CENTRO PER LA FORMAZIONE IN ECONOMIA E POLITICA DELLO SVILUPPO RURALE
UNIVERSITÀ DI NAPOLI FEDERICO II - DIPARTIMENTO DI SCIENZE STATISTICHE
UNIVERSITÀ DI SALERNO - DIPARTIMENTO DI SCIENZE ECONOMICHE E STATISTICHE

Quaderni di STATISTICA

VOLUME 9 - 2007

LIGUORI EDITORE

Volume 9, anno 2007

ISSN 1594-3739 (edizione a stampa)

Registrazione al n. 5264 del 6/12/2001 presso il Tribunale di Napoli
ISBN-13 **978 - 88 - 207 - 4209 - 6**

Direttore responsabile: Gennaro Piccolo

Prima edizione italiana Dicembre 2007
Finito di stampare in Italia nel mese di Dicembre 2007 da OGL - Napoli

Il C.F.E.P.S.R. si avvale per la stampa dei Quaderni di Statistica del contributo dell'Istituto Banco di Napoli - Fondazione.

La carta utilizzata per la stampa di questo volume è inalterabile, priva di acidi, a PH neutro, conforme alle norme UNI EN Iso 9706 ×, realizzata con materie prime fibrose vergini provenienti da piantagioni rinnovabili e prodotti ausiliari assolutamente naturali, non inquinanti e totalmente biodegradabili.

Indice

FORUM

Quaderni di Statistica
Vol. 9, 2007

On variance reduction in some Bernstein-type approximations

Andrea Pallini
Dipartimento di Statistica e Matematica Applicata all'Economia, Università di Pisa,
E-mail: apallini@ec.unipi.it

Summary: A linear adjustment for more efficient versions of the Bernstein-type estimators proposed in Pallini (2005) is asymptotically studied, as the sample size becomes increasingly large. The Bernstein-type approximations in Pallini (2005) originate from the well-known definitions of univariate and multivariate Bernstein polynomials in the theory of approximation of continuous functions. The Bernstein-type approximations are defined through a constructive coefficient that conveniently characterizes the velocity of their uniform convergence. The statistical estimation of a smooth function of population means is considered, with conclusions about the asymptotic behaviour of the linear adjustment, as the sample size becomes increasingly large. Monte Carlo simulations on the ratio of means example are detailed and discussed.

Keywords: Bernstein Polynomial, Smooth Function of Means, Variance Reduction.

1. Introduction

The principles behind the Bernstein polynomials are typically read about the approximation of continuous functions that makes use of the binomial random variable. The Bernstein polynomials are regarded as very important tools in the theory of approximation of functions that are defined on a closed interval. From a historical point of view, the Bernstein polynomials were introduced in order to provide a simple proof of the Weierstrass approximation theorem, see Korovkin (1960), chapter 1, Davis (1963), chapter 6, Feller (1971), chapter 7, Lorentz (1986), chapter 1, Pinkus (2000), and Phillips (2003), chapter 7.

On the other hand, the uniform convergence of the Bernstein polynomials to a target function is known to be fairly slow. Better Bernstein-type approximations, based on the binomial and the multivariate binomial distributions, have been recently proposed in Pallini (2005), with conclusions about their uniform convergence. These Bernstein-type approximations can be defined by making a convenient choice of the real value for a constructive coefficient. The constructive coefficient defines different kernels from the integer values of the binomial random variables. The convergence of these Bernstein-type approximations can generally be set, through the constructive coefficient, in order to be faster than the convergence of the basic Bernstein polynomials.

Here, we want to study a linear adjustment for variance reduction in statistical versions of the Bernstein-type approximations of Pallini (2005), as the sample size becomes increasingly large. Specifically, in section 2, we overview the main features of the Bernstein-type approximations of Pallini (2005). In section 3, we describe and study the Bernstein-type estimators of smooth functions of means, as the sample size becomes increasingly large. In section 4, we study a linear adjustment for variance reduction in the classical estimation of smooth functions of population means, as the sample size becomes increasingly large. In section 5, we show the efficiency, and the effectiveness, of the Bernstein-type estimators of smooth functions of means, by a simulation study on the square of mean example, on the cube of mean example, on the ratio of means example.

2. Bernstein-type approximations

Let P_m be the space of polynomials $P(x)$ of degree at most m, for all real numbers x. Let g be a bounded, continuous, and real-valued function that is defined on the closed interval $[0,1]$. The Bernstein-type approximation $B_m^{(s)}(g;x)$ of order m for the function $g(x)$ is defined as

$$B_m^{(s)}(g;x) = \sum_{v=0}^{m} g(m^{-s}(m^{-1}v - x) + x)\binom{m}{v} x^v (1-x)^{m-v} , \qquad (1)$$

where $s > -1/2$ is a constructive coefficient, m is a positive integer, and $x \in [0,1]$. See Pallini (2005).

It is seen that $B_m^{(s)}(g;x) \in \mathrm{P}_m$, where $s > -1/2$, for every $x \in [0,1]$. See Pallini (2005).

The Bernstein-type approximation $B_m^{(s)}(g;x)$ is defined in (1) by the binomial distribution $Bi(m;x) = \binom{m}{v} x^v (1-x)^{m-v}$, that takes on the integer values $v = 0,1,\ldots,m$, for every $x \in [0,1]$.

The Bernstein polynomial $B_m(g;x)$ can be obtained as the Bernstein-type approximation $B_m^{(0)}(g;x)$, by setting $s = 0$ in the definition (1) of $B_m^{(s)}(g;x)$, for every $x \in [0,1]$.

Under the condition $s > -1/2$, where the constructive coefficient s is fixed, the Bernstein-type approximation $B_m^{(s)}(g;x)$ converges to $g(x)$, $B_m^{(s)}(g;x) \to g(x)$, as $m \to \infty$, uniformly at any point $x \in [0,1]$. See Pallini (2005), Appendix 8.2, therein.

The Bernstein-type approximation $B_m^{(s)}(g;x)$, given by (1), improves on its numerical performance, as the constructive coefficient s increases, for every point $x \in [0,1]$.

Let g be a bounded, continuous, and real-valued function that is defined on the closed k-dimensional cube $[0,1]^k$. We let $\mathrm{x} = (x_1,\ldots,x_k)^T$, where $\mathrm{x} \in [0,1]^k$. The multivariate Bernstein-type approximation $B_{\mathrm{m}}^{(s)}(g;\mathrm{x})$ for the function $g(\mathrm{x})$ is defined as

$$B_{\mathrm{m}}^{(s)}(g;\mathrm{x}) = \sum_{v_1=0}^{m_1} \cdots \sum_{v_k=0}^{m_k} g\begin{pmatrix} m_1^{-s}(m_1^{-1}v_1 - x_1) + x_1 \\ \vdots \\ m_k^{-s}(m_k^{-1}v_k - x_k) + x_k \end{pmatrix}$$

$$\cdot \binom{m_1}{v_1} \cdots \binom{m_k}{v_k} x_1^{v_1}(1-x_1)^{m_1-v_1} \cdots x_k^{v_k}(1-x_k)^{m_k-v_k}, \qquad (2)$$

where $s > -1/2$ is a constructive coefficient, $\mathrm{m} = (m_1, \ldots, m_k)^T$ are positive integers, and $\mathrm{x} \in [0,1]^k$. See Pallini (2005).

It is seen that the multivariate Bernstein-type approximation $B_{\mathrm{m}}^{(s)}(g;\mathrm{x}) \in \mathrm{P}_m$, where $m = \sum_{i=1}^{k} m_i$ is the global degree in $B_{\mathrm{m}}^{(s)}(g;\mathrm{x})$, P_m is the space of polynomials $P(\mathrm{x})$ of degree at most m, and $s > -1/2$, for every $\mathrm{x} \in [0,1]^k$. See Pallini (2005).

