STATISTICA, anno LXV, n. 3, 2005

AVERAGE LINEAR APPROXIMATIONS FOR SMOOTH FUNCTIONS

Andrea Pallini

1. INTRODUCTION

Smooth functions g are continuous functions admitting an appropriate number of bounded and continuous derivatives. Feller (1971), chapter 7, provides an elegant exposition of select topics in approximation theory for continuous functions, focussing on Bernstein polynomials and uniform approximation in the sense of Weierstrass. See also Davis (1963), chapter 6, Cheney (1982), chapter 3, Cheney (1986), chapter 1, Lorentz (1986), Beckner and Regev (1998), and Pinkus (2000). Bernstein polynomials are positive linear operators defined with Binomial probabilities. Bernstein polynomials converge uniformly to a continuous function of interest. Binomial probabilities in Bernstein polynomials work as a weighting scheme, and asymptotics for these linear operators can be studied by Tchebychev inequalities. On the other hand, with inferential paradigms, Bernstein polynomials are not generally applicable to random samples drawn from a population distribution F. More precisely, Binomial probabilities in Bernstein polynomials have always to play some role in linking the population distribution F with a smooth function g to approximate.

Alternative approximating linear operators that converge uniformly to a smooth function g can be introduced following Pallini (2000, 2002). These approximating linear operators are defined as sums of n smooth functions of the same analytic form as the smooth function g to approximate, calculated on n independent and identically distributed (i.i.d.) random observations with population distribution F. These approximating linear operators for smooth functions g are positive and can use various probability models for their weighting schemes. Following the terminology in Bogachev (2000), these approximating linear operators can be named average linear approximations. The weighting schemes in the proposed average linear approximations do not depend on the population distribution F.

In section 2, we introduce average linear approximations for smooth functions g, with empirical probabilities that define their weighting schemes. These average linear approximations have errors in probability of order $O_p(n^{-s})$ and

$O_p(n^{-2s})$, where $s>0$ is a finite constant, as the sample size n diverges, as $n\to\infty$. In section 3, we introduce average linear approximations using Gaussian probabilities, with errors in probability of order $O_p(n^{-s})$ and $O_p(n^{-2s})$, where $s>0$ is a constant, as $n\to\infty$. In section 4, we study convergence of these average linear approximations, as $n\to\infty$, in the sense of the Weierstrass approximation theorem. We show that their convergence is uniform, as $n\to\infty$, with straightforward applications of Tchebychev inequalities to the empirical and Gaussian probabilities that define their weighting schemes. In section 5, we study the average linear approximations of sections 2 and 3 with situations that require a prearranged error function. In section 6, we will conclude the present contribution with some general remarks.

2. AVERAGE LINEAR APPROXIMATIONS

Let $\chi_n=\{\mathrm{X}_1,\ldots,\mathrm{X}_n\}$ be a random sample of n i.i.d. observations drawn from a q-variate random variable X, taking values in a space $\aleph$ with population distribution F. Let g be a real-valued smooth function of u, where $\mathrm{u}\in V\subseteq\Re^q$. That is, $g:V\to\Re^1$. Let us suppose that the smooth function g satisfies the conditions in appendix (7.1). We want to approximate the values $g(\mathrm{u})$ of the smooth function g on a closed subset U of V, where $\mathrm{u}\in U$ and $U\subseteq V$, by average linear approximations defined on the space $\aleph\times U$.

Smooth function g can also be studied by a random variable X of lower dimension q_0, say $q_0<q$, following, for instance, the asymptotic statistical theory for the class of smooth functions of means. See Bhattacharya and Ghosh (1978), Hall (1992), chapters 2 and 3, and Pallini (2000, 2002).

For any finite constant $s>0$, we can define the average linear approximation $L_n[g(\mathrm{u})]$ of the values $g(\mathrm{u})$ of a smooth function g, where $\mathrm{u}\in U$, with an error in probability of order $O_p(n^{-s})$, as $n\to\infty$. Assigning the empirical probability n^{-1} to every random observation X_{i_1} in the sample χ_n, where $i_1=1,\ldots,n$, we can define the average linear approximation $L_n[g(\mathrm{u})]$ as

$$L_n[g(\mathrm{u})]=n^{-1}\sum_{i_1=1}^{n} g(n^{-s}\mathrm{X}_{i_1}+\mathrm{u}), \tag{1}$$

where $\mathrm{u}\in U$. Observe that $(n^{-s}\mathrm{X}_{i_1}+\mathrm{u})\in V$, for every $i_1=1,\ldots,n$. Similarly, for any finite constant $s>0$, we can define the average linear approximations $M_n[g(\mathrm{u})]$ and $N_n[g(\mathrm{u})]$, where $\mathrm{u}\in U$, with errors in probability of order $O_p(n^{-2s})$, as $n\to\infty$. In particular, we have the definitions

$$M_n[g(\mathrm{u})] = n^{-2} \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} g(n^{-s}(\mathrm{X}_{\mathrm{i}_1} - \mathrm{X}_{\mathrm{i}_2}) + \mathrm{u}), \tag{2}$$

$$N_n[g(\mathrm{u})] = n^{-2q} \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \cdots \sum_{i_{2q-1}=1}^{n} \sum_{i_{2q}=1}^{n} g \begin{pmatrix} n^{-s}(\mathrm{X}_{\mathrm{i}_1}^{(1)} - \mathrm{X}_{\mathrm{i}_2}^{(1)}) + \mathrm{u}^{(1)} \\ \cdots \\ n^{-s}(\mathrm{X}_{\mathrm{i}_{2q-1}}^{(q)} - \mathrm{X}_{\mathrm{i}_{2q}}^{(q)}) + \mathrm{u}^{(q)} \end{pmatrix}, \tag{3}$$

where $\mathrm{u} \in U$.

The average linear approximation $M_n[g(\mathrm{u})]$ uses all differences between the pairs $(\mathrm{X}_{\mathrm{i}_1}, \mathrm{X}_{\mathrm{i}_2})$ of the sample observations in χ_n, where $\mathrm{u} \in U$, and $i_1, i_2 = 1, \ldots, n$. The average linear approximation $N_n[g(\mathrm{u})]$ uses all differences between the pairs $(\mathrm{X}_{\mathrm{i}_1}^{(\mathrm{t}_1)}, \mathrm{X}_{\mathrm{i}_2}^{(\mathrm{t}_1)})$ of the sample observations in χ_n, where $\mathrm{u} \in \mathrm{U}$, $t_1 = 1, \ldots, q$, and $i_1, i_2 = 1, \ldots, n$. Use of pairs in $M_n[g(\mathrm{u})]$ and $N_n[g(\mathrm{u})]$ improves over the error in probability achievable with the average linear approximation $L_n[g(\mathrm{u})]$, for all $\mathrm{u} \in U$.

Errors in probability for average linear approximations $L_n[g(\mathrm{u})]$, $M_n[g(\mathrm{u})]$ and $N_n[g(\mathrm{u})]$, where $\mathrm{u} \in U$, depends on the finite constant $s > 0$, and can be very low. The constant s in definitions (1), (2) and (3) can be used for tuning up the randomness that originates from the sample χ_n. In appendix (7.2), for all $\mathrm{u} \in U$, it is shown that

$$L_n[g(\mathrm{u})] - g(\mathrm{u}) = O_p(n^{-s}), \tag{4}$$

$$M_n[g(\mathrm{u})] - g(\mathrm{u}) = O_p(n^{-2s}), \tag{5}$$

$$N_n[g(\mathrm{u})] - g(\mathrm{u}) = O_p(n^{-2s}), \tag{6}$$

as $n \to \infty$, for every choice of the constant s, where $s > 0$. Observe that, for a fixed s, where $s > 0$, the average linear approximations $M_n[g(\mathrm{u})]$ and $N_n[g(\mathrm{u})]$ perform better than the average linear approximation $L_n[g(\mathrm{u})]$, for all $\mathrm{u} \in U$, as $n \to \infty$.

3. AVERAGE LINEAR APPROXIMATIONS USING GAUSSIAN PROBABILITIES

We denote by ϕ the multivariate Gaussian probability on $\Re^q$, obtained as the product of q standard Gaussian probabilities on $\Re^1$. For any finite constant

$s>0$, we can define the average linear approximation $L_n[g(\mathrm{u});\phi]$, for the values $g(\mathrm{u})$ of a smooth function g, where $\mathrm{u}\in U$, with an error in probability of order $O_p(n^{-s})$, as $n\to\infty$. Assigning the probability

$$\left\{\sum_{i_1=1}^{n}\phi(n^{-s}\mathrm{X}_{\mathrm{i}_1})\right\}^{-1}\phi(n^{-s}\mathrm{X}_{\mathrm{i}_1}),$$

to every random observation $\mathrm{X}_{\mathrm{i}_1}$ in the sample χ_n, where $i_1=1,\ldots,n$, we can define the average linear approximation $L_n[g(\mathrm{u});\phi]$ as

$$L_n[g(\mathrm{u});\phi]=\left\{\sum_{i_1=1}^{n}\phi(n^{-s}\mathrm{X}_{\mathrm{i}_1})\right\}^{-1}\sum_{i_1=1}^{n}g(n^{-s}\mathrm{X}_{\mathrm{i}_1}+\mathrm{u})\,\phi(n^{-s}\mathrm{X}_{\mathrm{i}_1}),\tag{7}$$

where $\mathrm{u}\in U$. Similarly, for any finite constant $s>0$, we can define the average linear approximations $M_n[g(\mathrm{u});\phi]$ and $N_n[g(\mathrm{u});\phi]$, where $\mathrm{u}\in U$, with errors in probability of order $O_p(n^{-2s})$, as $n\to\infty$. In particular, we have the definitions

$$M_n[g(\mathrm{u});\phi]=\left\{\sum_{i_1=1}^{n}\sum_{i_2=1}^{n}\phi(n^{-s}(\mathrm{X}_{\mathrm{i}_1}-\mathrm{X}_{\mathrm{i}_2}))\right\}^{-1}$$

