(x) \rightarrow r(x) \) uniformly in \( x \) (see [15]). Splitting the range of integration in Eq. (7) into two parts yields,

\[
\| r_{a, 1} - r \| \leq \sup_{0 \leq \delta \leq 1} \left( \int_{|u - x| \leq \delta} |r(u) - r(x)| \beta(u, [a\alpha] + 1, \alpha - [a\alpha] + 1)du \right).
\]

Denoting \( \Delta(r, \delta) = \sup_{|u - x| \leq \delta} |r(u) - r(x)| \) we obtain,

\[
\| r_{a, 1} - r \| \leq \Delta(r, \delta) + R_\alpha \quad \text{with } 0 \leq \delta \leq 1.
\]

Using Chebyshev’s inequality we get the upper bound of \( R_\alpha \) as,

\[
R_\alpha \leq 2\|r\| \sup_{0 \leq \delta \leq 1} \int_{|u - x| > \delta} \beta(u, [a\alpha] + 1, \alpha - [a\alpha] + 1)du \leq \frac{2\|r\|}{\delta^2(a + 2)}.
\]

Substituting (9) in (8) completes the proof.
(ii) Note that, $\beta(\cdot, [\alpha x] + 1, \alpha - [\alpha x] + 1)$ forms a $\delta$-sequence as $\alpha \to \infty$ with mean, $\theta_\alpha = \frac{[\alpha x] + 1}{\alpha + 2}$ and variance

$$\sigma^2_\alpha = \frac{([\alpha x] + 1)(\alpha - [\alpha x] + 1)}{(\alpha + 2)^2(\alpha + 1)}.$$ 

Here $\theta_\alpha - x = \frac{1 - 2x}{\alpha + 2} + \Delta_{1,\alpha}(x)$ with $\Delta_{1,\alpha}(x) = \frac{[\alpha x] - \alpha x}{\alpha + 2}$ and $\sigma^2_\alpha = \frac{\chi(1 - x)}{(\alpha + 2)^2} + \Delta_{2,\alpha}(x)$, where $\Delta_{2,\alpha} \leq \frac{2}{(\alpha + 2)^2}$ (see [16, 17]). The Taylor series expansion of $r$ at $x$ and the continuity property of $r''$ proves the theorem. The second statement of (ii) follows immediately from the first one, since for any $x \in [0, 1]$ we have $-2 \leq 1 - 2x + [\alpha x] - \alpha x \leq 1$, and $0 \leq x(1 - x) \leq 1/4$. □

**Corollary 1.** If $\tilde{\alpha} = 2\alpha$, $\tilde{r}_{a,1}(x) = 2r_{a,1}(x) - r_{a,1}(x)$, and $r''$ is a continuous function, then for each $x \in (0, 1)$

$$\tilde{r}_{a,1}(x) - r(x) = \frac{1}{(\alpha + 1)(\alpha + 2)} \left[ r''(x) \left( 1 - 2x + (\alpha + 3([\alpha x] - \alpha x) \right) + \frac{1}{2} r'(x)(1 - x) \right]$$

as $\alpha \to \infty$.

**Proof.** The difference between $\tilde{r}_{a,1}(x)$ and $r(x)$ can be expressed as

$$r_{a,1}(x) - r(x) = \frac{1}{(\alpha + 1)(\alpha + 2)} \left[ r''(x) \left( 1 - 2x + (\alpha + 3([\alpha x] - \alpha x) \right) + \frac{1}{2} r'(x)(1 - x) \right]$$

Applying the Taylor series expansion of $r$ in Eq. (10) and using the properties of $\beta(\cdot, [\alpha x] + 1, \alpha - [\alpha x] + 1)$ yields

$$\tilde{r}_{a,1}(x) - r(x) = \frac{1}{(\alpha + 1)(\alpha + 2)} \left[ r''(x) \left( 1 - 2x + (\alpha + 3([\alpha x] - \alpha x) \right) + \frac{1}{2} r'(x)(1 - x) \right]$$

Theorem 1. Analogous to the conditional expectation, two approximations of the conditional variance $\sigma^2(x) = E(Y^2|X = x) - (E(Y|X = x))^2$ can be proposed as well:

$$\sigma^2_{a,1}(x) = (\varphi^{-1}_\alpha m_2)(x) - \left( (\varphi^{-1}_\alpha m_1)(x) \right)^2$$

and

$$\tilde{\sigma}^2_{a,1}(x) = 2\sigma^2_{a,1}(x) - \sigma^2_{a,1}(x) \quad \text{with } \tilde{\alpha} = 2\alpha.$$ 

The proofs of statements similar to the ones presented in Theorem 1 and Corollary 1, as well as in the next subsection are omitted.

**Example 1.** Let us specify the marginal distributions of $X$ and $Y$ as follows: $X \sim U(0, 1)$; $Y \sim F(y) = y^{1/2}$. Define the joint density of $X$ and $Y$ according to the Farlie–Gumbel–Morgenstern copulas with $\theta = \frac{1}{2}$ (see, for example, [18]). In this case, $h(x, y) = \frac{1}{2} y^{1/2}$, where $0 < x, y < 1$. Corresponding regression function is $r(x) = (2x + 3)/12$, for $x \in [0, 1]$, and the product moments $m(1, j) = (5j + 8)/(12(j + 1)(j + 2))$. $j = 0, 1, \ldots$. The approximated regression functions $r_{a,1}$ and $\tilde{r}_{a,1}$ are constructed utilizing Eq. (4) and Corollary 1 respectively for $\alpha = 30$ and $\alpha = 15$. In Fig. 1, the curves of $r_{a,1}$ and $\tilde{r}_{a,1}$ along with the true regression function $r$ are displayed.