The multivariate Bernstein-type approximation $B_{\mathrm{m}}^{(s)}(g;\mathrm{x})$ is defined in (2) by a multivariate binomial distribution $Bi(\mathrm{m};\mathrm{x}) = \prod_{i=1}^{k} Bi(m_i;x_i)$ that is obtained as the product of k independent binomial distributions $Bi(m_1;x_1),\ldots,Bi(m_k;x_k)$. The distribution $Bi(\mathrm{m};\mathrm{x})$ takes values on the Cartesian product $\{0,1,\ldots,m_1\} \times \cdots \times \{0,1,\ldots,m_k\}$, for every $\mathrm{x} \in [0,1]^k$.

The multivariate Bernstein polynomial $B_{\mathrm{m}}(g;\mathrm{x})$ can be obtained as the Bernstein-type approximation $B_{\mathrm{m}}^{(0)}(g;\mathrm{x})$, by setting $s = 0$ in the definition (2) of $B_{\mathrm{m}}^{(s)}(g;\mathrm{x})$, for every $\mathrm{x} \in [0,1]^k$.

The multivariate Bernstein-type approximation $B_{\mathrm{m}}^{(s)}(g;\mathrm{x})$ converges to $g(\mathrm{x})$ uniformly, where $s > -1/2$ is fixed, at any k-dimensional point of continuity $\mathrm{x} \in [0,1]^k$, as $m_i \to \infty$, where $i = 1,\ldots,k$. See Pallini (2005), Appendix 8.2, therein.

The Bernstein-type approximation $B_{\mathrm{m}}^{(s)}(g;\mathrm{x})$, given by (2), improves on its numerical performance, as the constructive coefficient s increases, for every point $\mathrm{x} \in [0,1]^k$.

3. Estimation of smooth functions of means

The Bernstein-type approximations $B_m^{(s)}(g;x)$ and $B_{\mathrm{m}}^{(\mathrm{s})}(g;\mathrm{x})$, given by (1) and (2), respectively, where $s > -1/2$, $x \in [0,1]$ and $\mathrm{x} \in [0,1]^k$, can be used for an efficient estimation of smooth functions of the population means, on a random sample of n independent and identically distributed (i.i.d.) observations of a random variable (r.v.).

Let X be a r.v., with values $x \in [0,1]$, with population distribution function F and finite mean $\mu = E[X]$.

We want to estimate a population characteristic $g(\mu)$, where g is a smooth function, $g:[0,1] \to R^1$. The natural estimator of $g(\mu)$ is $g(\bar{x})$, where $\bar{x} = n^{-1}\sum_{j=1}^{n} X_j$ is the sample mean, calculated on a random sample of n i.i.d. observations X_j of the r.v. X, where $j = 1,\dots,n$. See Hall (1992), chapter 2.

An alternative estimator of $g(\mu)$ can be obtained as the Bernstein-type estimator $B_m^{(s)}(g;\bar{x})$ defined as

$$B_m^{(s)}(g;\bar{x}) = \sum_{v=0}^{m} g(m^{-s}(m^{-1}v - \bar{x}) + \bar{x}) \binom{m}{v} \bar{x}^{v}(1-\bar{x})^{m-v}, \tag{3}$$

where $s > -1/2$. See Pallini (2005).

The Bernstein-type estimator (3) follows from the definition (1) of $B_m^{(s)}(g;x)$, by substituting the argument $x \in [0,1]$ with the sample mean $\bar{x}$, where $\bar{x}$ ranges in $[0,1]$.

Let X be a k-variate r.v., with values $\mathrm{x} \in [0,1]^k$, where $\mathrm{X} = (X_1,\dots,X_k)^T$, with k-variate distribution function F and finite k-variate mean $\mu = E[X]$, where $\mu = (\mu_1,\dots,\mu_k)^T$.

We want to estimate $g(\mu)$, where $g:[0,1]^k \to R^1$. The natural estimator of $g(\mu)$ is $g(\bar{\mathrm{x}})$, where $\bar{\mathrm{x}} = (\bar{x}_1,\dots,\bar{x}_k)^T$ is the k-variate sample mean, with components $\bar{x}_i = n^{-1}\sum_{j=1}^{n} X_{ij}$, where $i = 1,\ldots,k$, on a random sample of n i.i.d. observations $\mathrm{X_j} = (X_{1j},\ldots,X_{kj})^T$ of the k-variate r.v. X, where $j = 1,\ldots,n$. See Hall (1992), chapter 2.

An alternative estimator of $g(\mu)$ can be obtained as the Bernstein-type estimator $B_{\mathrm{m}}^{(s)}(g;\bar{\mathrm{x}})$ defined as

$$B_{\mathrm{m}}^{(s)}(g;\bar{\mathrm{x}}) = \sum_{v_1=0}^{m_1}\cdots\sum_{v_k=0}^{m_k} g\begin{pmatrix} m_1^{-s}(m_1^{-1}v_1 - \bar{x}_1) + \bar{x}_1 \\ \vdots \\ m_k^{-s}(m_k^{-1}v_k - \bar{x}_k) + \bar{x}_k \end{pmatrix}\binom{m_1}{v_1}\cdots\binom{m_k}{v_k}$$
$$\cdot\, \bar{x}_1^{v_1}(1-\bar{x}_1)^{m_1-v_1}\cdots\bar{x}_k^{v_k}(1-\bar{x}_k)^{m_k-v_k}, \qquad (4)$$

where $s > -1/2$, and $\mathrm{m} = (m_1,\ldots,m_k)^T$. See Pallini (2005).

The Bernstein-type estimator (4) follows from the definition (2) of $B_{\mathrm{m}}^{(s)}(g;\mathrm{x})$, by substituting the argument $\mathrm{x} \in [0,1]^k$ with the k-variate sample mean $\bar{\mathrm{x}}$, where $\bar{\mathrm{x}}$ ranges in $[0,1]^k$.

4. Variance reduction

We denote by V the binomial $(m;\bar{x})$ r.v. that takes on the integer values $v = 0,1,\ldots,m$, with probability $Bi(m;\bar{x})$.

In the definition (3) of the Bernstein-type estimator $B_m^{(s)}(g;\bar{x})$, the sample mean $\bar{x}$ and the sample size n can be regarded as fixed with respect to $Bi(m;\bar{x})$.

The r.v. V has first moment about zero $E_m[V] = m\bar{x}$, where E_m is the expected value with respect the binomial $(m;\bar{x})$ distribution. The r.v. V has the first three central moments $E_m[V - m\bar{x}] = 0$, $E_m[(V - m\bar{x})^2] = m\bar{x}(1-\bar{x})$, and $E_m[(V - m\bar{x})^3] = m\bar{x}(1-\bar{x})(1-2\bar{x})$.

In the definition (3) of the Bernstein-type estimator $B_m^{(s)}(g;\bar{x})$, for every $v = 0,1,\ldots,m$, the Taylor expansion of the smooth function g of $m^{-s}(m^{-1}v-\bar{x})+\bar{x}$, around $\bar{x}$, yields

$$\begin{aligned}
B_m^{(s)}(g;\bar{x}) &= \sum_{v=0}^{m} g(m^{-s}(m^{-1}v-\bar{x})+\bar{x})\binom{m}{v}\bar{x}^{v}(1-\bar{x})^{m-v} \\
&= g(\bar{x}) + g'(\bar{x})\sum_{v=0}^{m} m^{-s}(m^{-1}v-\bar{x})\binom{m}{v}\bar{x}^{v}(1-\bar{x})^{m-v} \\
&+\frac{1}{2}g''(\bar{x})\sum_{v=0}^{m} m^{-2s}(m^{-1}v-\bar{x})^{2}\binom{m}{v}\bar{x}^{v}(1-\bar{x})^{m-v} \\
&+\frac{1}{6}g'''(\bar{x})\sum_{v=0}^{m} m^{-3s}(m^{-1}v-\bar{x})^{3}\binom{m}{v}\bar{x}^{v}(1-\bar{x})^{m-v} \\
&+O(m^{-4s-3}),
\end{aligned}$$

as $m \to \infty$, with n and $\bar{x}$ fixed. Thus, we have

$$\begin{aligned}
B_m^{(s)}(g;\bar{x}) &= g(\bar{x}) \\
&+\frac{1}{2}g''(\bar{x})m^{-2s-1}\bar{x}(1-\bar{x}) \\
&+\frac{1}{6}g'''(\bar{x})m^{-3s-2}\bar{x}(1-\bar{x})(1-2\bar{x})+O(m^{-4s-3}) \\
&= g(\bar{x})+O(m^{-2s-1}),
\end{aligned} \tag{5}$$

as $m \to \infty$, with n and $\bar{x}$ fixed.