$$\cdot\sum_{i_1=1}^{n}\sum_{i_2=1}^{n}g(n^{-s}(\mathrm{X}_{\mathrm{i}_1}-\mathrm{X}_{\mathrm{i}_2})+\mathrm{u})\,\phi(n^{-s}(\mathrm{X}_{\mathrm{i}_1}-\mathrm{X}_{\mathrm{i}_2})),\tag{8}$$

$$N_n[g(\mathrm{u});\phi]=\left\{\sum_{i_1=1}^{n}\sum_{i_2=1}^{n}\cdots\sum_{i_{2q-1}=1}^{n}\sum_{i_{2q}=1}^{n}\phi(n^{-s}(\mathrm{X}^{(1)}_{\mathrm{i}_1}-\mathrm{X}^{(1)}_{\mathrm{i}_2}))\cdot\cdots\cdot\phi(n^{-s}(\mathrm{X}^{(\mathrm{q})}_{\mathrm{i}_{2\mathrm{q}\text{-}1}}-\mathrm{X}^{(\mathrm{q})}_{\mathrm{i}_{2\mathrm{q}}}))\right\}^{-1}$$

$$\cdot\sum_{i_1=1}^{n}\sum_{i_2=1}^{n}\cdots\sum_{i_{2q-1}=1}^{n}\sum_{i_{2q}=1}^{n}g\begin{pmatrix}n^{-s}(\mathrm{X}^{(1)}_{\mathrm{i}_1}-\mathrm{X}^{(1)}_{\mathrm{i}_2})+\mathrm{u}^{(1)}\\ \cdots\\ n^{-s}(\mathrm{X}^{(\mathrm{q})}_{\mathrm{i}_{2\mathrm{q}}}-\mathrm{X}^{(\mathrm{q})}_{\mathrm{i}_{2\mathrm{q}\text{-}1}})+\mathrm{u}^{(\mathrm{q})}\end{pmatrix}$$

$$\cdot\phi(n^{-s}(\mathrm{X}^{(1)}_{\mathrm{i}_1}-\mathrm{X}^{(1)}_{\mathrm{i}_2}))\cdot\cdots\cdot\phi(n^{-s}(\mathrm{X}^{(\mathrm{q})}_{\mathrm{i}_{2\mathrm{q}\text{-}1}}-\mathrm{X}^{(\mathrm{q})}_{\mathrm{i}_{2\mathrm{q}}})),\tag{9}$$

where $\mathrm{u}\in U$.

Errors in probability for the average linear approximations $L_n[g(\mathrm{u});\phi]$, $M_n[g(\mathrm{u});\phi]$ and $N_n[g(\mathrm{u});\phi]$, where $\mathrm{u}\in U$, can be very low. More precisely, in appendix (7.3), for all $\mathrm{u}\in U$, it is shown that

$$L_n[g(\mathrm{u});\phi]\text{-g(u)} = O_p(n^{-s}), \tag{10}$$

$$M_n[g(\mathrm{u});\phi]\text{-g(u)} = O_p(n^{-2s}), \tag{11}$$

$$N_n[g(\mathrm{u});\phi]\text{-g(u)} = O_p(n^{-2s}), \tag{12}$$

as $n \to \infty$, for every s, where $s > 0$.

Following appendix (7.3), it is seen that the average linear approximations $L_n[g(\mathrm{u})]$ and $L_n[g(\mathrm{u});\phi]$, given by (1) and (7), $M_n[g(\mathrm{u})]$ and $M_n[g(\mathrm{u});\phi]$, given by (2) and (8), and $N_n[g(\mathrm{u})]$ and $N_n[g(\mathrm{u});\phi]$, given by (3) and (9), have the same asymptotic variance, for all $\mathrm{u} \in U$, as $n \to \infty$. To go into some detail, about the average linear approximations $L_n[g(\mathrm{u})]$ and $L_n[g(\mathrm{u});\phi]$, we have that

$$\begin{aligned}\sigma^2(\mathrm{u}) &= E[\{L_n[g(\mathrm{u})]\text{-g(u)}\}^2] \\ &= E[\{L_n[g(\mathrm{u});\phi]\text{-g(u)}\}^2],\end{aligned}$$

as $n \to \infty$, for all $\mathrm{u} \in U$. In any case, the average linear approximations $L_n[g(\mathrm{u})]$, $M_n[g(\mathrm{u})]$, $N_n[g(\mathrm{u})]$, $L_n[g(\mathrm{u});\phi]$, $M_n[g(\mathrm{u});\phi]$ and $N_n[g(\mathrm{u});\phi]$, defined in (1), (2), (3), (7), (8) and (9), may perform differently, whereas the sample size n is finite. See subsection (5.3) below.

4. UNIFORM APPROXIMATION

The convergence of the average linear approximations $L_n[g(\mathrm{u})]$, $M_n[g(\mathrm{u})]$ and $N_n[g(\mathrm{u})]$ to the value g(u) of a smooth function g is uniform, for all $\mathrm{u} \in U$. The rates of convergence in (4), (5) and (6) do not depend on u, for all $\mathrm{u} \in U$. In appendix (7.4), it is shown that

$$L_n[g(\mathrm{u})] \to [g(\mathrm{u})], \tag{13}$$

$$M_n[g(\mathrm{u})] \to [g(\mathrm{u})], \tag{14}$$

$$N_n[g(\mathrm{u})] \to [g(\mathrm{u})], \tag{15}$$

uniformly on $\mathrm{u} \in U$, as $n \to \infty$.

The average linear approximations $L_n[g(\mathrm{u});\phi]$, $M_n[g(\mathrm{u});\phi]$ and $N_n[g(\mathrm{u});\phi]$ have rates of convergence to the value g(u) of a smooth function g, in (10), (11) and (12), that do not depend on u, for all $\mathrm{u} \in U$. In appendix (7.5), it is shown that

$$L_n[g(\mathrm{u});\phi] \to [g(\mathrm{u})], \tag{16}$$

$$M_n[g(\mathrm{u});\phi] \to [g(\mathrm{u})], \tag{17}$$

$$N_n[g(\mathrm{u});\phi] \to [g(\mathrm{u})], \tag{18}$$

uniformly on $\mathrm{u} \in U$, as $n \to \infty$.

5. AVERAGE LINEAR APPROXIMATIONS WITH PREARRANGED ERROR

Let $\Lambda : U \to \Re^1$ be a real-valued error function, with values $\Lambda(\mathrm{u})$, where $\mathrm{u} \in U$. The value $\mathrm{g(u)}$ of a smooth function g can be approximated with a prearranged error $\Lambda(\mathrm{u})$ by average linear approximations obtained from the definitions (1), (2), (3), (7), (8) and (9) of $L_n[g(\mathrm{u})]$, $M_n[g(\mathrm{u})]$, $N_n[g(\mathrm{u})]$, $L_n[g(\mathrm{u});\phi]$, $M_n[g(\mathrm{u});\phi]$ and $N_n[g(\mathrm{u});\phi]$, respectively, for all $\mathrm{u} \in U$. From the definition (1) of $L_n[g(\mathrm{u})]$, it follows that the value $\mathrm{g(u)}$ of a smooth function g with prearranged error can be obtained as

$$g(\mathrm{u}) - \Lambda(\mathrm{u}) = A_n[g(\mathrm{u})] L_n[g(\mathrm{u})], \tag{19}$$

for all $\mathrm{u} \in U$. Observe that the bridge function

$$A_n[g(\mathrm{u})] = \{L_n[g(\mathrm{u})]\}^{-1} \{g(\mathrm{u}) - \Lambda(\mathrm{u})\}, \tag{20}$$

with $L_n[g(\mathrm{u})] \neq 0$, exactly solves the linear equation (19), for every $\mathrm{u} \in U$. If $\Lambda(\mathrm{u}) = 0$, for all $\mathrm{u} \in U$, the linear equation (19) determines the exact bridge between the average linear approximation $L_n[g(\mathrm{u})]$ and the value $\mathrm{g(u)}$, for all $\mathrm{u} \in U$.

From the definitions (2), (3), (7), (8) and (9) of $M_n[g(\mathrm{u})]$, $N_n[g(\mathrm{u})]$, $L_n[g(\mathrm{u});\phi]$, $M_n[g(\mathrm{u});\phi]$ and $N_n[g(\mathrm{u});\phi]$, we can similarly obtain the bridge functions

$$B_n[g(\mathrm{u})] = \{M_n[g(\mathrm{u})]\}^{-1} \{g(\mathrm{u}) - \Lambda(\mathrm{u})\}, \tag{21}$$

$$C_n[g(\mathrm{u})] = \{N_n[g(\mathrm{u})]\}^{-1} \{g(\mathrm{u}) - \Lambda(\mathrm{u})\}, \tag{22}$$

with $M_n[g(\mathrm{u})], N_n[g(\mathrm{u})] \neq 0$, for every $\mathrm{u} \in U$, and

$$A_n[\mathrm{g(u)};\phi] = \{L_n[g(\mathrm{u});\phi]\}^{-1} \{g(\mathrm{u}) - \Lambda(\mathrm{u})\}, \tag{23}$$

$$B_n[g(u);\phi] = \{M_n[g(u);\phi]\}^{-1}\{g(u) - \Lambda(u)\}, \tag{24}$$

$$C_n[g(u);\phi] = \{N_n[g(u);\phi]\}^{-1}\{g(u) - \Lambda(u)\}, \tag{25}$$

with $L_n[g(u);\phi], M_n[g(u);\phi], N_n[g(u);\phi] \neq 0$, for every $u \in U$.

Setting $\Lambda(u) \equiv 0$, the bridge functions $A_n[g(u)]$, $B_n[g(u)]$, $C_n[g(u)]$, $A_n[g(u);\phi]$, $B_n[g(u);\phi]$ and $C_n[g(u);\phi]$, given by (20), (21), (22), (23), (24) and (25), are exact in yielding the value $g(u)$, for all $u \in U$. See appendix (7.6), for further theoretical details.

5.1. *The ratio example*

The smooth function g is defined as $g(u) = u^{(1)}(u^{(2)})^{-1}$, for every $u \in U$, where $u = (u^{(1)}, u^{(2)})$, $U \subseteq V$, and $V = (-\infty, +\infty) \times (0, +\infty)$.