As an application, let us apply (2) for approximating the conditional variance $\sigma^2(x)$. In this example we have $E(Y^2|X = x) = (40x + 28)/210$ and $m(2, j) = (68 + 96)/210$, $j = 0, 1, \ldots$. Fig. 2 displays the approximated conditional variances $\sigma^2_{a,1}(x)$ and $\tilde{\sigma}^2_{a,1}(x)$ when $\alpha = 30$, and $\alpha = 15$, respectively, as well as the true $\sigma^2(x)$. 


Proof. (iii) It is known that $r(x)$ is continuous on $[0, 1]$ and

$$r_{\alpha, 2}(x) = \frac{\Gamma(\alpha + 2)}{\Gamma([ax] + 1)} \int_0^1 u^{[ax]} \left( \sum_{j=0}^{a-[ax]} \frac{(-u)^j}{j!(\alpha-[ax]-j)!} \right) f(Y|X = u) \frac{f(u)}{f(x)} \, du. \tag{12}$$
Changing the order of integration and summation of (12) we get,
\[ r_{a,2}(x) = \int_0^1 \beta(u, [\alpha x] + 1, \alpha - [\alpha x] + 1) r(u) \frac{f(u)}{f(x)} du \]
\[ = \frac{1}{f(x)} \int_0^1 \beta(u, [\alpha x] + 1, \alpha - [\alpha x] + 1) \psi(u) du. \]  
(13)
Therefore from (13) we get \( r_{a,2}(x) = r(x) \) uniformly in \( x \) (see [15]).
(ii) Following the footsteps of Theorem 1, we split the above integration into two parts
\[ \| r_{a,2} - r \| \leq \sup_{0 \leq x \leq 1} \int (\int_{|u-x| \leq \delta} + \int_{|u-x| > \delta}) |\psi(u) - \psi(x)| \beta(u, [\alpha x] + 1, \alpha - [\alpha x] + 1) du \]  
(14)
In addition let us denote,
\[ \Delta(\psi, \delta) = \sup_{|u-x| \leq \delta} |\psi(u) - \psi(x)| \leq \|r\| \Delta(f, \delta) + \|f\| \Delta(r, \delta). \]  
(15)
From (14) and (15) we have,
\[ \| r_{a,2} - r \| \leq \frac{\Delta(\psi, \delta)}{\beta} + R_a \quad \text{where} \quad 0 < \delta \leq 1. \]
Using Chebyshev’s inequality we finally get,
\[ R_a \leq \frac{2 \| \psi \| \Delta(\psi, \delta)}{\beta} \sup_{0 \leq x \leq 1} \int \beta(u, [\alpha x] + 1, \alpha - [\alpha x] + 1) du \leq \frac{2 \| r \| \| f \|}{\delta^2 \beta(\alpha + 2)} \]
which completes the proof of (ii). The proofs of statements in (iii) can be easily derived. \( \square \)

**Corollary 2.** Suppose \( \psi = tf, \alpha = 2 \alpha, \tilde{r}_{a,2}(x) = 2r_{a,2}(x) - r_{a,2}(x), \) and \( \psi'' \) is continuous, then for each \( x \in (0, 1) \)
\[ \tilde{r}_{a,2}(x) - r(x) = \frac{2}{(\alpha + 1)(\alpha + 2)} \left[ \left( 1 - 2x + (\alpha + 3)([\alpha x] - \alpha x) \right) + \frac{1}{2f(x)} \psi''(x)(1 - x) \right] \]
\[ + \frac{1}{f(x)} \left( 2\Delta_{2,\alpha}(x) - \Delta_{2,\alpha}(x) \right) \psi''(x) + O \left( \frac{1}{\alpha^2} \right), \]
as \( \alpha \to \infty. \)

**Proof.** Let us rewrite \( \tilde{r}_{a,2}(x) - r(x) \) as follows:
\[ \tilde{r}_{a,2}(x) - r(x) = 2(r_{a,2}(x) - r(x)) - (r_{a,2}(x) - r(x)). \]  
(16)
Similar to Corollary 1, Eq. (16) yields,
\[ \tilde{r}_{a,2}(x) - r(x) = \frac{2}{(\alpha + 1)(\alpha + 2)} \left[ \left( 1 - 2x + \Delta_{1,\alpha}(x) + \Delta_{1,\alpha}(x) \right) \psi''(x) + \frac{1}{2f(x)} \psi''(x)(1 - x) \right] \]
\[ + \frac{1}{f(x)} \left( 2\Delta_{1,\alpha}(x) - \Delta_{1,\alpha}(x) \right) \psi''(x) + O \left( \frac{1}{\alpha^2} \right). \]  
(17)

**Example 2.** Suppose \( h(x, y) = \left( \frac{x}{(x+y)^{\alpha-1}} \right)(1-x-y)^{\alpha-1} \) is a bivariate Dirichlet joint density, with \( 0 < x, y < 1 \) and \( x + y < 1 \). In this example \( r(t) = \frac{a(1-t)}{a+t}, \) for \( t \in [0, 1] \) and \( m(1, j) = \frac{(a)(b)}{(a+b+c)} \), where \( (a) \) and \( (b) \) are Gamma functions. Assume \( a = 1/3, \ b = 1 \) and \( c = 1/2 \). The approximated regression functions \( r_{a,2} \) (when \( \alpha = 40 \)) and \( \tilde{r}_{a,2} \) (when \( \alpha = 20 \)) along with the true regression function \( r \) are displayed in Fig. 3.