Hereafter, we compare estimators that are defined by Bernstein-like schemes around an observed sample of random observations, that are based on the artificial integer m. The size n, the observations in the sample, and the sample mean $\bar{x}$, are taken as fixed.

From (3) and (5), an adjusted Bernstein-type estimator $C_m^{(s)}(g;\bar{x})$ can be defined as

$$C_m^{(s)}(g;\bar{x}) = B_m^{(s)}(g;\bar{x}) - \frac{1}{2} g''(\bar{x}) m^{-2s-1} \bar{x}(1-\bar{x}) . \tag{6}$$

Comparing (6) with (3), the Bernstein-type estimator $C_m^{(s)}(g;\bar{x})$ is obtained from $B_m^{(s)}(g;\bar{x})$ by a linear adjustment in the Taylor expansion that produces (5). The linear adjustment uses the value $g''(\bar{x})$ of the second derivative g'' of the smooth function g, at the known point $\bar{x}$, when we aim at estimating the population value $g(\mu)$. From (5), it follows that we have the approximation

$$\begin{aligned} C_m^{(s)}(g;\bar{x}) &= g(\bar{x}) + \frac{1}{6} g'''(\bar{x}) m^{-3s-2} \bar{x}(1-\bar{x})(1-2\bar{x}) \\ &= g(\bar{x}) + O(m^{-3s-2}) , \end{aligned} \tag{7}$$

as $m \to \infty$, with n and $\bar{x}$ fixed.

We can observe that $B_m^{(s)}(g;\bar{x}) - g(\bar{x}) = O(m^{-2s-1})$, as $m \to \infty$, and $C_m^{(s)}(g;\bar{x}) - g(\bar{x}) = O(m^{-3s-2})$, as $m \to \infty$, with n and $\bar{x}$ fixed.

We denote by σ^2 the asymptotic variance, with respect to the population distribution F, defined as

$$\sigma^2 = (g'(\mu))^2 E[(X-\mu)^2] .$$

An asymptotic conclusion in statistical inference is that $g(\bar{x}) = g(\mu) + O_p(n^{-1/2})$, as the size n of the random sample, drawn from the population distribution F, becomes increasingly large, namely, as $n \to \infty$.

From the definitions (3) and (6), of the Bernstein-type estimator $B_m^{(s)}(g;\bar{x})$ and the adjusted estimator $C_m^{(s)}(g;\bar{x})$, and (7), it follows that

$$B_m^{(s)}(g;\bar{x}) - g(\mu) = O(m^{-2s-1}) + O_p(n^{-1/2}) ,$$

as $m \to \infty$, and $n \to \infty$, and

$$C_m^{(s)}(g;\bar{x}) - g(\mu) = O(m^{-3s-2}) + O_p(n^{-1/2}),$$

as $m \to \infty$, and $n \to \infty$. Finally, we have the asymptotic variance

$$VAR[B_m^{(s)}(g;\bar{x})] = O(m^{-2(2s+1)}) + n^{-1}\sigma^2 + O(m^{-2s-1}n^{-1}), \tag{8}$$

as $m \to \infty$ and $n \to \infty$, and the asymptotic variance

$$VAR[C_m^{(s)}(g;\bar{x})] = O(m^{-2(3s+2)}) + n^{-1}\sigma^2 + O(m^{-3s-2}n^{-1}), \tag{9}$$

as $m \to \infty$ and $n \to \infty$.

The asymptotic variances $VAR[B_m^{(s)}(g;\bar{x})]$ and $VAR[C_m^{(s)}(g;\bar{x})]$, given by (8) and (9), can exhibit interesting efficiencies in the statistical estimation of a smooth function g of the population mean μ, as $m \to \infty$ and $n \to \infty$.

Comparing (9) with (8), we can prefer $VAR[C_m^{(s)}(g;\bar{x})]$ to $VAR[B_m^{(s)}(g;\bar{x})]$, and thus the adjusted estimator $C_m^{(s)}(g;\bar{x})$, given by (6), to the Bernstein-type estimator $B_m^{(s)}(g;\bar{x})$, given by (3), as $m \to \infty$ and $n \to \infty$.

We denote the rth derivative $\partial^r / \partial^{i_1} \cdots \partial^{i_r} g(\mathrm{x})$ of the smooth function $g:[0,1]^k \to R^1$ by $g_{i_1 \cdots i_r}(\mathrm{x})$, where $r = i_1 + \cdots + i_r$, and $\mathrm{x} \in [0,1]^k$.

Following the definition (6), we can obtain the adjusted estimator $C_{\mathrm{m}}^{(\mathrm{s})}(g;\bar{\mathrm{x}})$ from the Bernstein-type estimator $B_{\mathrm{m}}^{(\mathrm{s})}(g;\bar{\mathrm{x}})$, given by (4). In particular, we have

$$C_{\mathrm{m}}^{(s)}(g;\bar{\mathrm{x}}) = B_{\mathrm{m}}^{(s)}(g;\bar{\mathrm{x}}) - \frac{1}{2}\sum_{i=1}^{k} g_{ii}(\bar{\mathrm{x}}) m_i^{-2s-1} \bar{x}_i (1-\bar{x}_i). \tag{10}$$

From (10), it follows that

$$C_{\mathrm{m}}^{(s)}(g;\overline{\mathrm{x}}) = g(\overline{\mathrm{x}}) + \frac{1}{6}\sum_{i=1}^{k} g_{iii}(\overline{\mathrm{x}})\, m_i^{-3s-2}\overline{x}_i(1-\overline{x}_i)(1-2\overline{x}_i)$$
$$= g(\overline{\mathrm{x}}) + \sum_{i=1}^{k} O(m_i^{-3s-2}), \qquad (11)$$

as $m_i \to \infty$, where $i = 1,\ldots,k$, and $n \to \infty$, with n and $\overline{\mathrm{x}}$ fixed.

We can observe that $B_{\mathrm{m}}^{(s)}(g;\overline{\mathrm{x}}) - g(\overline{\mathrm{x}}) = \sum_{i=1}^{k} O(m_i^{-2s-1})$, as $m_i \to \infty$, where $i = 1,\ldots,k$, and $C_{\mathrm{m}}^{(s)}(g;\overline{\mathrm{x}}) - g(\overline{\mathrm{x}}) = \sum_{i=1}^{k} O(m_i^{-3s-2})$, as $m_i \to \infty$, where $i = 1,\ldots,k$, and $n \to \infty$, with n and $\overline{\mathrm{x}}$ fixed.

We denote by σ^2 the asymptotic variance, with respect to the k-variate population distribution F, defined as

$$\sigma^2 = \sum_{i_1=1}^{k}\sum_{i_2=1}^{k} g_{i_1}(\mu) g_{i_2}(\mu) E[(X_{i_1} - \mu_{i_1})(X_{i_2} - \mu_{i_2})].$$

An asymptotic conclusion in multivariate statistical inference is that $g(\overline{\mathrm{x}}) = g(\mu) + O_p(n^{-1/2})$, where $\overline{\mathrm{x}} = (x_1,\ldots,x_k)^T$ and $\mu = (\mu_1,\ldots,\mu_k)^T$, as the size n of the random sample, drawn from the k–variate distribution F, becomes increasingly large, namely, as $n \to \infty$.