The random sample χ_n consists of n i.i.d. observations $X_{i_1} = (X_{i_1}^{(1)}, X_{i_1}^{(2)})$, where $q = 2$, $i_1 = 1, \ldots, n$. The marginal $X_{i_1}^{(2)}$ ranges in a set of positive real values, for every $i_1 = 1, \ldots, n$.

The average linear approximations $L_n[g(u)]$, $M_n[g(u)]$ and $N_n[g(u)]$, given by (1), (2) and (3), where $u \in U$, are defined as

$$L_n[g(u)] = n^{-1} \sum_{i_1=1}^{n} \frac{n^{-s} X_{i_1}^{(1)} + u^{(1)}}{n^{-s} X_{i_1}^{(2)} + u^{(2)}},$$

$$M_n[g(u)] = n^{-2} \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \frac{n^{-s}(X_{i_1}^{(1)} - X_{i_2}^{(1)}) + u^{(1)}}{n^{-s}(X_{i_1}^{(2)} - X_{i_2}^{(2)}) + u^{(2)}},$$

$$N_n[g(u)] = n^{-4} \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \sum_{i_3=1}^{n} \sum_{i_4=1}^{n} \frac{n^{-s}(X_{i_1}^{(1)} - X_{i_2}^{(1)}) + u^{(1)}}{n^{-s}(X_{i_3}^{(2)} - X_{i_4}^{(2)}) + u^{(2)}},$$

respectively, where $u \in U$. Accordingly, the average linear approximations $L_n[g(u);\phi]$, $M_n[g(u);\phi]$ and $N_n[g(u);\phi]$, given by (7), (8) and (9), where $u \in U$, are defined with Gaussian weights

$$\phi(n^{-s} X_{i_1}) = (2\pi)^{-1} \exp\left\{ n^{-2s} \frac{1}{2} \sum_{t_1=1}^{2} (X_{i_1}^{(t_1)})^2 \right\},$$

$$\phi(n^{-s}(\mathrm{X}_{\mathrm{i}_1}\text{-}\mathrm{X}_{\mathrm{i}_2})) = (2\pi)^{-1}\exp\left\{n^{-2s}\frac{1}{2}\sum_{t_1=1}^{2}(\mathrm{X}_{\mathrm{i}_1}^{(\mathrm{t}_1)} - \mathrm{X}_{\mathrm{i}_2}^{(\mathrm{t}_1)})^2\right\},$$

$$\phi(n^{-s}(\mathrm{X}_{\mathrm{i}_1}\text{-}\mathrm{X}_{\mathrm{i}_2}))\,\phi(n^{-s}(\mathrm{X}_{\mathrm{i}_3}\text{-}\mathrm{X}_{\mathrm{i}_4})) =$$

$$= (2\pi)^{-2}\exp\left\{n^{-2s}\frac{1}{2}\sum_{t_1=1}^{2}\sum_{t_2=1}^{2}(\mathrm{X}_{\mathrm{i}_1}^{(\mathrm{t}_1)} - \mathrm{X}_{\mathrm{i}_2}^{(\mathrm{t}_1)} + \mathrm{X}_{\mathrm{i}_3}^{(\mathrm{t}_2)} - \mathrm{X}_{\mathrm{i}_4}^{(\mathrm{t}_2)})^2\right\},$$

respectively, where $i_1, i_2, i_3, i_4 = 1,\ldots,n$.

5.2. *Some error functions*

For a fixed and finite subset U of the domain V of the smooth function $\mathrm{g(u)} = \mathrm{u}^{(1)}(\mathrm{u}^{(2)})^{-1}$, we define the constant τ as $\tau = \max_{\mathrm{u}\in\mathrm{U}}(|\,\mathrm{u}^{(1)} + \mathrm{u}^{(2)}|)$, where $\mathrm{u} = (\mathrm{u}^{(1)},\mathrm{u}^{(2)})$, and $\mathrm{u}\in U$. We consider the following prearranged error functions

$$\Lambda_{1,\mathrm{n}}(\mathrm{u}) = \sin(2\pi\,\tau^{-1}n^{-1/2}(\mathrm{u}^{(1)} + \mathrm{u}^{(2)})), \tag{26}$$

$$\Lambda_{2,\mathrm{n}}(\mathrm{u}) = \sin(2\pi\,\tau^{-1}n^{-1}(\mathrm{u}^{(1)} + \mathrm{u}^{(2)})), \tag{27}$$

$$\Lambda_{3}(\mathrm{u}) = \sin(2\pi\,\tau^{-1}(\mathrm{u}^{(1)} + \mathrm{u}^{(2)})), \tag{28}$$

where $\mathrm{u} = (\mathrm{u}^{(1)},\mathrm{u}^{(2)})$, and $\mathrm{u}\in U$. Observe that, for all $\mathrm{u}\in\mathrm{U}$, $\Lambda_{1,\mathrm{n}}(\mathrm{u})\to 0$, $\Lambda_{2,\mathrm{n}}(\mathrm{u})\to 0$, as $n\to\infty$. See also figure 1, where the error functions $\Lambda_{1,\mathrm{n}}(\mathrm{u})$, $\Lambda_{2,\mathrm{n}}(\mathrm{u})$, and $\Lambda_{3}(\mathrm{u})$, defined in (26), (27) and (28), are plotted together on $\mathrm{u}^{(1)} + \mathrm{u}^{(2)}$, where $\mathrm{u}^{(1)},\mathrm{u}^{(2)}\in[0,5]$.

5.3. *Empirical results*

We consider the smooth function $\mathrm{g(u)} = \mathrm{u}^{(1)}(\mathrm{u}^{(2)})^{-1}$, where $\mathrm{u} = (\mathrm{u}^{(1)},\mathrm{u}^{(2)})$ and $\mathrm{u}\in U$, and where the subset $U\subseteq V$ is determined by the equation $\mathrm{u}^{(2)} = (\mathrm{u}^{(1)})^2 + \mathrm{u}^{(1)} + 1$, for every $\mathrm{u} = (\mathrm{u}^{(1)},\mathrm{u}^{(2)})\in V$.

Simulated data were generated from a bivariate folded-normal distribution, where the dimension q is $q = 2$. In particular, let $W_1 = |N(0,1)|$, $W_2 = |N(0,1)|$ and $W_3 = |N(0,1)|$ be i.i.d. random variables, where $q = 1$. The folded-normal random variable $\mathrm{X}_{\mathrm{i}_1} = (\mathrm{X}_{i_1}^{(1)},\mathrm{X}_{\mathrm{i}_1}^{(2)})$ in the sample χ_n of size n is obtained as $\mathrm{X}_{\mathrm{i}_1}^{(1)} = W_1 + W_3$ and $\mathrm{X}_{\mathrm{i}_1}^{(2)} = W_2 + W_3$, where $i_1 = 1,\ldots,n$. Observe that the components $X^{(1)}_{i_1}$ and $X^{(2)}_{i_1}$ in $(X^{(1)}_{i_1},X^{(2)}_{i_1})$ have correlation coefficient $\rho = 0.5$, for every $i_1 = 1,\ldots,n$.

Figure 2 and figure 3 show the performance of the average linear approximations $L_n[g(\mathrm{u})]$, $M_n[g(\mathrm{u})]$ and $N_n[g(\mathrm{u})]$, given by (1), (2) and (3), respectively, for $s = 1$ and $s = 1.5$, where $\mathrm{u} = (\mathrm{u}^{(1)},\mathrm{u}^{(2)})$, and $\mathrm{u} \in U$. The value $s = 1.5$ yields the best results. Approximations $M_n[g(\mathrm{u})]$ and $N_n[g(\mathrm{u})]$ always outperform $L_n[g(\mathrm{u})]$, where $\mathrm{u} = (\mathrm{u}^{(1)},\mathrm{u}^{(2)})$, and $\mathrm{u} \in U$.

Figure 4 shows the performance of the average linear approximations $L_n[g(\mathrm{u});\phi]$, $M_n[g(\mathrm{u});\phi]$ and $N_n[g(\mathrm{u});\phi]$, given by (7), (8) and (9), respectively, for $s = 1.5$, where $\mathrm{u} \in U$. These approximations may be preferable with situations that require higher levels of randomness from the sample χ_n. Approximation $N_n[g(\mathrm{u});\phi]$, where $\mathrm{u} \in U$, always produces the best result.

Figure 5 shows the performance of the average linear approximation $L_n[g(\mathrm{u})]$, given by (1), with prearranged error functions $\Lambda_{1,n}(\mathrm{u})$, $\Lambda_{2,n}(\mathrm{u})$ and $\Lambda_3(\mathrm{u})$, given by (26), (27) and (28), where $\mathrm{u} \in U$. More precisely, the functions $L_n[g(\mathrm{u})] + \Lambda_{1,n}(\mathrm{u})$, $L_n[g(\mathrm{u})] + \Lambda_{2,n}(\mathrm{u})$ and $L_n[g(\mathrm{u})] + \Lambda_3(\mathrm{u})$ are plotted with the corresponding bridge function $A_n[g(\mathrm{u})]$, given by (20), for exactness, where $\mathrm{u} \in U$.