The approximation of the conditional variance: \( \sigma^2 = E(Y^2 | x = x) - (E(Y | x = x))^2 \) with \( E(Y^2 | x = x) = \frac{b(b+1)(1-x)^2}{(b+c)(b+c+1)} \) is also considered. Let us define the approximated conditional variance \( \sigma_{a,2}^2(x) \) as
\[ \sigma_{a,2}^2 = \frac{1}{f(x)} \left( \phi_a^{-1} m_2 \right)(x) - \left( \frac{1}{f(x)} \left( \phi_a^{-1} m_1 \right)(x) \right)^2 \]
Fig. 3. Approximated regression function (blue dots) in (a) $r_{\alpha,2}$ when $\alpha = 40$; and in (b) $\bar{r}_{\alpha,2}$ when $\alpha = 20$. The true regression $r(x) = 2(1 - x)/5$ (orange curve). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Fig. 4. Approximated conditional variance (blue dots) in (a) $\sigma_{\alpha,2}^2(x)$ when $\alpha = 30$; and in (b) $\bar{\sigma}_{\alpha,2}^2(x)$ when $\alpha = 15$. The orange curve represents the true conditional variance $\sigma^2(x)$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

and its modified version as follows:

$$\bar{\sigma}_{\alpha,2}^2(x) = 2\sigma_{\alpha,2}^2(x) - \sigma_{\alpha,2}^2(x), \quad \text{with } \bar{\alpha} = 2\alpha.$$  

Note that

$$m(2,j) = \frac{(a)_j(b)_j}{(a + b + c)_{2+j}}.$$  

The plots (a) and (b) in Fig. 4 display the approximated and true conditional variances for $\alpha = 30$ and $\alpha = 15$, respectively. In both plots the values of $a$, $b$ and $c$ are taken as $1/3$, $1$ and $1/2$, respectively. The approximations resemble the true regression function for $\alpha \geq 30$.

2.2. Distribution of $X$ is unknown

In this subsection we consider

Model 3: Let $X \sim F$ and assume $F$ is an absolutely continuous distribution with the density $f(x) > 0$. Consider the case when supp$[F] = \mathbb{R}_+$, the moments of $F$ are finite and $F$ is $M$-determinate. The case when support of $F$ is finite: $[0, T], \ 0 < T < \infty$, and $F$ is $M$-determinate, can be handled using a similar argument. Namely, one can apply (17) and (18) (see below) with $x \in [0, T], \ 0 \leq x \leq T$.

To approximate the conditional mean $r(x)$ and conditional variance $\sigma^2(x)$, let us introduce the following sequences of moments $m_k = \{m(k,j), j = 0, 1, \ldots, \alpha\}$ with

$$m(k,j) = E(Y^k[\phi(X)^j]).$$

Here $\phi(x) = b^{-x}$ for some $b > 1$, and $k = 0, 1, 2$. We will evaluate the values of all proposed approximants at $x_j = \log\left(\frac{a}{a+1}\right) / \log(b)$ for $j = 1, \ldots, \alpha$.

Assuming that the two moment sequences $m_0 = \{m(0,j), j = 0, 1, \ldots, \alpha\}$ and $m_1 = \{m(1,j), j = 0, 1, \ldots, \alpha\}$ are given, we suggest the approximation of $r(x)$:

$$r_{\alpha,b}(x) = \left(\left(\sigma_{\alpha,b}^{-1}m_1 \circ \phi\right)(x) / \left(\sigma_{\alpha,b}^{-1}m_0 \circ \phi\right)(x)\right)^{\frac{1}{\alpha}}, \quad x \in \mathbb{R}_+.$$  

(17)
Here, in (17) both the numerator and denominator can be written as follows:

\[
(\varphi^{-1}_a m_k \circ \phi)(x) = C_a(x) \sum_{j=0}^{\alpha - [\alpha \phi(x)]} \frac{(-1)^j m(k, j + [\alpha \phi(x)])}{j! (\alpha - [\alpha \phi(x)] - j)!},
\]

(18)

where \( C_a(x) = \frac{[\alpha \phi(x)]!}{\alpha^{[\alpha \phi(x)]}} \frac{\Gamma(\alpha + 2)}{\Gamma(\alpha + 1)} \) and \( k = 0, 1 \). The expression (18) with \( k = 2 \) will be used for recovering the conditional variance.

**Example 3.** Let the density of \( X \) be a mixture of two gamma densities \( g(\cdot, 1, 1) \) and \( g(\cdot, 3, 1) \) with the equal weights: \( f(x) = g(x, 1, 1)/2 + g(x, 3, 1)/2 \), while the joint density is \( h(x, y) = g(x, 1, 1)g(y, 4, 1)/2 + g(x, 3, 1)g(y, 2, 1)/2 \). In this example \( r(t) = (2t^2 + 8)/(t^2 + 2) \) for \( t > 0 \). It is easy to calculate \( m_0 \) and \( m_1 \) explicitly as

\[
m(0, j) = \frac{1}{2(j \ln b + 1)} + \frac{1}{2(j \ln b + 1)^3},
\]

\[
m(1, j) = \frac{2}{j \ln b + 1} + \frac{1}{(j \ln b + 1)^3}.
\]

The approximated regression curves \( r_{a,b} \) (with \( \alpha = 40 \)) and \( \tilde{r}_{a,b} \) (with \( \alpha = 20 \)) are presented in Fig. 5. The parameter \( b \) is taken as 1.15 in both approximates.