From the definitions (4) and (10) of the Bernstein-type estimator $B_{\mathrm{m}}^{(s)}(g;\overline{\mathrm{x}})$ and the adjusted estimator $C_{\mathrm{m}}^{(s)}(g;\overline{\mathrm{x}})$, and (11), it follows that

$$B_{\mathrm{m}}^{(s)}(g;\overline{\mathrm{x}}) - g(\mu) = \sum_{i=1}^{k} O(m_i^{-2s-1}) + O_p(n^{-1/2}),$$

as $m_i \to \infty$, where $i = 1,\ldots,k$, and $n \to \infty$, and

$$C_{\mathrm{m}}^{(s)}(g;\overline{\mathrm{x}}) - g(\mu) = \sum_{i=1}^{k} O(m_i^{-3s-2}) + O_p(n^{-1/2}),$$

as $m_i \to \infty$, where $i = 1,\ldots,k$, and $n \to \infty$. Finally, we have

$$VAR[B_{\mathrm{m}}^{(s)}(g;\overline{\mathrm{x}})\,] = \sum_{i=1}^{k} O(m_i^{-2(2s+1)}) + n^{-1}\sigma^2 + \sum_{i=1}^{k} O(m_i^{-2s-1}n^{-1}), \qquad (12)$$

as $m_i \to \infty$, where $i = 1,\ldots,k$, and $n \to \infty$, and

$$VAR[C_{\mathrm{m}}^{(s)}(g;\overline{\mathrm{x}})\,] = \sum_{i=1}^{k} O(m_i^{-2(3s+2)}) + n^{-1}\sigma^2 + \sum_{i=1}^{k} O(m_i^{-3s-2}n^{-1}), \qquad (13)$$

as $m_i \to \infty$, where $i = 1,\ldots,k$, and $n \to \infty$.

The asymptotic variances $VAR[B_{\mathrm{m}}^{(s)}(g;\overline{\mathrm{x}})\,]$ and $VAR[C_{\mathrm{m}}^{(s)}(g;\overline{\mathrm{x}})\,]$, given by (12) and (13), can exhibit interesting efficiencies in the statistical estimation of a smooth function g of the population k-variate mean μ, as $m_i \to \infty$, where $i = 1,\ldots,k$, and $n \to \infty$.

Comparing (13) with (12), we can prefer the adjusted estimator $C_{\mathrm{m}}^{(s)}(g;\overline{\mathrm{x}})$, given by (10), to the Bernstein-type estimator $B_{\mathrm{m}}^{(s)}(g;\overline{\mathrm{x}})$, given by (4), as $m_i \to \infty$, where $i = 1,\ldots,k$, and $n \to \infty$.

5. Monte Carlo simulations

Let X be a univariate r.v. with real values x in the bounded domain $(0,\beta)$, where $\beta < +\infty$, $x \in (0,\beta)$, with univariate distribution function F and finite univariate mean $\mu = E[X]$. We have a random sample of n random observations X_j, where $j = 1,\ldots,n$. We want to estimate the smooth function of means $g(\mu) = \mu^2$. The natural estimator of $g(\mu)$ is $g(\overline{x}) = \overline{x}^2$.

Following the results in Lorentz (1986), chapter 2, the definition (3) of the Bernstein-type estimator $B_m^{(s)}(g;\bar{x})$ and the definition (6) of the adjusted estimator $C_m^{(s)}(g;\bar{x})$ can be extended to definitions on the bounded cube $x \in (0,\beta)$, where $\beta^{-1}x \in (0,1)$.

From the definition (3), it follows that the Bernstein-type estimator $B_{\mathrm{m}}^{(s)}(g;\bar{\mathrm{x}})$ of $g(\mu)=\mu^2$ can be defined as

$$B_m^{(s)}(g;\bar{x}) = \sum_{v=0}^{m} \left\{ m^{-s} (m^{-1}\beta v - \bar{x}) \right\}^2 \binom{m_1}{v_1} (\beta^{-1}\bar{x}_1)^{v_1} (1-\beta^{-1}\bar{x}_1)^{m_1 - v_1} , \quad (14)$$

where one can write $\beta^{-1}\bar{x} \in (0,1)$.

From the definition (6), it follows that the adjusted Bernstein-type estimator $C_{\mathrm{m}}^{(s)}(g;\bar{\mathrm{x}})$ of $g(\mu)=\mu^2$ is

$$C_m^{(s)}(g;\bar{x}) = B_m^{(s)}(g;\bar{x}) - m^{-2s-1}\beta^2 \bar{x}^2 (1-\bar{x}) . \quad (15)$$

An empirical study on the Monte Carlo simulation of random samples from the folded normal $|N(0,1)|$ distribution, with sizes from $n=10$ to $n=100$, shows that the technique can not be suggested with the $g(\mu)=\mu^2$ example. We had a minimum difference, between the Monte Carlo variance and its corrected version, that equals -0.0014272, and a maximum difference that equals -0.0013567. A similar situation may occur with the $g(\mu)=\mu^3$ example. We had a minimum difference between the variances that equals -0.0000064 and a maximum difference between the variances that equals -0.0000061.

The proposed variance reduction technique can be regarded as competitive for smooth functions of means g defined on a multivariate domain, $g:[0,1]^k \to R^1$. We recall that the natural estimator of $g(\mu)$, where $\mu=(\mu_1,\ldots,\mu_k)^T$, is $g(\bar{\mathrm{x}})$, where $\bar{\mathrm{x}}=(\bar{x}_1,\ldots,\bar{x}_k)^T$.

Let $\mathrm{X} = (X_1, X_2)^T$ be a bivariate r.v. with real values $\mathrm{x} \in (x_1, x_2)^T$ in the bounded domain $(0, \beta)^2$, where $\beta < +\infty$, $\mathrm{x} \in (0, \beta)^2$, with bivariate distribution function F and finite bivariate mean $\mu = E[X]$, $\mu = (\mu_1, \mu_2)^T$. We have a random sample of n random observations $\mathrm{X}_j = (\mathrm{X}_{1j}, \mathrm{X}_{2j})^T$, where $j = 1, \ldots, n$. We want to estimate the population ratio of means $g(\mu) = \mu_1 / \mu_2$. The natural estimator of $g(\mu)$ is $g(\overline{\mathrm{x}}) = \overline{x}_1 / \overline{x}_2$, where $\overline{\mathrm{x}} = (\overline{x}_1, \overline{x}_2)^T$ is the bivariate sample mean, with components $\overline{x}_i = n^{-1} \sum_{j=1}^{n} X_{ij}$, where $i = 1,2$.

Following the results in Lorentz (1986), chapter 2, and following the definition (4), the Bernstein-type estimator $B_{\mathrm{m}}^{(s)}(g; \overline{\mathrm{x}})$ of $g(\mu) = \mu_1 / \mu_2$ can be defined as

$$B_{\mathrm{m}}^{(s)}(g; \overline{\mathrm{x}}) = \sum_{v_1=0}^{m_1} \sum_{v_2=0}^{m_2} \frac{m_1^{-s}(m_1^{-1}\beta v_1 - \overline{x}_1) + \overline{x}_1}{m_2^{-s}(m_2^{-1}\beta v_2 - \overline{x}_2) + \overline{x}_2} \binom{m_1}{v_1} \binom{m_2}{v_2}$$
$$\cdot (\beta^{-1}\overline{x}_1)^{v_1} (1 - \beta^{-1}\overline{x}_1)^{m_1 - v_1} (\beta^{-1}\overline{x}_2)^{v_2} (1 - \beta^{-1}\overline{x}_2)^{m_2 - v_2}, \qquad (16)$$

where $\overline{\mathrm{x}} \in (0, \beta)^2$.