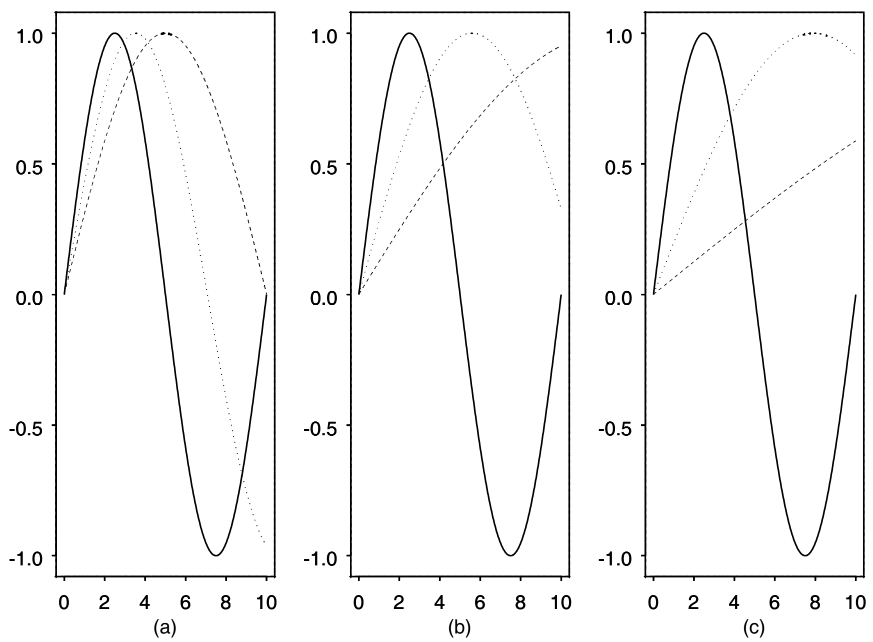

Figure 1 – Error functions $\Lambda_{1,n}(\mathrm{u})$, $(\cdots)$, $\Lambda_{2,n}(\mathrm{u})$, $(---)$, and $\Lambda_3(\mathrm{u})$, $(\textit{solid})$, defined in (26), (27) and (28), where $\mathrm{u} = (\mathrm{u}^{(1)},\mathrm{u}^{(2)})$; horizontal axes given by the range of $\mathrm{u}^{(1)} + \mathrm{u}^{(2)}$, where $\mathrm{u}^{(1)},\mathrm{u}^{(2)} \in [0,5]$. Simulation parameter n set as $n = 2$ (panel (a)), $n = 5$ (panel (b)), $n = 10$ (panel (c)).

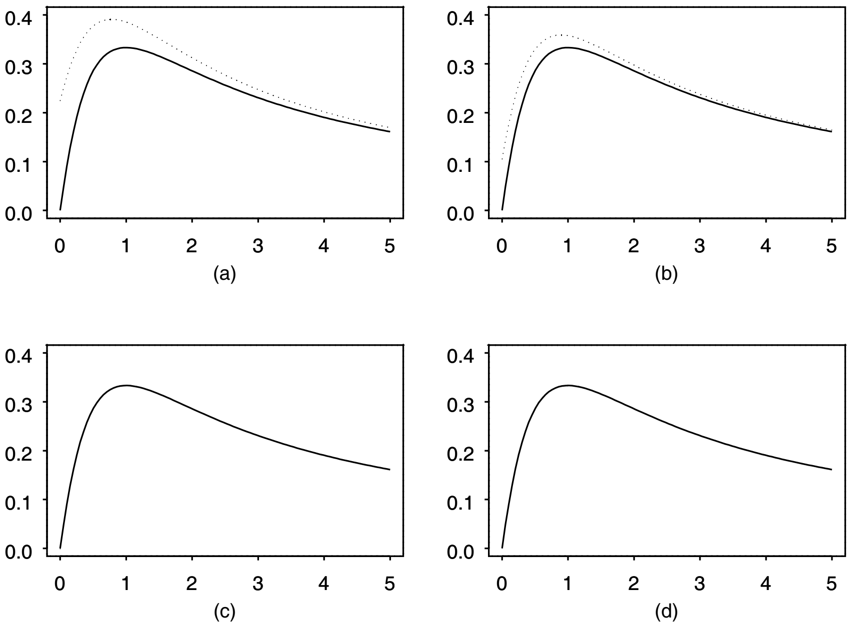

Figure 2 – Average linear approximations $L_n[g(\mathrm{u})]$, $M_n[g(\mathrm{u})]$ and $N_n[g(\mathrm{u})]$, $(\cdots)$, for the ratio $\mathrm{g(u)} = \mathrm{u}^{(1)}(\mathrm{u}^{(2)})^{-1}$, $(solid)$, given by (1), (2) and (3), with $s = 1$, where $\mathrm{u} = (\mathrm{u}^{(1)}, \mathrm{u}^{(2)})$, $\mathrm{u}^{(2)} = (\mathrm{u}^{(1)})^2 + \mathrm{u}^{(1)} + 1$, and $\mathrm{u} \in U$. Approximations $L_n[g(\mathrm{u})]$, for $n = 5$ (panel (a)), for $n = 10$ (panel (b)), $M_n[g(\mathrm{u})]$, for $n = 10$ (panel (c)), and $N_n[g(\mathrm{u})]$, for $n = 10$ (panel (d)), where $\mathrm{u} \in U$.

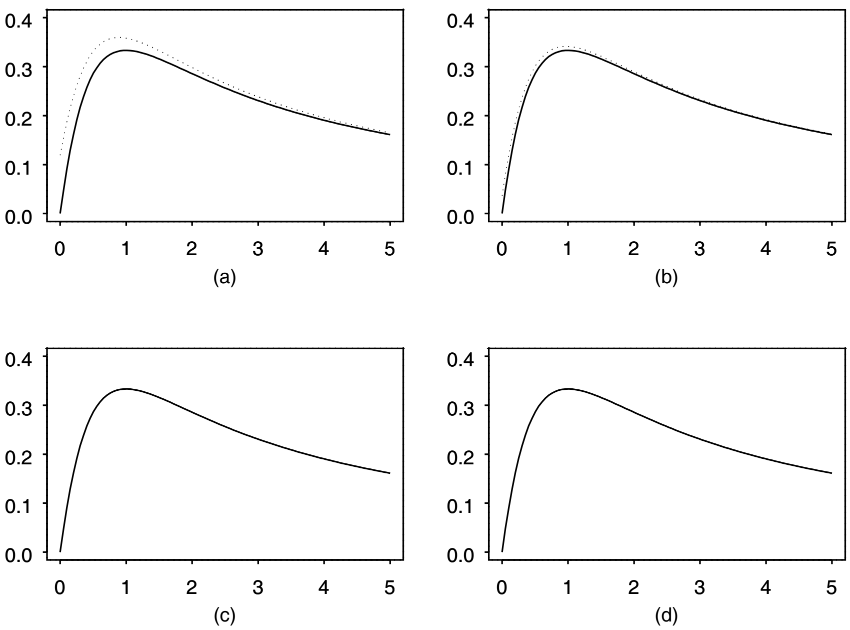

Figure 3 – Average linear approximations $L_n[g(\mathrm{u})]$, $M_n[g(\mathrm{u})]$ and $N_n[g(\mathrm{u})]$, $(\cdots)$, for the ratio $\mathrm{g(u)} = \mathrm{u}^{(1)}(\mathrm{u}^{(2)})^{-1}$, $(solid)$, given by (1), (2) and (3), with $s = 1.5$, where $\mathrm{u} = (\mathrm{u}^{(1)}, \mathrm{u}^{(2)})$, $\mathrm{u}^{(2)} = (\mathrm{u}^{(1)})^2 + \mathrm{u}^{(1)} + 1$, and $\mathrm{u} \in U$. Approximations $L_n[g(\mathrm{u})]$, for $n = 5$ (panel (a)), for $n = 10$ (panel (b)), $M_n[g(\mathrm{u})]$, for $n = 10$ (panel (c)), and $N_n[g(\mathrm{u})]$, for $n = 10$ (panel (d)), where $\mathrm{u} \in U$.

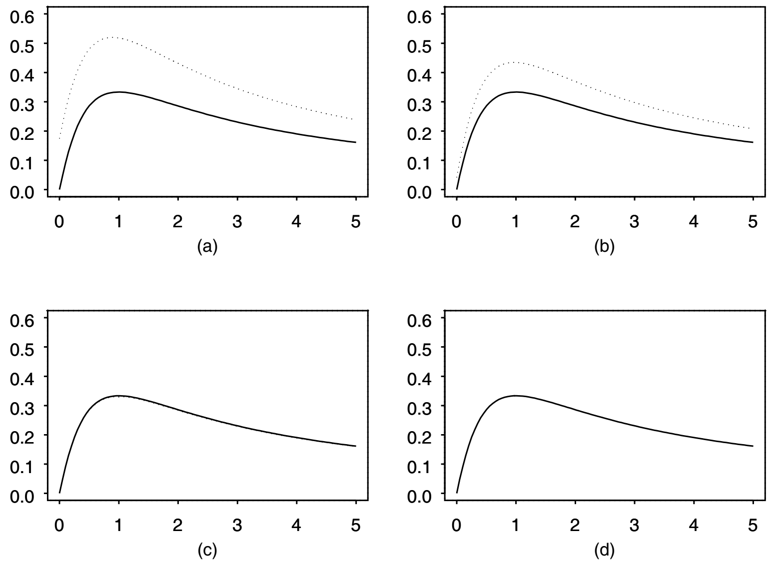

Figure 4 – Average linear approximations $L_n[g(\text{u});\phi]$, $M_n[g(\text{u});\phi]$ and $N_n[g(\text{u});\phi]$, $(\cdots)$, for the ratio $\text{g(u)} = \text{u}^{(1)}(\text{u}^{(2)})^{-1}$, (*solid*), given by (7), (8) and (9), with $s = 1.5$, where $\text{u} = (\text{u}^{(1)},\text{u}^{(2)})$, $\text{u}^{(2)} = (\text{u}^{(1)})^2 + \text{u}^{(1)} + 1$, and $\text{u} \in U$. Approximations $L_n[g(\text{u});\phi]$, for $n = 5$ (panel (a)), for $n = 10$ (panel (b)), $M_n[g(\text{u});\phi]$, for $n = 10$ (panel (c)), and $N_n[g(\text{u});\phi]$, for $n = 10$ (panel (d)), where $\text{u} \in U$.