The approximations of conditional variance \( \sigma^2(x) \) with \( E(Y^2 | X = x) = (6x^2 + 40)/(x^2 + 2) \) are defined as

\[
\sigma^2_{r,a,b}(x) = \frac{(\varphi^{-1}_a m_2 \circ \phi)(x)}{(\varphi^{-1}_a m_0 \circ \phi)(x)} - \left( \frac{(\varphi^{-1}_a m_1 \circ \phi)(x)}{(\varphi^{-1}_a m_0 \circ \phi)(x)} \right)^2
\]

and

\[
\tilde{\sigma}^2_{r,a,b}(x) = 2\sigma^2_{r,a,b}(x) - \sigma^2_{r,a,b}(x)
\]
The sequences with Example 4. The approximated conditional variance is shown in Fig. 8. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

with $m(2,j) = \frac{10}{1+j\ln b} + \frac{3}{(1+j\ln b)^2}$. Fig. 6 shows the approximated conditional variances $\sigma^2_{\alpha,b}(x)$ and $\tilde{\sigma}^2_{\alpha,b}(x)$ when $\alpha = 40$ and $\alpha = 20$, respectively, and $b = 1.15$. The true conditional variance $\sigma^2(x)$ (the orange curve) is shown as well.

Example 4. Assume that the joint distribution of $X$ and $Y$ is defined according to the copula function $C_{\theta}(u, v) = uv + \theta u(1-u)v(1-v)$ with parameter $\theta$ (see also [18]). The marginal distributions of $X$ and $Y$ considered in this example are exponential with $G(x) = 1 - e^{-\lambda_1 x}$ and $F(y) = 1 - e^{-\lambda_2 y}$, respectively. Here $\lambda_1, \lambda_2 > 0$ are known parameters.

The conditional expectation of $Y$ given $X = x$ has the following form: $r(x) = \frac{\lambda_1}{\lambda_2} + \frac{\theta \lambda_1}{2\lambda_2} (1 - 2e^{-\lambda_1 x})$. It is easy to see that the sequences $m_0$ and $m_1$ are defined as

$$m(0,j) = \frac{\lambda_1}{\lambda_1 + j\ln(b)},$$

$$m(1,j) = \frac{\lambda_1}{\lambda_2 (\lambda_1 + j\ln(b))} + \frac{\theta \lambda_1}{2\lambda_2 (\lambda_1 + j\ln(b))} - \frac{\theta \lambda_1}{\lambda_2 (2\lambda_1 + j\ln(b))}.$$

A comparison between approximated regression functions $r_{\alpha,b}$ (when $\alpha = 30$), $\tilde{r}_{\alpha,b}$ (when $\alpha = 15$) with $r$ for $\theta = -0.9$, $\lambda_1 = 0.5$, $\lambda_2 = 0.7$, and $b = 1.2$ is displayed in Fig. 7.

The approximated conditional variance is shown in Fig. 8. Note that in Example 4, we have $E(Y^2|X = x) = \frac{\lambda_1}{\lambda_2^2} + \frac{3\theta \lambda_1}{2\lambda_2^2} (1 - 2e^{-\lambda_1 x})$ and

$$m(2,j) = \frac{2\lambda_1}{\lambda_2^2 (\lambda_1 + j\ln b)} + \frac{3\theta \lambda_1}{2\lambda_2^2 (\lambda_1 + j\ln b)} - \frac{3\theta \lambda_1}{\lambda_2^2 (2\lambda_1 + j\ln b)}.$$
To recover
Let us denote
nodes.
used the values of
in
Example 4.
The uniform rate of approximation of \( r_{a,b} \) when \( \alpha \) is varying in the range \([30, 230]\) and \( b \in [1.1, 2.5] \).
(For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

\[ f(x) = \frac{3x + 3}{12} \]

The records of the distances in\( sup \)-norm between \( r_{a,1}(x) \) and \( r(x) = (2x + 3)/12 \) (see Example 1) as a function of the parameter \( \alpha \in [30, 230] \). The upper bound of approximation (orange curve) derived in Theorem 1(ii):

\[ \frac{1}{\alpha + 2} \left( 2 \| r' \| + \frac{1}{8} \| r'' \| \right) = \frac{1}{3(\alpha + 2)} \]

is plotted as well. We see that the approximation errors almost mimic the values of upper bound \((32 + 2)^{-1}\).

The plot in Fig. 9(b) shows the surface of the distance in \( sup \)-norm between \( r_{a,b} \) and \( r \) as a function of \( (\alpha, b) \) (see Example 4) for \( \alpha \in [30, 230] \) and \( b \in [1.1, 2.5] \). The best choices of parameters \( \alpha \) and \( b \) in our experiment are \( \alpha_0 = 228 \) and \( b_0 = 2.0 \), at which the minimum distance becomes 0.00796570.

Note that the approximates \( \mu_{a,m} \), \( \sigma_{a,m}^2 \) (\( m = 1, 2 \), introduced in Corollaries 1 and 2, as well as \( \mu_{a,b} \) and \( \sigma_{a,b}^2 \) always perform better if compared to \( r_{a,m} \), \( r_{a,b} \) and \( \sigma_{a,m}^2 \), \( \sigma_{a,b}^2 \).

