Pointwise Convergence of a Nonparametric Estimator of Regression in a Measurable Space Used in Contingent Valuation Method

Taïbi-Hassani Salima1,2, Dimitri Laroutis1 and S. L. Adigaw-E-Touck2
1. Esitpa, 3, Rue du Tronquet CS 40118, 76134 Mont Saint Aignan France.
2. Laboratoire de Mathématiques LMRS UMR 6085 CNRS-Université de Rouen France.

Received: October 26, 2011 / Accepted: February 23, 2015 / Published: May 25, 2015.

Abstract: The Contingent Valuation Method is used to evaluate individual preferences for a change concerning a public non-market resource or property. The objective is to build a nonparametric forecasting model of an individual's Willingness To Pay according to geographical location. Within this framework, an estimator (of type Nadaraya-Watson) is proposed for the regression of the variable related to geolocation. The specific characteristics of the location variable lead us to a more general regression model than the traditional models. Results are established for convergence of our estimator.

Keywords: Regression, nonparametric estimation, mixing process, almost complete convergence, contingent valuation method.

1. Introduction

It has become usual to apply the Contingent Valuation Method (CVM) to evaluate individuals' preferences, in particular when considering non-market resources or property. This method is based on a questionnaire which reveals the individual's Willingness To Pay (WTP). CVM is one of the methods most frequently used in the evaluation of wetlands (Loomis et al. 2000; Yang et al. 2008; Laroutis and Taïbi, 2010), threatened species (Loureiro and Ojea 2008), biodiversity (Christie et al. 2006), and so on.

The surveys include a detailed description of the resource to be evaluated, and based on this information, the individuals are asked to state their WTP, that is say how much of their revenue they would be willing to pay to maintain the current level of services received from the resource. Other questions include the usual socio-economic and demographic information and these are used as potential explanatory indicators for the observed variations in the individual preferences.

Recently, several studies have underlined the significance of an individual's location on the WTP (Garcia et al. 2009; Loomis et al. 2000; Loureiro and Ojea 2008; Wang et al. 2007). Beaumais et al. (2008), evaluating preservation of the Seine Estuary wetlands, found significant differences in WTP between inhabitants in the rural countryside and those of large cities (who were ready to contribute more to protect the wetlands).

Similar observations led Spash et al. (2009) to stress the need for more precise analysis of urban/rural differences and the impact of individuals' geographical location on WTP. Our objective is to build a nonparametric model of Willingness To Pay based on geographical location. By considering WTP to be a (real) random variable \( Y \) with \( Y = r(x) \) where \( x \) is a fixed point (with
random variable $X$ taking values in a measurable space $(E, E)$, we then estimate the regression $r$ of $Y$ in relation to $X$. This can be seen to be an extension of the well-studied case with $R^p$.

Then, after defining an estimator of type Nadaraya-Watson, we give results for pointwise convergence of this estimator. Several prediction models for the WTP (often parametric) are generally used in Contingent Valuation: simple or multiple logistic regression (Alvarez-Farizo et al. 2009; Bekele and Drake 2003; Calkins et al. 2002), Probit model (Maruyama and Takimoto 2008) or again Tobit model (Beaumais et al. 2008). Comparisons between the models have also been made - for example Crooker and Herriges (2004) seek to evaluate and compare the variations of WTP estimations as a function of the distribution of individual preferences and the type of estimation procedure used.

In section 2 and section 3, we discuss nonparametric estimation of the regression function and identify our assumptions for the following sections. Then, in section 4 and section 5, we describe our WTP estimator and establish convergence properties, with proofs in section 6. Finally in section 7 we give some other potential applications of the estimator.

2. Nonparametric Estimation of the Regression Function

Let $(\Omega, A, P)$ be a probability space, $(E, \mathcal{E}, \mu)$ a measurable space where $\mu$ is a positive and bounded measure, and $(X_i, Y_i)_{i \in \mathbb{N}}$ a series of pairs of random variable defined on $(\Omega, A, P)$ with values in $(E \times R, \mathcal{E} \times B_R)$. We suppose that for all $i \in \mathbb{N}$ $Y_i$ is integrable, that there is a real function $r$ defined on $E$ and satisfying

$$E(Y_i/X_i) = r(X_i) \quad \forall i \in \mathbb{N} \hspace{2cm}(2.1)$$

and we consider the problem of estimating this regression $r$.