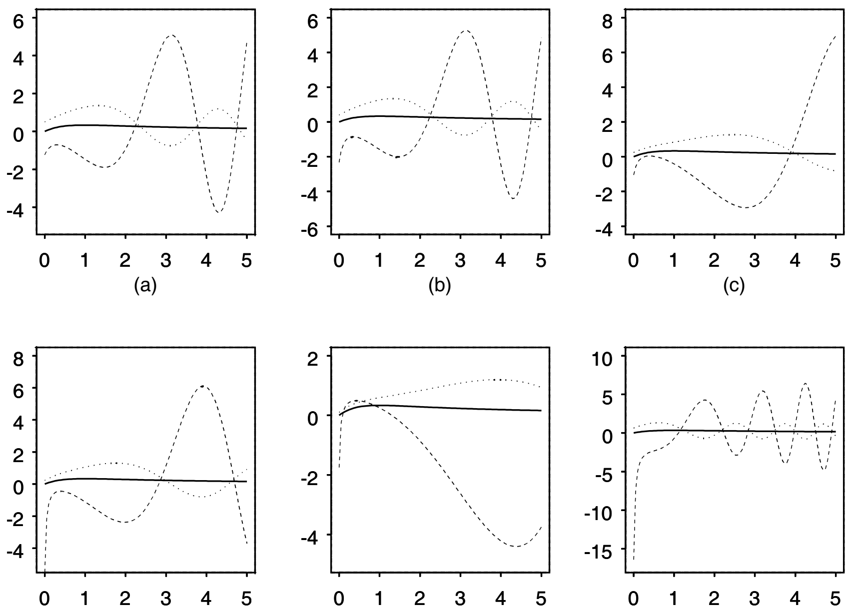

Figure 5 – Average linear approximation with prearranged error $L_n[g(\text{u})] + \Lambda(\text{u})$, $(\cdots)$, where $L_n[g(\text{u})]$ is given by (1), for the ratio $\text{g(u)} = \text{u}^{(1)}(\text{u}^{(2)})^{-1}$, (*solid*), where $\text{u} = (\text{u}^{(1)},\text{u}^{(2)})$, $\text{u}^{(2)} = (\text{u}^{(1)})^2 + \text{u}^{(1)} + 1$, and $\text{u} \in U$. Approximation $L_n[g(\text{u})]$, for $n = 5$ with $s = 1$ and error function $\Lambda_{1,n}(\text{u})$, (panel (a)), for $n = 5$ with $s = 1.5$ and $\Lambda_{1,n}(\text{u})$, (panel (b)), for $n = 5$ with $s = 1.5$ and error function $\Lambda_{2,n}(\text{u})$, (panel (c)), for $n = 10$ with $s = 1.5$ and $\Lambda_{1,n}(\text{u})$, (panel (d)), for $n = 10$ with $s = 1.5$ and $\Lambda_{2,n}(\text{u})$, (panel (e)), for $n = 10$ with $s = 1.5$ and error function $\Lambda_3(\text{u})$, (panel (f)), where $\text{u} \in U$. Errors $\Lambda_{1,n}(\text{u})$, $\Lambda_{2,n}(\text{u})$ and $\Lambda_3(\text{u})$ are given by (26), (27) and (28), respectively, where $\text{u} \in U$. Bridge $A_n[g(\text{u})]$ for exactness, (- - -), (panels (a) to (f)), is defined in (20), where $\text{u} \in U$.

6. CONCLUDING REMARKS

(6.1). Proofs in appendixes (7.4) and (7.5) of uniform convergence of the average linear approximations $L_n[g(\mathrm{u})]$, $M_n[g(\mathrm{u})]$, $N_n[g(\mathrm{u})]$, $L_n[g(\mathrm{u});\phi]$, $M_n[g(\mathrm{u});\phi]$ and $N_n[g(\mathrm{u});\phi]$, where $\mathrm{u} \in U$, given by (1), (2), (3), (7), (8) and (9), in the sense of the Weierstrass approximation theorem, are based on straightforward applications of Tchebychev inequalities to the probabilities that define their weighting schemes. See Lorentz (1986), chapter 1, for a parallel proof of uniform convergence of the Bernstein polynomials. Alternative ways of showing uniform convergence may be deduced from the review of Pinkus (2000) of the most relevant proofs of the Weierstrass approximation theorem.

The average linear approximations (1), (2), (3), (7), (8) and (9) are positive linear operators. For instance, the average linear approximation $L_n[g(\mathrm{u})]$, where $\mathrm{u} \in U$, defined in (1), is a linear operator, because $g(\mathrm{u}) = a\eta(\mathrm{u}) + b\theta(\mathrm{u})$, where a and b are reals, and $\eta(\mathrm{u})$ and $\theta(\mathrm{u})$ are functions, implies that $L_n[g(\mathrm{u})] = aL_n[\eta(\mathrm{u})] + bL_n[\theta(\mathrm{u})]$, for all $\mathrm{u} \in U$, where $U \subseteq V$. Moreover, the linear operator $L_n[g(\mathrm{u})]$, where $\mathrm{u} \in U$, is positive, because $g(\mathrm{u}) \geq 0$, for all $\mathrm{u} \in V$, implies that $L_n[g(\mathrm{v})] \geq 0$, for all $\mathrm{v} \in U$, where $U \subseteq V$. The theory of positive linear operators can provide unifying concepts for the uniform convergence of the average linear approximations (1), (2), (3), (7), (8) and (9). See, specifically, the results in Korovkin (1960), chapters 1 and 4. See also Stone (1932), Akhiezer and Glazman (1993), and Small and McLeish (1994).

(6.2). Following the definition (7) of the average linear approximation $L_n[g(\mathrm{u});\phi]$, where $\mathrm{u} \in U$, we can define the average linear approximation $L_n^{\perp}[g(\mathrm{u});\phi]$ as

$$L_n^{\perp}[g(\mathrm{u});\phi] = \left\{ \sum_{i_1=1}^{n} \phi(g(n^{-s}\,\mathrm{X}_{\mathrm{i}_1} + \mathrm{u})) \right\}^{-1} \sum_{i_1=1}^{n} g(n^{-s}\,\mathrm{X}_{\mathrm{i}_1} + \mathrm{u})\, \phi(g(n^{-s}\,\mathrm{X}_{\mathrm{i}_1} + \mathrm{u})),$$

where ϕ is the standard Gaussian probability on $\Re^1$, and $\mathrm{u} \in U$. Following the definitions (8) and (9) for $M_n[g(\mathrm{u});\phi]$ and $N_n[g(\mathrm{u});\phi]$, we can similarly define the average linear approximations $M_n^{\perp}[g(\mathrm{u});\phi]$ and $N_n^{\perp}[g(\mathrm{u});\phi]$, where $\mathrm{u} \in U$. Alternative definitions can directly use multivariate Gaussian probabilities ϕ on $\Re^q$, and specific covariance structures. Alternative continuous probability models can be used as well.

(6.3). Separation of variables can be very important to apply the average linear approximations given by (1), (2), (3), (7), (8) and (9). Let ψ be a function, such that $n^{-1}\sum_{i_1=1}^{n}\psi(n^{-s}\mathrm{X}_{\mathrm{i}_1})=1+O_p(n^{-s})$, as $n\to\infty$. For instance, $L_n[g(\mathrm{u})]$, given by (1), may become

$$L_n^*[g(\mathrm{u})]=n^{-1}g(\mathrm{u})\sum_{i_1=1}^{n}\psi(n^{-s}\mathrm{X}_{\mathrm{i}_1})$$

where $\mathrm{u}\in U$. Observe that, for all $\mathrm{u}\in U$, $L_n^*[g(\mathrm{u})]=g(\mathrm{u})+O_p(n^{-s})$, where $\mathrm{u}\in U$, as $n\to\infty$.

(6.4). Average linear approximations (1), (2), (3), (7), (8) and (9) can be studied according to the mathematical theory of discretization errors. See, among others, Korovkin (1960), chapter 4, Cheney (1982), chapter 3, and Lorentz (1986), chapter 1.

(6.5). Bogachev (2000) theoretically explores projections of Gaussian measures on linear spaces, by assuming the existence of individual functionals to be added together. See also Bogachev (1998), chapter 3. Average linear approximations (1), (2), (3), (7), (8) and (9) are constructive and directly applicable. In particular, they do not require projection techniques for their definition and their properties of uniform convergence.

Following Bogachev (2000), let $\Im$ be the class of permissible information functions. We denote by $g^{(t_1)}(\mathrm{u})$ the t_1 th information input, where $t_1=1,\ldots,p$, and $\mathrm{u}\in\mathrm{U}$. Average linear approximations (1), (2), (3), (7), (8) and (9) can be used for the information operator $\chi(\mathrm{u})=\{g^{(1)}(\mathrm{u}),\ldots,g^{(q)}(\mathrm{u})\}$, defined as a vector of p information inputs, where p can depend on u, for all $\mathrm{u}\in U$. These average linear approximations can similarly work for the information operator

$$\chi(\mathrm{u})=\{g^{(1)}(\mathrm{u}),g^{(2)}(\mathrm{u},g^{(1)}(\mathrm{u})),\ldots,g^{(p)}(\mathrm{u},g^{(1)}(\mathrm{u}),\ldots,g^{(p-1)}(\mathrm{u}))\},$$

that is defined by recursive inputs, for all $\mathrm{u}\in U$.

(6.6). Bridge functions $A_n[g(\mathrm{u})]$, $B_n[g(\mathrm{u})]$, $C_n[g(\mathrm{u})]$, $A_n[\mathrm{g}(\mathrm{u});\phi]$, $B_n[\mathrm{g}(\mathrm{u});\phi]$ and $C_n[\mathrm{g}(\mathrm{u});\phi]$, given by (20), (21), (22), (23), (24) and (25), can also be defined with a prearranged error function $\Lambda_{\mathrm{n}}(\mathrm{u})$ that depends on the random sample χ_n of size n, where $\mathrm{u}\in U$. The error function $\Lambda_{\mathrm{n}}(\mathrm{u})$ can be defined so that $\Lambda_{\mathrm{n}}(\mathrm{u})\to\lambda$, in probability, as $n\to\infty$, for all $\mathrm{u}\in U$, where λ is a real value.