The records of the distances in \( sup \)- and \( L_2 \)-norms of approximation errors for \( r_{a,2}, \mu_{a,2} \) (Example 2) as well as of \( r_n \) defined in (3) are presented in Table 1. Here we set up the number of used product moments equal to 30 and 60. In this table we used the values of \( r_n \) as given in [13]. The distances considered approximation (3) were evaluated using the Chebyshev’s nodes.

Table 1
The errors of approximates of \( r_{a,2}, \mu_{a,2}, \) and \( r_n \) in \( sup \)- and \( L_2 \)-norms.

<table>
<thead>
<tr>
<th>Approximation</th>
<th>( sup )-norm</th>
<th>( L_2 )-norm</th>
</tr>
</thead>
<tbody>
<tr>
<td>( r_{a,2} ) with ( \alpha = 30 )</td>
<td>0.0140185</td>
<td>0.0075029</td>
</tr>
<tr>
<td>( \mu_{a,2} ) with ( \alpha = 30 )</td>
<td>0.0055180</td>
<td>0.0032724</td>
</tr>
<tr>
<td>( r_n ) with ( n = 30 )</td>
<td>0.0510259</td>
<td>0.0039579</td>
</tr>
</tbody>
</table>

2.3. A comparative study

In this subsection we study the asymptotic behavior of errors (in \( sup \)-norm) for approximants \( r_{a,1}, r_{a,2}, \) and \( r_{a,b} \) as a function of parameters \( \alpha \) and \( b \) in the models considered in Examples 1, 2 and 4, respectively.

Blue dots in Fig. 9(a) display the distances in \( sup \)-norm between \( r_{a,1}(x) \) and \( r(x) = (2x + 3)/12 \) (see Example 1) as a function of the parameter \( \alpha \in [30, 230] \). The uniform rate of approximation (orange curve) derived in Theorem 1(ii):

\[ \frac{1}{\alpha + 2} \left( 2 \| r' \| + \frac{1}{8} \| r'' \| \right) = \frac{1}{3(\alpha + 2)} \]

is plotted as well. We see that the approximation errors almost mimic the values of upper bound \((3 \alpha + 2)^{-1}\).

The plot in Fig. 9(b) shows the surface of the distance in \( sup \)-norm between \( r_{a,b} \) and \( r \) as a function of \( (\alpha, b) \) (see Example 4) for \( \alpha \in [30, 230] \) and \( b \in [1.1, 2.5] \). The best choices of parameters \( \alpha \) and \( b \) in our experiment are \( \alpha_0 = 228 \) and \( b_0 = 2.0 \), at which the minimum distance becomes 0.00796570.

Note that the approximates \( \mu_{a,m} \), \( \sigma_{a,m}^2 \) (\( m = 1, 2 \), introduced in Corollaries 1 and 2, as well as \( \mu_{a,b} \) and \( \sigma_{a,b}^2 \) always perform better if compared to \( r_{a,m} \), \( r_{a,b} \) and \( \sigma_{a,m}^2 \), \( \sigma_{a,b}^2 \).

The records of the distances in \( sup \)- and \( L_2 \)-norms of approximation errors for \( r_{a,2}, \mu_{a,2} \) (Example 2) as well as of \( r_n \) defined in (3) are presented in Table 1. Here we set up the number of used product moments equal to 30 and 60. In this table we used the values of \( r_n \) as given in [13]. The distances considered approximation (3) were evaluated using the Chebyshev’s nodes.

From the results of Table 1 we see that constructions \( r_{a,2} \) and \( \mu_{a,2} \) if compared with \( r_n \) given by (3) perform much better in the terms of \( sup \)-norm when \( \alpha \) is small and equals 30.

3. Approximating the derivative function

Consider a random variable \( Y \sim G \) that has absolute continuous density \( g \). Assume also for simplicity \( supp[g] = [0, 1] \). In this section we recover \( g' \) given the moments of \( g \) up to order \( \alpha \):

\[ \mu_{\alpha}(j) = E(Y^j), \quad j = 1, 2, \ldots, \alpha. \]

Let us denote

\[ \mu_{\alpha}(j) = (-j) \mu_{\alpha}(j - 1), \quad j = 1, 2, \ldots, \alpha. \]

To recover \( g' \) consider \( \tilde{g}_{\alpha}^{-1} \mu_{\alpha} \) where \( \tilde{g}_{\alpha}^{-1} \) is defined in (5) with \( m_{\alpha} \) replaced by \( \mu_{\alpha} \) and \( \phi(x) = x \). Consider

\[ g_{\alpha}(y) := (\tilde{g}_{\alpha}^{-1} \mu_{\alpha})(y), \quad 0 \leq y \leq 1. \]
Theorem 3. (i) If \( g(1) = 0 \) and \( g' \) is continuous on \((0, 1)\), then \( g'_\alpha \to g' \) as \( \alpha \to \infty \) and
\[
\|g'_\alpha - g'\| \leq \Delta(g', \delta) + \frac{2\|g'\|}{\delta^2(\alpha + 2)}.
\]