We will study the pointwise convergence of the sequence of estimators $(r_n)_{n \in \mathbb{N}}$, defined by

$$r_n(x) = \begin{cases} \sum_{i=1}^n Y_i K_n(x, X_i) / \sum_{i=1}^n K_n(x, X_i) & \text{for } \sum_{i=1}^n K_n(x, X_i) \neq 0 \\ 0 & \text{otherwise} \end{cases}$$

(2.2)

where $x$ is a fixed member of $E$ and $(K_n)_{n \in \mathbb{N}}$ is a sequence of real measurable functions defined on $E \times E$, and subject to two assumptions which will be specified in section 3 below.

When $E = R^p$ and the pairs $(X_i, Y_i)_{i \in \mathbb{N}}$, are identically distributed, we wish to estimate the regression $r$ of $Y_i$ on $X_i$; this problem has been widely studied $(X_i, Y_i)_{i \in \mathbb{N}}$, see the review of bibliography by Collomb (1981, 1985a), Ferraty et al. (2002) and Härdle (1991) and the most usual nonparametric estimator of $r$ is then the Nadaraya-Watson estimator defined by (2.2) with

$$K_n(x, z) = h_n^{-p} K \left( \frac{x - z}{h_n} \right)$$

(2.3)

where $K$ is a kernel of $R^p$.

The properties of almost sure convergence of such estimators were extended to the case of pairs $(X_i, Y_i)_{i \in \mathbb{N}}$, which are not necessarily independent but form a uniformly strongly mixing process (Collomb, 1984, 1985b; Collomb et al., 1986a,b; Härdle and Vieu, 1990). Some of the latter results have been extended in this article to estimators of $r(x)$ defined with the help of sequences $(K_n(x_i))_{n \in \mathbb{N}}$, which are more general than (2.3), and for random variables $(X_i)_{i \in \mathbb{N}}$, taking values in a space $E$ which is not necessarily $R^p$.

Let $(X_n)_{n \in \mathbb{N}}$ have values in $E = \{1, 2, \ldots, m\}$, $m \in \mathbb{N}$, and suppose that

$$P(X_{n+1} = j / X_n = i) = \pi(i, j), \forall (i, j) \in E, \forall n \in \mathbb{N}$$

which is satisfied whenever $(X_n)_{n \in \mathbb{N}}$, is a stationary
Markov process.

We estimate the transition probability

\[ \pi(i,j) = E \left[ 1_{\{X_2=j\}} / X_1 = i \right] \]

by

\[ \pi_n(i,j) = \frac{\sum_{k=1}^{n-1} 1_{\{X_k=i\}} 1_{\{X_{k+1}=j\}}}{\sum_{k=1}^{n-1} 1_{\{X_k=i\}}} \]

and in this case

\[ K_n(x,z) = 1_{\{x=z\}} \]

The generalization (2.2) of the Nadaraya-Watson estimator allows us to handle the case of the regresogram found by taking

\[ K_n(x,z) = \begin{cases} \frac{1}{h_n} & \text{if } z \in \prod_{i=1}^{p} \left[ \frac{x_i}{h_n}, \frac{x_i}{h_n} \right] + 1 \\ 0 & \text{otherwise} \end{cases} \]

\[ \forall z \in \mathbb{R}^p, \forall x = (x_1, \ldots, x_p) \in \mathbb{R}^p \text{ where } \lfloor x \rfloor \text{ is the integer part of } x. \]

3. Assumptions

We now take up the assumptions which were mentioned above (Section 2). We suppose:

(Assumption A1) The probability law of \( X_i \) has a density \( f_i \) with respect to the measure \( \mu \) such that:

\[ \exists \gamma > 0, \Gamma > 0 : \gamma < f_i < \Gamma, \forall i \in \mathbb{N} \]

\[ \exists M > 0, |Y_i| < M, \forall i \in \mathbb{N} \]