From the definition (10), it follows that the adjusted Bernstein-type estimator $C_{\mathrm{m}}^{(s)}(g; \overline{x})$ of $g(\mu) = \mu_1 / \mu_2$ is

$$C_{\mathrm{m}}^{(s)}(g; \overline{x}) = B_{\mathrm{m}}^{(s)}(g; \overline{x}) - m_2^{-2s-1} \beta^2 \, \overline{x}_2^2 (1 - \overline{x}_2). \qquad (17)$$

We report on a empirical study on this ratio of means example. Simulated random samples were always obtained from a bivariate r.v. with the same distribution F, the folded normal distribution, with the non-negative components $X_1 = |W_1|$ and $X_2 = |W_2|$, in $\mathrm{X} = (X_1, X_2)^T$, where W_1 and W_2 denote stochastically independent normal (0,1) r.v.'s. All these simulated random samples contained i.i.d. observations from the folded normal distribution F and were independently generated by the classical Monte Carlo simulation. See Becker, Chambers and Wilks (1988).

In figures 1 to 4, we compare the Monte Carlo measure of discrepancy $DI_M(B_{\mathrm{m}}^{(s)}(g;\overline{\mathrm{x}});g(\overline{\mathrm{x}}))$ that was defined as

$$DI_M(B_{\mathrm{m}}^{(s)}(g;\overline{\mathrm{x}});g(\overline{\mathrm{x}}))=(M-1)^{-1}\sum_{l=1}^{M}(B_{\mathrm{m}}^{(s)}(g;\overline{\mathrm{x}})-g(\overline{\mathrm{x}}))^2 ,$$

with the Monte Carlo measure of discrepancy $DI_M(C_{\mathrm{m}}^{(s)}(g;\overline{\mathrm{x}});g(\overline{\mathrm{x}}))$, for the Bernstein-type estimator $B_{\mathrm{m}}^{(s)}(g;\overline{\mathrm{x}})$ and the adjusted estimator $C_{\mathrm{m}}^{(s)}(g;\overline{\mathrm{x}})$, given by (16) and (17), respectively. Bernstein-type estimators $B_{\mathrm{m}}^{(s)}(g;\overline{\mathrm{x}})$ and $C_{\mathrm{m}}^{(s)}(g;\overline{\mathrm{x}})$ were viewed as approximations to the sample smooth function $g(\overline{\mathrm{x}})=\overline{x}_1/\overline{x}_2$. The Monte Carlo measures of discrepancy were calculated on M independent samples, of various sizes n, for different values of the simulation parameters, $\mathrm{m}=(m_1,m_2)^T$ and $s>-1/2$. We always considered the smooth function $g(\overline{\mathrm{x}})=\overline{x}_1/\overline{x}_2$ on $\overline{\mathrm{x}}\in(0,\beta)^2$, where $\beta=3$.

In figures 5 to 8, we compare the efficiencies of the Bernstein-type estimator $B_{\mathrm{m}}^{(s)}(g;\overline{\mathrm{x}})$ and the adjusted estimator $C_{\mathrm{m}}^{(s)}(g;\overline{\mathrm{x}})$, given by (16) and (17), respectively, in the classical estimation for the population ratio of means $g(\mu)=\mu_1/\mu_2$. We calculated the values of the Monte Carlo variance $VAR_M(B_{\mathrm{m}}^{(s)}(g;\overline{\mathrm{x}});g(\mu))$ that was defined as

$$VAR_M(B_{\mathrm{m}}^{(s)}(g;\overline{\mathrm{x}});g(\mu))=(M-1)^{-1}\sum_{l=1}^{M}(B_{\mathrm{m}}^{(s)}(g;\overline{\mathrm{x}})-g(\mu))^2 ,$$

and the values of the Monte Carlo variance $VAR_M(C_{\mathrm{m}}^{(s)}(g;\overline{\mathrm{x}});g(\mu))$, on M independent samples, of various sizes n, for different values of the simulation parameters, $\mathrm{m}=(m_1,m_2)^T$ and $s>-1/2$. Finally, we calculated the values of the Monte Carlo variance $VAR_M(g(\overline{\mathrm{x}});g(\mu))$ that was defined as

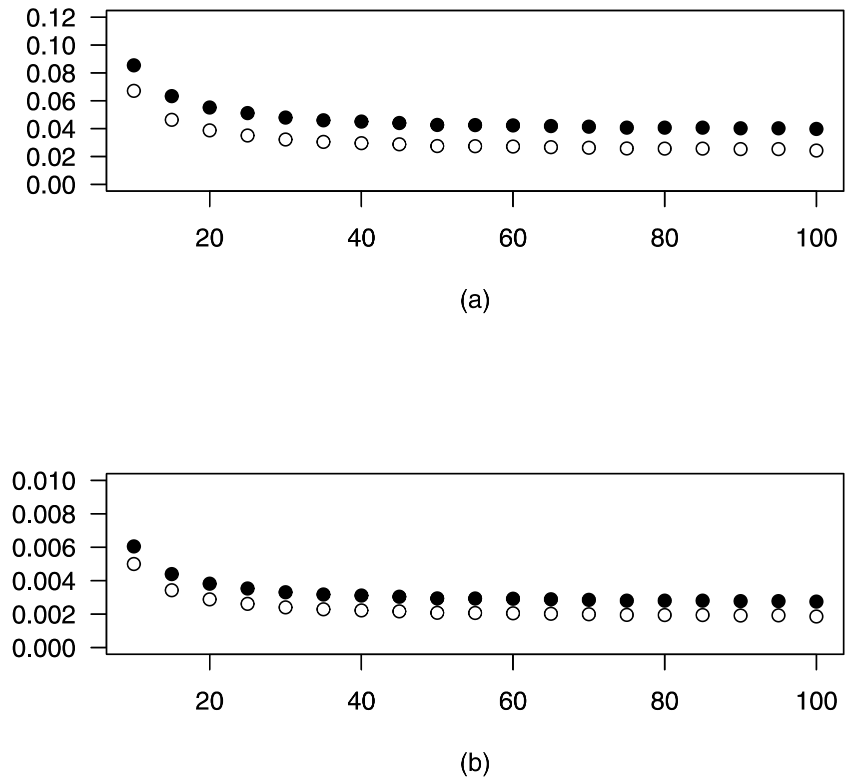


Figure 1. Ratio of means $g(\bar{x}) = \bar{x}_1 / \bar{x}_2$. *Monte Carlo measures of discrepancy* $DI_M(B_m^{(s)}(g;\bar{x}); g(\bar{x}))$, *(●), and* $DI_M(C_m^{(s)}(g;\bar{x}); g(\bar{x}))$, *(○), from* $M = 10000$ *simulations. Simulation parameters, for sample sizes* n *that range from* $n = 10$ *to* $n = 100$; $m_1 = m_2 = 5$ *and* $s = 0.8$ *(panel (a)),* $m_1 = m_2 = 6$ *and* $s = 11$ *(panel (b)).*