(ii) If \( g^{(3)} \) is continuous on \((0, 1)\), then for each \( y \in (0, 1) \)
\[
g'_\alpha(y) - g'(y) = \frac{1}{\alpha + 2} \left[ (1 - 2y + [\alpha y] - \alpha y) g^{(2)}(y) + \frac{1}{2} (1 - y) g^{(3)}(y) \right] + o \left( \frac{1}{\alpha} \right)
\]
and
\[
\|g'_\alpha - g'\| \leq \frac{1}{\alpha + 2} \left[ 2 \|g^{(2)}\| + \frac{1}{8} \|g^{(3)}\| \right] + o \left( \frac{1}{\alpha} \right), \quad \text{as} \ \alpha \to \infty.
\]

Proof. Using the integration by parts and the condition \( g(1) = 0 \), we can easily obtain
\[
\mu_-(j) = -j \int_0^1 t^{j-1} g(t) \, dt = \int_0^1 t^j g'(t) \, dt.
\]
Now, substituting \( \mu_-(j) \) in the approximation \( g'_\alpha \) defined earlier, we arrive at
\[
g'_\alpha(y) = \int_0^1 \beta(u, [\alpha y] + 1, \alpha - [\alpha y] + 1) g(u) \, du.
\]
The rest of the proof is based on the properties of functions \( \beta(\cdot, [\alpha y] + 1, \alpha - [\alpha y] + 1) \), the continuity of function \( g' \) and \( g^{(3)} \), and repeats the steps from Theorem 1. \( \square \)

Corollary 3. If \( \tilde{\alpha} = 2\alpha \), \( \tilde{g}'(x) = 2 g'_\alpha(x) - g'_\alpha(x) \) and \( g^{(3)} \) is a continuous function, then
\[
\tilde{g}'(x) - g'(x) = \frac{1}{(\alpha + 1)(\alpha + 2)} \left[ (1 - 2x + (\alpha + 3)(\alpha x - \alpha x) g^{(2)}(x) + \frac{1}{2} x (1 - x) g^{(3)}(x) \right]
\]
\[
+ \frac{1}{2} \left( 2 \tilde{\Delta}_{2,\alpha}(x) - \Delta_{2,\alpha}(x) \right) g^{(3)}(x) + O \left( \frac{1}{\alpha^2} \right), \quad \text{as} \ \alpha \to \infty.
\]

Proof. The proof mimics the steps of Corollary 1. \( \square \)

Remark 2. The results of Theorems 1–3 can be extended and applied for approximations of the \( k \)th order derivatives of the regression and density functions. For example, in Model 1, let us assume for simplicity that \( r^{(0)}(1) = 0 \) for \( l = 0, 1, \ldots, k - 1 \), with \( r^{(0)} = r \). For each \( k = 1, \ldots \), denote
\[
\mu_-(j) = (-1)^k j (j - 1) \cdots (j - k + 1) m(1, j - k), \quad j = k, k + 1, \ldots.
\]

To approximate \( r^{(k)}(x) \), the following construction is proposed:
\[
r^{(k)}_{\alpha}(x) = (\alpha^{-1} \mu_-(x)), \quad x \in [0, 1].
\]

Now consider the problem of determining the value, say \( \alpha_0 \), that guarantees the approximation’s degree \( \epsilon \) in the regression case. Upper bounds derived in Theorem 1(ii) and Theorem 2(iii) involve unknown derivatives of \( r \), and cannot be evaluated in practice. To overcome this problem one can approximate these bounds by applying (19) with \( k = 1, 2 \). For instance, in Model 1, one can determine \( \alpha_0 \) from the equation
\[
\epsilon = \frac{1}{\alpha_0 + 2} \left[ 2 \|r'_{\alpha,1}\| + \frac{1}{8} \|r''_{\alpha,1}\| \right].
\]

Here, the approximates \( r'_{\alpha,1} \) and \( r''_{\alpha,1} \) should be evaluated with sufficiently large values of parameter \( \alpha \).

Example 5. Assume \( g \) represents the density of a \( B(a, b) \) distribution with \( a = b = 3 \). To recover \( g'(y) = 60y(1 - y)(1 - 2y) \), note that \( \mu_-(j) = (-j) B(a + j - 1, b) / B(a, b) \). Fig. 10 shows the approximated curves \( g'_\alpha \) (see plot (a) with \( \alpha = 50 \)) and \( g'_\alpha \) (see plot (b) with \( \alpha = 25 \)) and true \( g' \).

3.1. Multiplicative-censoring model

In this subsection we estimate the underlying df \( F \) and corresponding density function \( f \) in a model where we observe \( Y \sim g(y) \) given by equation
\[
g(y) = \int_0^y \frac{f(x)}{x} \, dx, \quad 0 < y \leq 1.
\]
Theorem 4. In the following statement assuming that $i, j \leq n$ respectively. From simulation study we found out that the asymptotic behavior of $\hat{f}_\alpha$ and $\hat{g}_\alpha$ is to estimate $\hat{df}$ and density $f$ given the sample $Y_1, \ldots, Y_n$ from $g$. Note that in this model $f(x) = -xg'(x)$. Using this relationship one can recover $f$ via

$$f_\alpha(x) = -xg'(x) \text{ as } \alpha \to \infty.$$ 

Hence, by denoting $\hat{g}_\alpha = \hat{g}_\alpha$ where the empirical counterpart $\hat{\mu}_-$ is used, we obtain the estimate of $f$ as,

$$\hat{f}_\alpha(x) = -\hat{g}_\alpha.$$ 

Here

$$\hat{\mu}_c(j) = \frac{1}{n} \sum_{i=1}^{n} Y_i.$$ 

Remark 3. It is worth mentioning that the moments of $f$ and $g$ are related in a very simple way:

$$E(X^n) = \mu_c(j)/E(U^n),$$ 

where $E(U^n) = 1/(j+1)$. Hence, to estimate $F$ and $f$ one can again apply Eqs. (1) and (2), where $\phi(x) = x$ and $m_\phi$ is replaced by

$$\hat{\mu}_c(j) = (j+1)\hat{\mu}_c(j).$$ 

Therefore the estimates of $F$ and $f$ can be written as follows:

$$\hat{F}_\alpha(x) := \left(\hat{g}_\alpha - \hat{\mu}_c\right)(x) \text{ and } \hat{f}_\alpha(x) := \left(\hat{g}_\alpha - \hat{\mu}_c\right)(x),$$ 

respectively. From simulation study we found out that the asymptotic behavior of $\hat{F}_\alpha$ and $\hat{f}_\alpha$ is very similar to each other. In the following statement assuming that $g(0)$ is finite, the weak consistency of $\hat{F}_\alpha$ and $\hat{f}_\alpha$ is provided.

Theorem 4. (i) $\hat{F}_\alpha(n) \to F(x) \text{ as } \sqrt{\alpha}/n \to 0 \text{ and } \alpha, n \to \infty$.

(ii) If $f$ is a continuous function, then $\hat{f}_\alpha(n) \to f(x) \text{ as } \alpha^{3/2}/n \to 0 \text{, } \alpha, n \to \infty$, uniformly on any interval $[a, b]$ with $0 < a, b < 1$.

Proof. (i). Since the operator $\hat{g}_\alpha$ is linear, we can rewrite $\hat{F}_\alpha$ as the sum

$$\hat{F}_\alpha(x) = \left(\hat{\mu}_c - \hat{\mu}_c\right)(x) + \left(\hat{g}_\alpha - \hat{\mu}_c\right)(x),$$

where $\hat{\mu}_c(j) = \hat{\mu}_c(j)$. Now, taking into account that

$$\sum_{k=0}^{\lfloor \alpha \rfloor} \left(\begin{array}{c} j \\ k \end{array}\right) (-1)^k = (-1)^{\lfloor \alpha \rfloor} \left(\begin{array}{c} j - 1 \\ \lfloor \alpha \rfloor \end{array}\right),$$

(see for example, [19]) and changing the sums in both terms of (20), after a simple algebra, we obtain the following representation

$$\hat{F}_\alpha(x) = \hat{g}_\alpha(x) - \hat{\mu}_c(x).$$
Here,

\[ \widehat{G}_{\alpha,n}(x) = \frac{1}{n} \sum_{i=1}^{n} B_{\alpha}(Y_i, x) \] with \( B_{\alpha}(t, x) = \sum_{k=0}^{[\alpha]} \binom{\alpha}{k} t^k (1-t)^{\alpha-k} \)

and

\[ \widehat{\eta}_{\alpha,n}(x) = \frac{1}{n} \sum_{i=1}^{n} Y_i \beta(Y_i, [\alpha x] + 1, \alpha - [\alpha x]) := \frac{1}{n} \sum_{i=1}^{n} \xi_i \]

with \( \beta (\cdot, c, d) \) to be the density function of a Beta \((c, d)\) distribution. Since, in the multiplicative-censoring model \( F(x) = G(x) - xg(x) \), it is sufficient to show that \( \widehat{G}_{\alpha,n}(x) \rightarrow G(x) \) and \( \widehat{\eta}_{\alpha,n}(x) \rightarrow xg(x) \) in probability at any continuity point \( x \) of \( G \) as \( \alpha, n \rightarrow \infty \). To complete the proof let us mention that the mean values of \( \widehat{G}_{\alpha,n}(x) \) and \( \widehat{\eta}_{\alpha,n}(x) \) converge uniformly to \( G(x) \) and \( xg(x) \), respectively. The later follows from the statements from [20] (see Theorem 3.1) and [15] (see v. II, Ch. VII), respectively.

For the variances we have

\[ \text{Var}(\widehat{G}_{\alpha,n}(x)) = \frac{1}{n} \text{Var}(B_{\alpha}(Y, x)) \leq \frac{1}{n} E(B_{\alpha}^2(Y, x)) \leq \frac{1}{n} \]

and

\[ \text{Var}(\widehat{\eta}_{\alpha,n}(x)) \leq \frac{1}{n} E(\xi_i^2) = \frac{1}{n} \int_0^1 y^2 \beta(y, [\alpha x] + 1, \alpha - [\alpha x]) g(y) \, dy \leq \frac{C_1 g(0) \sqrt{\alpha}}{n \sqrt{x(1-x)}} \] \hspace{1cm} (21)

In the last inequality of (21) we apply the following property of \( \beta (\cdot, c, d) \):

\[ \beta(y, [\alpha x] + 1, \alpha - [\alpha x]) \leq \frac{C_1 \sqrt{\alpha}}{\sqrt{x(1-x)}} , \quad x \in (0, 1) \] \hspace{1cm} (22)

for some positive constant \( C_1 \) (see [16]).