(6.7). Average linear approximation $N_n[g(\mathrm{u})]$, given by (3), can be generalized by using all the differences between the pairs of components $(\mathrm{X}_{\mathrm{i}_1}^{(t_1)},\mathrm{X}_{\mathrm{i}_2}^{(t_2)})$ of the sample observations in χ_n, where $\mathrm{u} \in U$, $t_1, t_2 = 1, \ldots, q$, and $i_1, i_2 = 1, \ldots, n$. For instance, we may define $N_n^*[g(\mathrm{u})]$ as

$$N_n^*[g(\mathrm{u})] = n^{-2s} \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \cdots \sum_{i_{2q-1}=1}^{n} \sum_{i_{2q}=1}^{n} g \begin{pmatrix} n^{-s}(\mathrm{X}_{\mathrm{i}_1}^{(1)} - \mathrm{X}_{\mathrm{i}_2}^{(2)}) + \mathrm{u}^{(1)} \\ \cdots \\ n^{-s}(\mathrm{X}_{\mathrm{i}_{2q-3}}^{(q-1)} - \mathrm{X}_{\mathrm{i}_{2q-2}}^{(q)}) + \mathrm{u}^{(q-1)} \\ n^{-s}(\mathrm{X}_{\mathrm{i}_{2q-1}}^{(1)} - \mathrm{X}_{\mathrm{i}_{2q}}^{(1)}) + \mathrm{u}^{(q)} \end{pmatrix},$$

where $\mathrm{u} \in U$. Following appendix (7.2), for all $\mathrm{u} \in U$, it can be shown that

$$N_n^*[g(\mathrm{u})] - g(\mathrm{u}) = O_p(n^{-2s}),$$

as $n \to \infty$. From the definition (9) of $N_n[g(\mathrm{u});\phi]$, we can obtain a similar version $N_n^*[g(\mathrm{u});\phi]$, where $\mathrm{u} \in U$. Following appendix (7.3), for all $\mathrm{u} \in U$, it can be shown that $N_n^*[g(\mathrm{u})]$ approximates $g(\mathrm{u})$ with an error $O_p(n^{-2s})$, as $n \to \infty$. Alternative average linear approximations can be defined on fixed functional relationships for the sample observations in χ_n, on reasonable estimates, with appropriate values of $s > 0$, of main effects, covariances, interactions, in the population distribution F. Following appendixes (7.4) and (7.5), it can be shown that all these generalizations are uniform approximations.

(6.8). Following the definitions of average linear approximations (1), (2), (3), (7), (8) and (9), the general theory of the Tchebychev approximation can yield alternative and relevant definitions. The Tchebychev approximation of the values $g(\mathrm{u})$ of a smooth function g, where $\mathrm{u} \in U$, can be accomplished by the minimax solution of an inconsistent system of linear equations. See Davis (1963), chapter 7, Feller (1971), chapter 7, Cheney (1982), chapters 1 to 3, and Cheney (1986). See also Bogachev (2000).

7. APPENDIX

(7.1). Assumptions

The smooth function g is continuous on its domain V, and bounded by a constant $G_1 > 0$, $|g(\mathrm{u})| \le G_1$, for all $\mathrm{u} \in V$.

Average approximations $L_n[g(\mathrm{u})]$, $M_n[g(\mathrm{u})]$, $N_n[g(\mathrm{u})]$, $L_n[g(\mathrm{u});\phi]$, $M_n[g(\mathrm{u});\phi]$ and $N_n[g(\mathrm{u});\phi]$, defined in (1), (2), (3), (7), (8) and (9), where $\mathrm{u} \in U$, and $U \subseteq V$, are bounded by constants $G_2, G_3, G_4, G_5, G_6, G_7 > 0$,

$|L_n[g(\mathrm{u})]| \le G_2$, $|M_n[g(\mathrm{u})]| \le G_3$, $|N_n[g(\mathrm{u})]| \le G_4$, $|L_n[g(\mathrm{u});\phi]| \le G_5$, $|M_n[g(\mathrm{u});\phi]| \le G_6$ and $|N_n[g(\mathrm{u});\phi]| \le G_7$, for all $\mathrm{u} \in V$.

We denote by $g_{t_1 \cdots t_k}(\mathrm{u})$ the derivative of order k of g,

$$g_{t_1 \cdots t_k}(\mathrm{u}) = \partial^k g(\mathrm{u}) \big/ \partial \mathrm{u}^{(\mathrm{t}_1)} \cdots \partial \mathrm{u}^{(\mathrm{t}_k)},$$

where $\mathrm{u} \in V$. The derivative $g_{t_1 \cdots t_k}(\mathrm{u})$ is bounded by a constant $G_8 > 0$, $|g_{t_1 \cdots t_k}(\mathrm{u})| \le G_8$, where $k \le 2$, for all $\mathrm{u} \in V$.

(7.2). Proof of $O_p(n^{-s})$ in (4), and $O_p(n^{-2s})$ in (5) and (6), as $n \to \infty$

The average linear approximations $L_n[g(\mathrm{u})]$, $M_n[g(\mathrm{u})]$ and $N_n[g(\mathrm{u})]$ are defined in (1), (2) and (3), where $\mathrm{u} \in U$.

By Taylor expanding around u the function

$$g(n^{-s}\,\mathrm{X}_{\mathrm{i}_1} + \mathrm{u}),$$

for every $i_1 = 1, \ldots, n$, where $\mathrm{u} \in U$, the function

$$g(n^{-s}\,(\mathrm{X}_{\mathrm{i}_1}\text{-}\mathrm{X}_{\mathrm{i}_2}) + \mathrm{u}),$$

for every pair (i_1, i_2), where $i_1, i_2 = 1, \ldots, n$, and $\mathrm{u} \in U$, and the function

$$g(n^{-s}\,(\mathrm{X}^{(1)}_{\mathrm{i}_1}\text{-}\mathrm{X}^{(1)}_{\mathrm{i}_2}) + \mathrm{u}^{(1)}, \ldots, n^{-s}\,(\mathrm{X}^{(\mathrm{q})}_{\mathrm{i}_{2\mathrm{q}}}\text{-}\mathrm{X}^{(\mathrm{q})}_{\mathrm{i}_{2\mathrm{q}\text{-}1}}) + \mathrm{u}^{(\mathrm{q})})$$

for every vector $(i_1, i_2, \ldots, i_{2q-1}, i_{2q})$, where $i_1, i_2, \ldots, i_{2q-1}, i_{2q} = 1, \ldots, n$, $\mathrm{u} = (\mathrm{u}^{(1)}, \ldots, \mathrm{u}^{(\mathrm{q})})$, and $\mathrm{u} \in U$, we have that

$$L_n[g(\mathrm{u})] - g(\mathrm{u}) = \Gamma_{1,n}(\mathrm{u}) + \Omega_{1,n}(\mathrm{u}),$$

$$M_n[g(\mathrm{u})] - g(\mathrm{u}) = \Gamma_{2,n}(\mathrm{u}) + \Omega_{2,n}(\mathrm{u}),$$

$$N_n[g(\mathrm{u})] - g(\mathrm{u}) = \Gamma_{3,n}(\mathrm{u}) + \Omega_{3,n}(\mathrm{u}),$$

for all $\mathrm{u} \in U$. The quantities $\Gamma_{1,n}(\mathrm{u})$, $\Gamma_{2,n}(\mathrm{u})$ and $\Gamma_{3,n}(\mathrm{u})$ are obtained as

$$\Gamma_{1,n}(\mathrm{u}) = n^{-1} \sum_{i_1=1}^{n} \sum_{t_1=1}^{q} g_{t_1}(\mathrm{u})\, n^{-s}\, \mathrm{X}^{(t_1)}_{\mathrm{i}_1},$$

$$\Gamma_{2,n}(\mathrm{u}) = n^{-2} \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \sum_{t_1=1}^{q} g_{t_1}(\mathrm{u})\, n^{-s}\, (\mathrm{X}^{(\mathrm{t}_1)}_{\mathrm{i}_1} - \mathrm{X}^{(\mathrm{t}_1)}_{\mathrm{i}_2}),$$

$$\Gamma_{3,n}(\mathrm{u}) = n^{-2q}\sum_{i_1=1}^{n}\sum_{i_2=1}^{n}\cdots\sum_{i_{2q-1}=1}^{n}\sum_{i_{2q}=1}^{n}\begin{pmatrix} g_1(\mathrm{u}) \\ \cdots \\ g_q(\mathrm{u}) \end{pmatrix}^T \begin{pmatrix} n^{-s}(\mathrm{X}_{\mathrm{i}_1}^{(1)} - \mathrm{X}_{\mathrm{i}_2}^{(1)}) \\ \cdots \\ n^{-s}(\mathrm{X}_{\mathrm{i}_{2q-1}}^{(q)} - \mathrm{X}_{\mathrm{i}_{2q}}^{(q)}) \end{pmatrix},$$

where $\Omega_{k,n}(\mathrm{u})$, $k = 1,2,3$, are the remainders, and $\mathrm{u} \in U$.