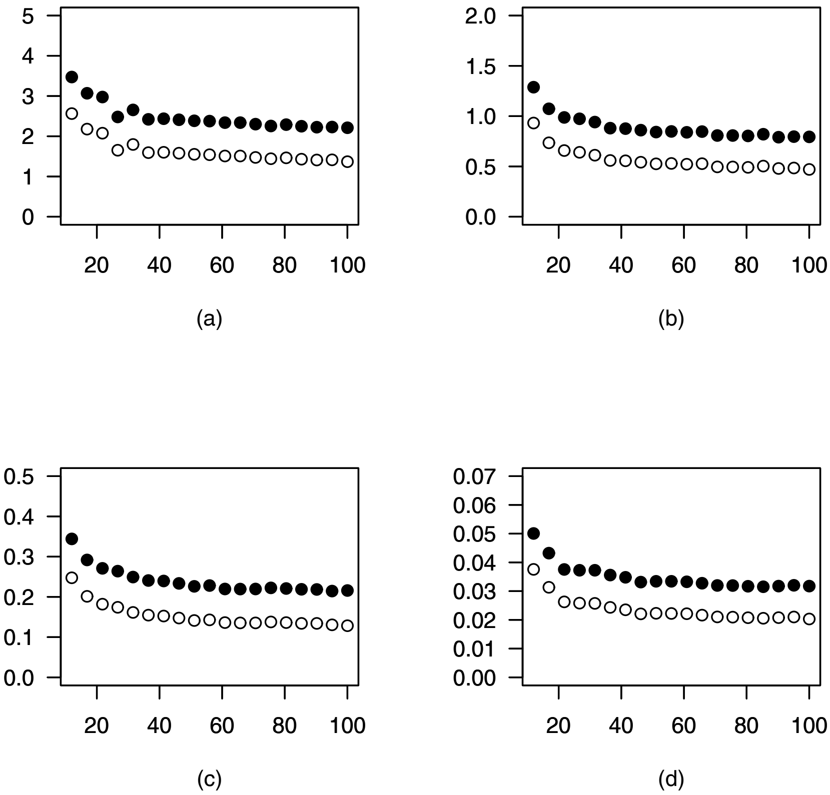


Figure 2. Ratio of means $g(\overline{\mathrm{x}}) = \bar{x}_1 / \bar{x}_2$. *Monte Carlo measures of discrepancy* $DI_M(B_{\mathrm{m}}^{(s)}(g;\overline{\mathrm{x}}); g(\overline{\mathrm{x}}))$, *(•), and* $DI_M(C_{\mathrm{m}}^{(s)}(g;\overline{\mathrm{x}}); g(\overline{\mathrm{x}}))$, *(○), from* $M = 1000$ *simulations. Simulation parameters, for sample sizes* n *that range from* $n = 10$ *to* $n = 100$*;* $m_1 = m_2 = 3$ *and* $s = 0.5$ *(panel (a)),* $m_1 = m_2 = 3$ *and* $s = 0.7$ *(panel (b)),* $m_1 = m_2 = 3$ *and* $s = 1$ *(panel (c)),* $m_1 = m_2 = 3$ *and* $s = 1.5$ *(panel (d)).*

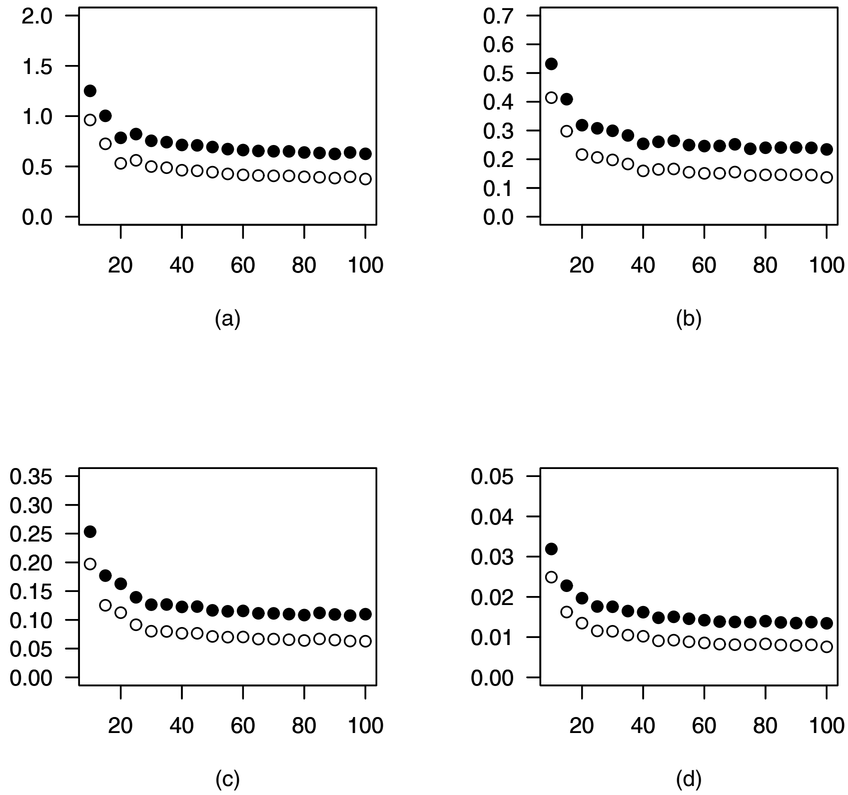


Figure 3. Ratio of means $g(\overline{\mathrm{x}}) = \bar{x}_1 / \bar{x}_2$. *Monte Carlo measures of discrepancy* $DI_M(B_{\mathrm{m}}^{(s)}(g;\overline{\mathrm{x}});g(\overline{\mathrm{x}}))$, *(•), and* $DI_M(C_{\mathrm{m}}^{(s)}(g;\overline{\mathrm{x}});g(\overline{\mathrm{x}}))$, *(○), from* $M = 1000$ *simulations. Simulation parameters, for sample sizes* n *that range from* $n = 10$ *to* $n = 100$; $m_1 = m_2 = 4$ *and* $s = 0.5$ *(panel (a)),* $m_1 = m_2 = 5$ *and* $s = 0.5$ *(panel (b)),* $m_1 = m_2 = 6$ *and* $s = 0.5$ *(panel (c)),* $m_1 = m_2 = 10$ *and* $s = 0.5$ *(panel (d)).*

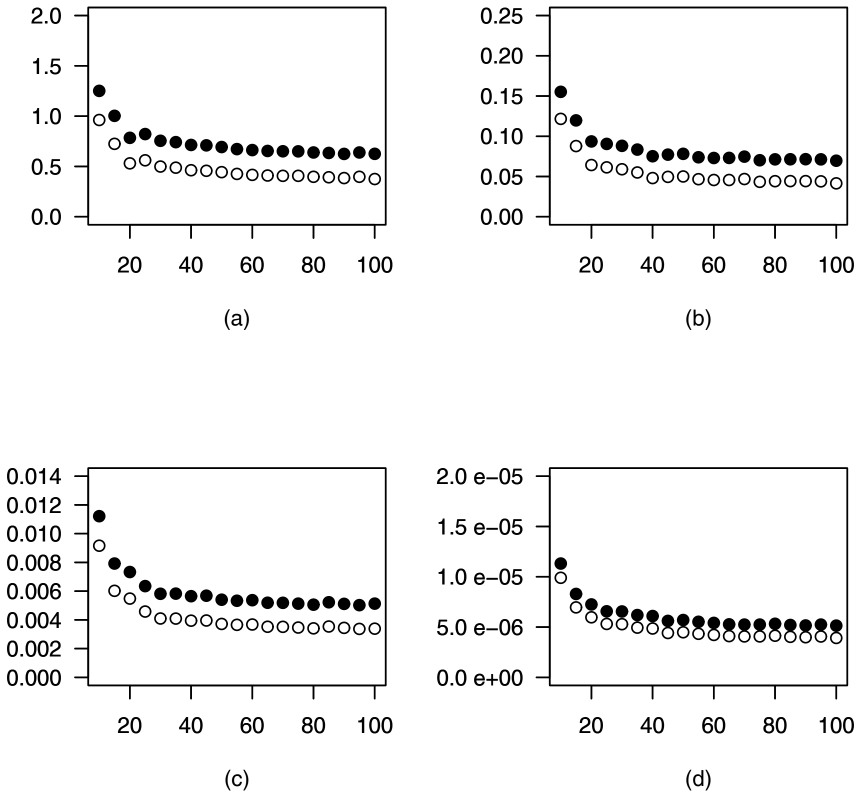
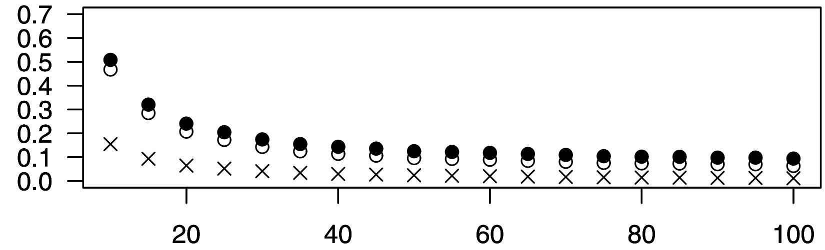
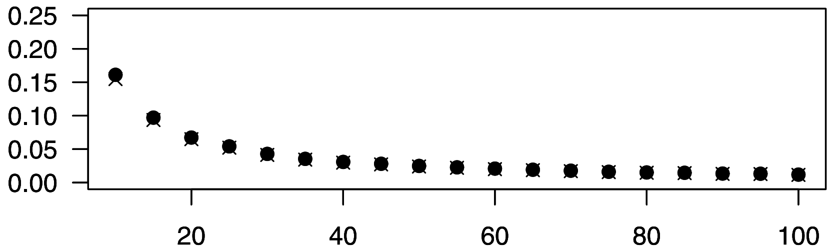