(ii) Note that \( g^c \) is continuous and \( E(\widehat{\eta}_{\alpha,n}) \rightarrow g^c \). According to Theorem 3(i) we have \( E(\widehat{\eta}_{\alpha,n}) \rightarrow g^c \) as \( \alpha, n \rightarrow \infty \). Now, let us prove that

\[ \text{Var}(\widehat{\eta}_{\alpha,n}(x)) \rightarrow 0 \quad \text{as} \quad \frac{\alpha^{3/2}}{n} \rightarrow 0. \] \hspace{1cm} (23)

After substitution of \( \widehat{\mu} \) into \( \frac{\alpha}{n} \widehat{\mu} \), and conducting a simple algebra, we derive:

\[ \widehat{G}_{\alpha,n}(x) = - \frac{\Gamma'(\alpha + 2)}{\Gamma(\alpha + 1)} \sum_{j=0}^{[\alpha x]} \frac{(-1)^j (j + [\alpha x]) (j + [\alpha x] - 1)}{j! (\alpha - [\alpha x] - j)!} \]

\[ = \frac{\alpha + 1}{n} \sum_{i=1}^{n} Y_i \left[ \frac{[\alpha x]}{[\alpha]} \right] \beta(Y_i, [\alpha x], \alpha - [\alpha x]) \]

\[ = \frac{\alpha + 1}{n} \sum_{i=1}^{n} \left( \frac{Y_i - [\alpha x]}{\alpha - [\alpha x]} \right) \beta(Y_i, [\alpha x], \alpha - [\alpha x]) \]

To prove (23), note that according to (22) the second moment of the summands in the last equation can be estimated from above as follows:

\[ \int \left( y - \frac{[\alpha x]}{\alpha} \right)^2 \beta^2(y, [\alpha x], \alpha - [\alpha x]) g(y) \, dy \leq \]

\[ = \frac{C_1 g(0) \sqrt{\alpha}}{\sqrt{x(1-x)}} \int \left( y - \frac{[\alpha x]}{\alpha} \right)^2 \beta(y, [\alpha x], \alpha - [\alpha x]) \, dy \]

\[ = \frac{C_1 g(0) \sqrt{\alpha}}{\sqrt{x(1-x)}} \frac{[\alpha x](\alpha - [\alpha x])}{\alpha^2(\alpha + 1)} , \quad x \in (0, 1). \] \hspace{1cm} [\Box]

**Example 5 (continued).** Assume \( X_i \sim \text{Beta}(3, 3) \), \( U_i \sim U(0, 1) \) with \( \{X_i\}_{i=1}^{n} \) and \( \{U_i\}_{i=1}^{n} \) to be mutually independent random samples. By taking \( Y_i = X_i U_i \), \( i = 1, \ldots, n \), we estimated \( F \) and \( f \) in the framework of multiplicative-censoring model. In our simulation study we set up several different sample sizes. We considered \( n \in \{500, 1000, 1500\} \) for graphical illustrations and \( n \in \{256, 512, 1024, 2048\} \) for evaluating the values of approximation errors in Table 2, respectively. In Fig. 11 we plotted the estimated df of \( X \) using \( F_{\alpha,n} \) when \( n = 500, \alpha = 35 \) and \( n = 1000, \alpha = 30 \), respectively. Fig. 12
provides estimated curves of $\hat{f}_{n,\alpha}(x)$ when $n = 1000$, $\alpha = 25$, and $n = 1500$, $\alpha = 30$, respectively. Our simulation study shows that proposed nonparametric estimates of $F$ and $f$ are consistent.

The question of consistent estimation of density $f$ has been studied by many authors, see, [21] and references therein. The further investigation of the rate of convergence will be conducted in a separate paper. Here, let us compare the $L_2$-errors of $f_{n,\alpha}$ with the ones constructed by means of wavelets. Define the average $L_2$-distance between estimate $f_{n,\alpha}$ and true $f$ as

$$d_{n,\alpha} := d(\hat{f}_{n,\alpha}, f) = \frac{1}{N} \sum_{r=1}^{N} \left( \frac{1}{\alpha} \sum_{j=1}^{\alpha} \left( \hat{f}_{n,\alpha}^{(r)} \left( \frac{j}{\alpha} \right) - f \left( \frac{j}{\alpha} \right) \right)^2 \right)^{1/2},$$

where $\hat{f}_{n,\alpha}^{(r)}$ denotes the value of $\hat{f}_{n,\alpha}$ on $r$th replication and $N$ is the number of replications. To compare our estimate $f_{n,\alpha}$ with the ones studied in [21], we took $N = 100$ and evaluated $d_{n,\alpha} := d(\hat{f}_{n,\alpha}, f)$. Table 2 also demonstrates the errors for estimates based on three different wavelet-type estimates borrowed from [21]. We see that our approximation behaves better when compared with the ones based on wavelets.
4. Conclusion

In this paper it is shown that the moment-determinate regression function $r$ as well as the conditional variance could be recovered with very high precision given the sequence of product moments (up to order $\alpha$) of the response and predictive variables. Utilizing this methodology, we show that finite order derivatives of $r$ and some moment-determinate density function $g$ can be recovered through the appropriately chosen sequence of moments.

Comparison of performances of our constructions based on Eq. (2) with the one given by Eq. (3) via numerical and graphical evaluations shows the superiority of the former one. The results are applied in the problem of estimating df and density function of underlying distribution in the framework of multiplicative-censoring model. In particular, the consistency in probability of proposed nonparametric estimates are established. The simulation results and the comparison with the other approximations based on the wavelets (see [21]) show that the asymptotic behavior of estimate $f_{n,\alpha}$ has a better performance in terms of average $L_2$-distance.

Acknowledgments

The authors are grateful to anonymous referees for useful suggestions and comments which led to a better presentation of the paper.

References