Note that

$$E[n^{-s}\,\mathrm{X}_{\mathrm{i}_1}^{(t_1)}] = O(n^{-s}),$$

as $n \to \infty$, for every $i_1 = 1,\ldots,n$, where $t_1 = 1,\ldots,q$,

$$E[n^{-s}\,(\mathrm{X}_{\mathrm{i}_1}^{(t_1)}\text{-}\mathrm{X}_{\mathrm{i}_2}^{(t_1)})] = 0,$$

for every pair (i_1, i_2), where $i_1, i_2 = 1,\ldots,n$, and $t_1 = 1,\ldots,q$. Note also that

$$E[n^{-s}\,(\mathrm{X}_{\mathrm{i}_1}^{(t_1)}\text{-}\mathrm{X}_{\mathrm{i}_2}^{(t_1)})\, n^{-s}\,(\mathrm{X}_{\mathrm{i}_3}^{(t_2)}\text{-}\mathrm{X}_{\mathrm{i}_4}^{(t_2)})] = O(n^{-2s}),$$

as $n \to \infty$, for every vector (i_1, i_2, i_3, i_4), where $i_1, i_2, i_3, i_4 = 1,\ldots,n$, and $t_1, t_2 = 1,\ldots,q$. Then, for every $\mathrm{u} \in U$,

$$E[|\Gamma_{k,n}(\mathrm{u}) + \Omega_{k,n}(\mathrm{u})|] \geq |E[\Gamma_{k,n}(\mathrm{u}) + \Omega_{k,n}(\mathrm{u})]| = \omega_{k,n}(\mathrm{u}) > 0,$$

where $k = 1,2,3$. There also exists a finite constant $M_{k,n}(\mathrm{u}) > 0$, such that

$$M_{k,n}(\mathrm{u})\,\omega_{k,n}(\mathrm{u}) = E[|\Gamma_{k,n}(\mathrm{u}) + \Omega_{k,n}(\mathrm{u})|],$$

where $k = 1,2,3$, and $\mathrm{u} \in U$. For every $t > 0$, and every $\mathrm{u} \in U$, we also have that

$$\begin{aligned} M_{k,n}(\mathrm{u})\,\omega_{k,n}(\mathrm{u}) &= \int_0^{+\infty} v\, dP(|\Gamma_{k,n}(\mathrm{u}) + \Omega_{k,n}(\mathrm{u})| \leq v) \\ &\geq \int_{t\,\omega_{k,n}(\mathrm{u})}^{+\infty} v\, dP(|\Gamma_{k,n}(\mathrm{u}) + \Omega_{k,n}(\mathrm{u})| \leq v) \\ &\geq t\,\omega_{k,n}(\mathrm{u}) \int_{t\,\omega_{k,n}(\mathrm{u})}^{+\infty} dP(|\Gamma_{k,n}(\mathrm{u}) + \Omega_{k,n}(\mathrm{u})| \leq v) \\ &= t\,\omega_{k,n}(\mathrm{u})\; P(|\Gamma_{k,n}(\mathrm{u}) + \Omega_{k,n}(\mathrm{u})| > t\,\omega_{k,n}(\mathrm{u})), \end{aligned}$$

where $k=1,2,3$, and $\mathrm{u}\in U$. Setting $t\,\omega_{k,n}(\mathrm{u})=\varepsilon$, where $\varepsilon>0$, we obtain the Tchebychev inequality

$$P(|\Gamma_{k,n}(\mathrm{u})+\Omega_{k,n}(\mathrm{u})|>\varepsilon)\leq\varepsilon^{-1}M_{k,n}(\mathrm{u})\omega_{k,n}(\mathrm{u}),$$

where $k=1,2,3$, and $\mathrm{u}\in U$. For every $s>0$, it follows that $\Gamma_{1,n}(\mathrm{u})+\Omega_{1,n}(\mathrm{u})=O_p(n^{-s})$, where $\mathrm{u}\in U$, as $n\to\infty$, because $M_{1,n}(\mathrm{u})=O(1)$ and $\omega_{1,n}(\mathrm{u})=O(n^{-s})$, where $\mathrm{u}\in U$, as $n\to\infty$. It also follows that $\Gamma_{k,n}(\mathrm{u})+\Omega_{k,n}(\mathrm{u})=O_p(n^{-2s})$, where $k=2,3$, and $\mathrm{u}\in U$, as $n\to\infty$, because $M_{k,n}(\mathrm{u})=O(1)$ and $\omega_{k,n}(\mathrm{u})=O(n^{-s})$, where $k=2,3$, and $\mathrm{u}\in U$, as $n\to\infty$. Accordingly, it holds that $\Omega_{1,n}(\mathrm{u})=O_p(n^{-2s})$, where $\mathrm{u}\in U$, as $n\to\infty$, and $\Omega_{k,n}(\mathrm{u})=O_p(n^{-3s})$, where $k=2,3$, and $\mathrm{u}\in U$, as $n\to\infty$.

(7.3). Proof of $O_p(n^{-s})$ in (10), and $O_p(n^{-2s})$ in (11) and (12), as $n\to\infty$

The average linear approximations $L_n[g(\mathrm{u});\phi]$, $M_n[g(\mathrm{u});\phi]$ and $N_n[g(\mathrm{u});\phi]$ are defined in (7), (8) and (9), where $\mathrm{u}\in U$.

Observe that

$$\left\{\sum_{i_1=1}^{n}\phi(n^{-s}\mathrm{X}_{\mathrm{i}_1})\right\}^{-1}\phi(n^{-s}\mathrm{X}_{\mathrm{i}_1})=n^{-1}+O_p(n^{-(1+s)}),$$

as $n\to\infty$, for every $i_1=1,\ldots,n$,

$$\left\{\sum_{i_1=1}^{n}\sum_{i_2=1}^{n}\phi(n^{-s}(\mathrm{X}_{\mathrm{i}_1}\text{-}\mathrm{X}_{\mathrm{i}_2}))\right\}^{-1}\phi(n^{-s}(\mathrm{X}_{\mathrm{i}_1}\text{-}\mathrm{X}_{\mathrm{i}_2}))=n^{-2}+O_p(n^{-2(1+s)}),$$

as $n\to\infty$, for every pair (i_1,i_2), where $i_1,i_2=1,\ldots,n$,

$$\left\{\sum_{i_1=1}^{n}\sum_{i_2=1}^{n}\cdots\sum_{i_{2q-1}=1}^{n}\sum_{i_{2q}=1}^{n}\phi(n^{-s}(\mathrm{X}^{(1)}_{\mathrm{i}_1}\text{-}\mathrm{X}^{(1)}_{\mathrm{i}_2}))\cdot\cdots\cdot\phi(n^{-s}(\mathrm{X}^{(\mathrm{q})}_{\mathrm{i}_{2\mathrm{q-1}}}\text{-}\mathrm{X}^{(\mathrm{q})}_{\mathrm{i}_{2\mathrm{q}}}))\right\}^{-1}$$

$$\cdot\,\phi(n^{-s}(\mathrm{X}^{(1)}_{\mathrm{i}_1}\text{-}\mathrm{X}^{(1)}_{\mathrm{i}_2}))\cdot\cdots\cdot\phi(n^{-s}(\mathrm{X}^{(\mathrm{q})}_{\mathrm{i}_{2\mathrm{q-1}}}\text{-}\mathrm{X}^{(\mathrm{q})}_{\mathrm{i}_{2\mathrm{q}}}))$$

$$=n^{-2q}+O_p(n^{-2(q+s)}),$$

as $n \to \infty$, for every vector $(i_1, i_2, \ldots, i_{2q-1}, i_{2q})$, where $i_1, i_2, \ldots, i_{2q-1}, i_{2q} = 1, \ldots, n$. Then, proofs of $O_p(n^{-s})$, as $n \to \infty$, in (10), and $O_p(n^{-2s})$, as $n \to \infty$, in (11) and (12), directly proceeds following appendix (7.2).

(7.4). Proof of uniform approximation in (13), (14) and (15)

The average linear approximation $L_n[g(\mathrm{u})]$ is defined in (1), where $\mathrm{u} \in U$. Under the assumption in appendix (7.1) that g is a bounded smooth function, it holds that

$$| g(n^{-s}\,\mathrm{X}_{\mathrm{i}_1} + \mathrm{u}) - g(\mathrm{u}) | \leq 2G_1,$$

for every $i_1 = 1, \ldots, n$, and all $\mathrm{u} \in U$. The sample variance $\hat{\sigma}^2(\mathrm{u})$ of $L_n[g(\mathrm{u})]$, with probability n^{-1} assigned to every sample vector $\mathrm{X}_{\mathrm{i}_1}$, where $i_1 = 1, \ldots, n$, can be defined by

$$n\hat{\sigma}^2(\mathrm{u}) = n^{-1} \sum_{i_1=1}^{n} \{g(n^{-s}\,\mathrm{X}_{\mathrm{i}_1} + \mathrm{u}) - g(\mathrm{u})\}^2,$$

for all $\mathrm{u} \in U$. For any constant $\delta > 0$, we obtain the Tchebychev inequalities

$$\sum_{|n^{-1}\{g(n^{-s}\mathrm{X}_{\mathrm{i}_1}+\mathrm{u}) - g(\mathrm{u})\}| \geq \delta} n^{-1}$$

$$\leq n^{-2}\delta^{-2} \sum_{|n^{-1}\{g(n^{-s}\mathrm{X}_{\mathrm{i}_1}+\mathrm{u}) - g(\mathrm{u})\}| \geq \delta} \{g(n^{-s}\,\mathrm{X}_{\mathrm{i}_1} + \mathrm{u}) - g(\mathrm{u})\}^2$$

$$\leq \delta^{-2}\,\hat{\sigma}^2(\mathrm{u})$$

$$\leq n^{-1}\delta^{-2}\,4G_1^2,$$

where $\mathrm{u} \in U$. We let $\|\mathrm{u}\| = (\mathrm{u}\mathrm{u}^{\mathrm{T}})^{1/2}$, for all $\mathrm{u} \in V$. Observe that, for a given constant $\varepsilon > 0$, we can find a constant $\delta_0 > 0$ such that $\|\mathrm{u}^*\text{-}\mathrm{u}\| < \delta_0$ implies that $|n^{-1}\{g(\mathrm{u}^*)\text{-g}(\mathrm{u})\}| < \delta$ and $|g(\mathrm{u}^*)\text{-g}(\mathrm{u})| < \varepsilon$, where $\delta > 0$ and $\mathrm{u}^*, \mathrm{u} \in V$. Then, it follows that

$$| L_n[g(\mathrm{u})] - g(\mathrm{u}) | = \left| n^{-1} \sum_{i_1=1}^{n} \{g(n^{-s}\,\mathrm{X}_{\mathrm{i}_1} + \mathrm{u}) - g(\mathrm{u})\} \right|$$

$$\leq n^{-1} \sum_{|n^{-1}\{g(n^{-s}\mathrm{X}_{\mathrm{i}_1}+\mathrm{u})\ -g(\mathrm{u})\}|<\delta} |\, g(n^{-s}\,\mathrm{X}_{\mathrm{i}_1}+\mathrm{u})-g(\mathrm{u})\,|$$

$$+\,n^{-1} \sum_{|n^{-1}\{g(n^{-s}\mathrm{X}_{\mathrm{i}_1}+\mathrm{u})\ -g(\mathrm{u})\}|\geq\delta} |\, g(n^{-s}\,\mathrm{X}_{\mathrm{i}_1}+\mathrm{u})-g(\mathrm{u})\,|$$

$$\leq \varepsilon + n^{-1}\delta^{-2}4G_1^2 .$$

For n sufficiently large, for all $\mathrm{u}\in U$, we also have that

$$|\, L_n[g(\mathrm{u})]-g(\mathrm{u})\,|<2\varepsilon .$$

Finally, δ is independent of u, for all $\mathrm{u}\in V$. Under the assumption in appendix (7.1) that g is a continuous function, it follows that $L_n[g(\mathrm{u})]\rightarrow g(\mathrm{u})$ uniformly, for all $\mathrm{u}\in V$, as $n\rightarrow\infty$.