Figure 4. Ratio of means $g(\overline{\mathrm{x}}) = \bar{x}_1 / \bar{x}_2$*. Monte Carlo measures of discrepancy* $DI_M(B_{\mathrm{m}}^{(s)}(g;\overline{\mathrm{x}}); g(\overline{\mathrm{x}}))$*, (•), and* $DI_M(C_{\mathrm{m}}^{(s)}(g;\overline{\mathrm{x}}); g(\overline{\mathrm{x}}))$*, (∘), from* $M = 1000$ *simulations. Simulation parameters, for sample sizes* n *that range from* $n = 10$ *to* $n = 100$*;* $m_1 = m_2 = 4$ *and* $s = 0.5$ *(panel (a)),* $m_1 = m_2 = 5$ *and* $s = 0.7$ *(panel (b)),* $m_1 = m_2 = 6$ *and* $s = 1$ *(panel (c)),* $m_1 = m_2 = 10$ *and* $s = 1.5$ *(panel (d)).*



(a)



(b)

Figure 5. Estimation of the ratio of means $g(\mu) = \mu_1 / \mu_2$. *Monte Carlo variances* $VAR_M(B_{\mathrm{m}}^{(s)}(g;\overline{\mathrm{x}});g(\mu))$, *(•),* $VAR_M(C_{\mathrm{m}}^{(s)}(g;\overline{\mathrm{x}});g(\mu))$, *(○), and* $VAR_M(g(\overline{\mathrm{x}});g(\mu))$, *(×), from* $M = 10000$ *simulations. Simulation parameters, for sample sizes* n *that range from* $n = 10$ *to* $n = 100$; $m_1 = m_2 = 5$ *and* $s = 0.7$ *(panel (a)),* $m_1 = m_2 = 6$ *and* $s = 1.7$ *(panel (b)).*

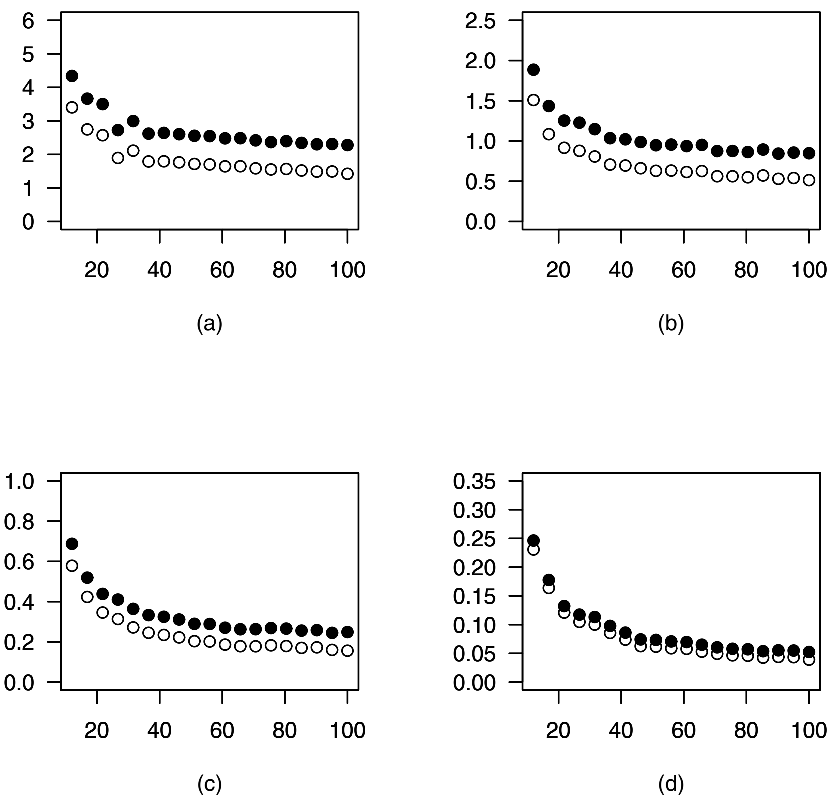


Figure 6. Estimation of the ratio of means $g(\mu) = \mu_1 / \mu_2$. *Monte Carlo variances* $VAR_M(B_{\mathrm{m}}^{(s)}(g;\overline{\mathrm{x}}); g(\mu))$, *(•), and* $VAR_M(C_{\mathrm{m}}^{(s)}(g;\overline{\mathrm{x}}); g(\mu))$, *(∘), from* $M = 1000$ *simulations. Simulation parameters, for sample sizes* n *that range from* $n = 10$ *to* $n = 100$; $m_1 = m_2 = 3$ *and* $s = 0.5$ *(panel (a)),* $m_1 = m_2 = 3$ *and* $s = 0.7$ *(panel (b)),* $m_1 = m_2 = 3$ *and* $s = 1$ *(panel (c)),* $m_1 = m_2 = 3$ *and* $s = 1.5$ *(panel (d)).*

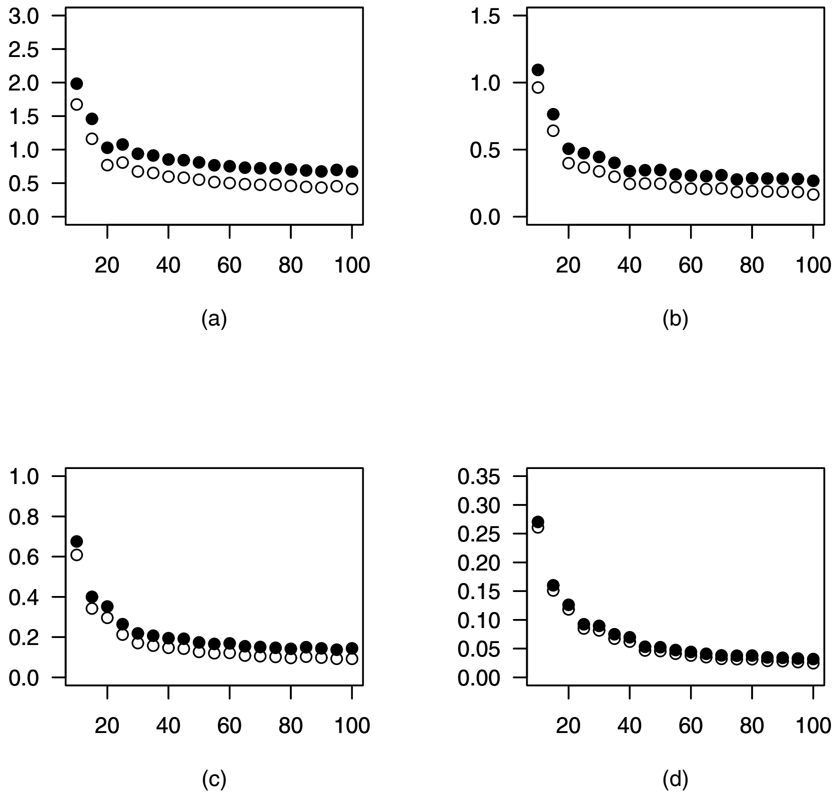


Figure 7. Estimation of the ratio of means $g(\mu) = \mu_1 / \mu_2$*. Monte Carlo variances* $VAR_M(B_{\mathrm{m}}^{(s)}(g;\overline{\mathrm{x}});g(\mu))$*, (•), and* $VAR_M(C_{\mathrm{m}}^{(s)}(g;\overline{\mathrm{x}});g(\mu))$*, (∘), from* $M = 1000$ *simulations. Simulation parameters, for sample sizes* n *that range from* $n = 10$ *to* $n = 100$*;* $m_1 = m_2 = 4$ *and* $s = 0.5$ *(panel (a)),* $m_1 = m_2 = 5$ *and* $s = 0.5$ *(panel (b)),* $m_1 = m_2 = 6$ *and* $s = 0.5$ *(panel (c)),* $m_1 = m_2 = 10$ *and* $s = 0.5$ *(panel (d)).*

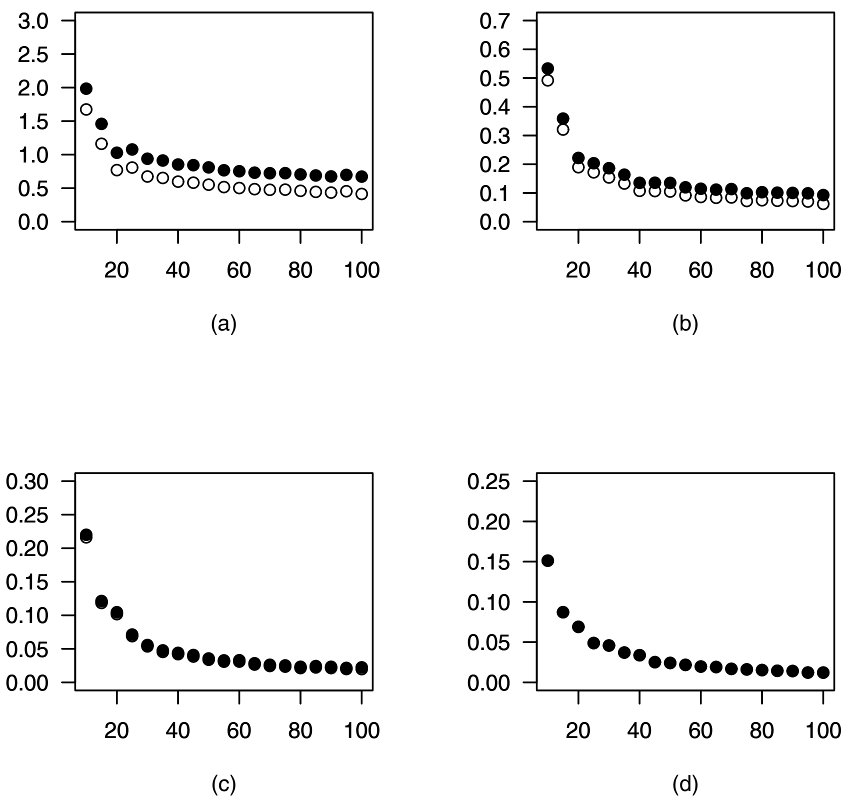


Figure 8. Estimation of the ratio of means $g(\mu)=\mu_1/\mu_2$. *Monte Carlo variances* $VAR_M(B_{\mathrm{m}}^{(s)}(g;\overline{\mathrm{x}});g(\mu))$, *(•), and* $VAR_M(C_{\mathrm{m}}^{(s)}(g;\overline{\mathrm{x}});g(\mu))$, *(○), from* $M=1000$ *simulations. Simulation parameters, for sample sizes* n *that range from* $n=10$ *to* $n=100$; $m_1=m_2=4$ *and* $s=0.5$ *(panel (a)),* $m_1=m_2=5$ *and* $s=0.7$ *(panel (b)),* $m_1=m_2=6$ *and* $s=1$ *(panel (c)),* $m_1=m_2=10$ *and* $s=2$ *(panel (d)).*

$$VAR_M(g(\overline{\mathrm{x}}); g(\mu)) = (M-1)^{-1} \sum_{l=1}^{M} (g(\overline{\mathrm{x}}) - g(\mu))^2 .$$

The adjusted Bernstein-type estimator $C_{\mathrm{m}}^{(s)}(g;\overline{\mathrm{x}})$, given by (17), always outperformed the Bernstein-type estimator $B_{\mathrm{m}}^{(s)}(g;\overline{\mathrm{x}})$, given by (16), in the empirical results we worked out. In any case, the differences were crucial for small values of the simulation parameters, $\mathrm{m} = (m_1, m_2)^T$ and s. Large values for the simulation parameters $\mathrm{m} = (m_1, m_2)^T$ diminishes the benefit of a Bernstein-like scheme based on the simulation parameter s.

6. Conclusions

The Bernstein-type estimators $B_m^{(s)}(g;\bar{x})$ and $B_{\mathrm{m}}^{(s)}(g;\overline{\mathrm{x}})$, given by (3) and (4), respectively, are naturally first-order unbiased. The adjusted estimators $C_m^{(s)}(g;\bar{x})$ and $C_{\mathrm{m}}^{(s)}(g;\overline{\mathrm{x}})$, given by (6) and (10), make use of the second derivative of the smooth function that defines the population smooth function of means.

Most importantly, the constructive coefficient s, where $s > -1/2$, allows a reasonable use of the Bernstein polynomials, through the definitions (3) and (4) of the Bernstein-type approximations $B_m^{(s)}(g;\bar{x})$ and $B_{\mathrm{m}}^{(s)}(g;\overline{\mathrm{x}})$, in general questions of efficient estimation. The constructive coefficient s, in (3) and (4), can make the Bernstein-type approximations effective in statistical inference. For $s > 0$, the Bernstein-type approximations $B_m^{(s)}(g;\bar{x})$ and $B_{\mathrm{m}}^{(s)}(g;\overline{\mathrm{x}})$ always outperform the corresponding Bernstein polynomials ($s = 0$).

Different and interesting fields of applications are density estimation and quantile estimation, beyond the simple framework of the smooth functions of means. See Vitale (1975) and Cheng (1995). These topics will be pursued elsewhere.

Acknowledgements: The Author is indebted to two anonymous Referees for helpful criticisms and suggestions on a previous version of the contribution.

References

Balakrishnan N., Nevzorov V.B. (2003), *A primer on statistical distributions*, Wiley and Sons, New York.

Becker R.A., Chambers J.M., Wilks A.R. (1988), *The new S language*, Wadsworth and Brooks/Cole, Pacific Grove, California.

Cheng C. (1995), The Bernstein polynomial estimator of a smooth quantile function, *Statistics and Probability Letters*, 24, 321-330.

Davis P.J. (1963), *Interpolation and approximation*, Blaisdell Publishing Company, Waltham, Massachusetts.

Feller W. (1971), An *Introduction to probability theory and its applications*, Volume II, 2nd Edition, Wiley and Sons, New York.

Hall P. (1992), *The bootstrap and Edgeworth expansion*, Springer-Verlag, New York.

Korovkin P.P. (1960), *Linear operators and approximation theory*, Hindustan Publishing Corporation, Delhi.

Lorentz G.G. (1986), *Bernstein polynomials*, Chelsea Publishing Company, New York.

Pallini A. (2005), Bernstein-type approximation of smooth functions, *Statistica*, 65, 169-191.

Phillips G.M. (2003), *Interpolation and approximation by polynomials*, Springer-Verlag, New York.

Pinkus A. (2000), Weierstrass and approximation theory, *Journal of Approximation Theory*, 107, 1-66.

Vitale R.A. (1975), A Bernstein polynomial approach to density estimation, In: *Statistical inference and related topics*, M.L. Puri ed., Vol. 2, Academic Press, New York, 87-100.