Uniformity in (14) and (15) for the average linear approximations $M_n[g(\mathrm{u})]$ and $N_n[g(\mathrm{u})]$, defined in (2) and (3), where $\mathrm{u}\in U$, can be shown in the same way.

(7.5). Proof of uniform approximation in (16), (17) and (18)

The average linear approximation $L_n[g(\mathrm{u});\phi]$ is defined in (7) and studied by (16), where $\mathrm{u}\in U$. The sample variance $\hat{\sigma}^2(\mathrm{u};\phi)$ of $L_n[g(\mathrm{u});\phi]$, with probability

$$\left\{\sum_{i_1=1}^{n}\phi(n^{-s}\,\mathrm{X}_{\mathrm{i}_1})\right\}^{-1}\phi(n^{-s}\,\mathrm{X}_{\mathrm{i}_1}),$$

assigned to every sample vector $\mathrm{X}_{\mathrm{i}_1}$, where $i_1=1,\ldots,n$, is defined by

$$\left\{\sum_{i_1=1}^{n}\phi(n^{-s}\,\mathrm{X}_{\mathrm{i}_1})\right\}\hat{\sigma}^2(\mathrm{u};\phi)=\left\{\sum_{i_1=1}^{n}\phi(n^{-s}\,\mathrm{X}_{\mathrm{i}_1})\right\}^{-1}\sum_{i_1=1}^{n}\{g(n^{-s}\,\mathrm{X}_{\mathrm{i}_1}+\mathrm{u})-g(\mathrm{u})\}^2\,\phi(n^{-s}\,\mathrm{X}_{\mathrm{i}_1}),$$

for all $\mathrm{u}\in U$. Following appendix (7.4), under the assumptions in appendix (7.1), for any constant $\delta>0$, we obtain

$$\left\{\sum_{i_1=1}^{n}\phi(n^{-s}\,\mathrm{X}_{\mathrm{i}_1})\right\}^{-1}\sum_{|n^{-1}\{g(n^{-s}\mathrm{X}_{\mathrm{i}_1}+\mathrm{u})\ -g(\mathrm{u})\}|\geq\delta}\phi(n^{-s}\,\mathrm{X}_{\mathrm{i}_1})$$

$$\le \left\{ \sum_{i_1=1}^{n} \phi(n^{-s} \mathrm{X}_{\mathrm{i}_1}) \right\}^{-2} \delta^{-2} \sum_{|n^{-1}\{g(n^{-s}\mathrm{X}_{\mathrm{i}_1}+\mathrm{u})\; -g(\mathrm{u})\}| \ge \delta} \{g(n^{-s}\mathrm{X}_{\mathrm{i}_1}+\mathrm{u})-g(\mathrm{u})\}^2 \phi(n^{-s}\mathrm{X}_{\mathrm{i}_1})$$

$$\le \delta^{-2} \hat{\sigma}^2(\mathrm{u};\phi)$$

$$\le \left\{ \sum_{i_1=1}^{n} \phi(n^{-s} \mathrm{X}_{\mathrm{i}_1}) \right\}^{-1} \delta^{-2} 4G_1^2,$$

where $\mathrm{u} \in U$. For a given constant $\varepsilon > 0$, we have that

$$|L_n[g(\mathrm{u});\phi] - g(\mathrm{u})|$$

$$\le \left\{ \sum_{i_1=1}^{n} \phi(n^{-s} \mathrm{X}_{\mathrm{i}_1}) \right\}^{-1} \sum_{|n^{-1}\{g(n^{-s}\mathrm{X}_{\mathrm{i}_1}+\mathrm{u})\; -g(\mathrm{u})\}| < \delta} |g(n^{-s}\mathrm{X}_{\mathrm{i}_1}+\mathrm{u})-g(\mathrm{u})| \phi(n^{-s}\mathrm{X}_{\mathrm{i}_1})$$

$$+ \left\{ \sum_{i_1=1}^{n} \phi(n^{-s} \mathrm{X}_{\mathrm{i}_1}) \right\}^{-1} \sum_{|n^{-1}\{g(n^{-s}\mathrm{X}_{\mathrm{i}_1}+\mathrm{u})\; -g(\mathrm{u})\}| \ge \delta} |g(n^{-s}\mathrm{X}_{\mathrm{i}_1}+\mathrm{u})-g(\mathrm{u})| \phi(n^{-s}\mathrm{X}_{\mathrm{i}_1})$$

$$\le \varepsilon + \left\{ \sum_{i_1=1}^{n} \phi(n^{-s} \mathrm{X}_{\mathrm{i}_1}) \right\}^{-1} \delta^{-2} 4G_1^2,$$

where $\mathrm{u} \in U$. Finally, $L_n[g(\mathrm{u});\phi] \to g(\mathrm{u})$ uniformly, for all $\mathrm{u} \in U$, as $n \to \infty$.

Uniformity in (17) and (18) for the average linear approximations $M_n[g(\mathrm{u});\phi]$ and $N_n[g(\mathrm{u});\phi]$, defined in (8) and (9), where $\mathrm{u} \in U$, as $n \to \infty$, can be shown in the same way.

(7.6). Bridge functions (1), (2), (3), (7), (8) and (9)

In appendix (7.4), the result (19) for the bridge function $A_n[g(\mathrm{u})]$, given by (20), can be regarded as exactness for the average linear approximation $L_n[g(\mathrm{u})]$ in the corresponding sequence of approximations, for all $\mathrm{u} \in U$, as $n \to \infty$. Similar results hold, in appendixes (7.4) and (7.5), for the bridge functions $B_n[g(\mathrm{u})]$, $C_n[g(\mathrm{u})]$, $A_n[\mathrm{g(u)};\phi]$, $B_n[\mathrm{g(u)};\phi]$ and $C_n[\mathrm{g(u)};\phi]$, given by (21), (22), (23), (24) and (25), and the corresponding sequences of approximations, for all $\mathrm{u} \in U$, as $n \to \infty$.

Dipartimento di Statistica e Matematica
Applicata all'Economia – Università di Pisa

ANDREA PALLINI

REFERENCES

N.I. AKHIEZER, I.M. GLAZMAN (1993), *Theory of Linear Operators in Hilbert Space*, Volumes I and II, Reprint of the 1961 and 1963 Editions, Dover Publications, Mineola, New York.

O.E. BARNDORFF-NIELSEN, D.R. COX (1989), *Asymptotic Techniques for Use in Statistics*, Chapman and Hall, London.

W. BECKNER, A. REGEV (1998), *Asymptotic estimates using probability*, "Advances in Mathematics", 14, 1-14.

R.N. BHATTACHARYA, J.K. GHOSH (1978), *On the validity of the formal Edgeworth expansion*, "Annals of Statistics", 6, 434-451.

V.I. BOGACHEV (1998), *Gaussian Measures*, American Mathematical Society, Providence, Rhode Island.

V. BOGACHEV (2000), *Average approximations and moments of measures*, "Journal of Complexity", 16, 390-410.

E.W. CHENEY (1982), *Introduction to Approximation Theory*, 2nd Edition, AMS Chelsea Publishing, American Mathematical Society, Providence, Rhode Island.

E.W. CHENEY (1986), *Multivariate Approximation Theory: Selected Topics*, SIAM, Society for Industrial and Applied Mathematics, Philadelphia, Pennsylvania.

P.J. DAVIS (1963), *Interpolation and Approximation*, Blaisdell Publishing Company, Waltham, Massachusetts.

W. FELLER (1971), *An Introduction to Probability Theory and Its Applications*, Volume II, 2nd Edition, Wiley and Sons, New York.

P. HALL (1992), *The Bootstrap and Edgeworth Expansion*, Springer-Verlag, New York.

J.P. KEENER (2000), *Principles of Applied Mathematics*, Rev. Ed., Perseus Books, Cambridge, Massachusetts.

P.P. KOROVKIN (1960), *Linear Operators and Approximation Theory*, Hindustan Publishing Corporation, Delhi.

G.G. LORENTZ (1986), *Bernstein Polynomials*, Chelsea Publishing Company, New York.

A. PALLINI (2000), *Efficient estimation of bootstrap distribution functions*, "Metron", 58 (1-2), 81-95.

A. PALLINI (2002), *On a linear method in bootstrap confidence intervals*, "Statistica", 62, 1-21.

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

P.K. SEN, J.M. SINGER (1993), *Large Sample Methods in Statistics*, Chapman and Hall, New York.

C.G. SMALL, D.L. MCLEISH (1994), *Hilbert Space Methods in Probability and Statistical Inference*, Wiley and Sons, New York.

M.H. STONE (1932), *Linear Transformations in Hilbert Space and Their Applications to Analysis*, AMS Colloquium Publications 15, American Mathematical Society, Providence, Rhode Island.

RIASSUNTO

Approssimazioni lineari medie per funzioni regolari

Vengono introdotte approssimazioni lineari medie per funzioni regolari, utilizzando probabilità empiriche e Gaussiane. Viene dimostrato come la convergenza di queste approssimazioni sia uniforme, nel senso del teorema di approssimazione di Weierstrass. Vengono infine studiate approsimazioni lineari medie con funzioni prestabilite di errore.

SUMMARY

Average linear approximations for smooth functions

Average linear approximations for smooth functions using empirical and Gaussian probabilities are introduced. The convergence of these approximations is shown to be uniform, in the sense of the Weierstrass approximation theorem. Average linear approximations with prearranged error functions are finally studied.