(Assumption A2) The process \((X_i, Y_i)_{i \in \mathbb{N}}\), is uniformly strongly mixing (or \( \phi \)-mixing) (Billingsley, 1968) i.e:

for all positive integers and \( k \) and for all events \( A \) (or \( B \), respectively) in the \( \sigma \)-algebra collection

\[ ((X_1, Y_1), \ldots, (X_i, Y_i)) \]

(or by \((X_{i+k}, Y_{i+k}), (X_{i+k+1}, Y_{i+k+1}), \ldots\)), respectively, we have:

\[ |P(A \cap B) - P(A)P(B)| \leq \phi_k P(A) \]

where \((\phi_n)_{n \in \mathbb{N}}\) is a real convergent sequence tending to zero.

We can suppose, without affecting the results, that the sequence \((\phi_n)_{n \in \mathbb{N}}\) is decreasing.

(Assumption A3) The functions \( K_n \) satisfy the following assumptions, for fixed \( x \) in \( E \).

\[ \forall x \in E \text{ and } \forall n \in \mathbb{N}, \int K_n(x,z) \mu(dz) = 1 \]

\[ \forall n \in \mathbb{N}, K_n \text{ is strictly positive on } E \times E \]

There exists a constant \( C \) such that for all \( n \in \mathbb{N} \) and all \((x,y)\) in \( E \times E \).

\[ K_n(x,y) \leq C \int K_n^2(x,) d\mu \]

Let \( \psi \) be a real measurable function defined on \( E \).

We will say that the sequence \((K_n)_{n \in \mathbb{N}}\), and the function \( \psi \) "satisfy the condition C", if for all \( x \) in \( E \),

\[ \int \psi(.) K_n(x,) d\mu \rightarrow \psi(x) \]

Remark:

When \( E = \mathbb{R}^p \) and \((K_n)_{n \in \mathbb{N}}, \) is a sequence of kernels defined in the following way from a kernel with compact support

\[ \forall n \in \mathbb{N}, K_n(x,)=h_n^{-1}K\left(\frac{x-\cdot}{h_n}\right) \]

the condition C is true for all continuous functions \( \psi \).

4. Nonparametric Estimation of Willingness To Pay

Suppose that we seek to predict the Willingness To Pay \( y \) based on geolocation \( x \) expressed by the latitude and longitude, given that we have a sample of
pairs \((X_i, Y_i) : i = 1, \ldots, n\) with

\[
\begin{align*}
X_i & : \text{the pair (latitude, longitude) i} \\
Y_i & : \text{the price that the individual is willing to pay}
\end{align*}
\]

the individuals \(i = 1, \ldots, n\) are in general relatively close to one another, so the pairs \((X_i, Y_i), i = 1, \ldots, n\) are not independent.

Nevertheless it is reasonable to suppose that for sufficiently separated individuals there is "almost independence", which is expressed by the \(\phi\)-mixing condition.

In addition we suppose that the pairs \((X_i, Y_i), i = 1, \ldots, n\) are identically distributed and writing \((X, Y)\) for a pair with the same law, it is natural to predict \(y\) from \(x\) by

\[
y = r(x) = E(Y / X = x)
\]

We are therefore estimating the regression of a real random variable \(Y\) with another random variable \(X\) taking values in the space \(E\) which we can easily consider to be metric. \(E\) is the unit sphere with the distance

\[
d(x, z) = \inf \{arc(x, z), 2\pi - arc(x, z)\}
\]

where \(arc(x, z)\) is the arc of the greatest circle passing through \(x\) and \(z\).

5. Properties of Pointwise Convergence of the Estimator of Willingness To Pay

Let \(i\) be the function of \(N\) in \(N\) associated with \(\phi = (\phi_n)_{n \in N}\) such that

\[
i_\phi(n) = n i(n) = \inf \left\{ j \leq n / \frac{\phi_j}{j} \leq \frac{1}{n} \right\}
\]

**Theorem 1.**

If \(r\) and \(K_n\) satisfy the condition C, and

\[
n^{-1} \int i(n) K_n^2(x, \cdot) d\mu \to 0\]

then, for every fixed point \(x\) in \(E\)

\[
r_n(x) \to P r(x)
\]

**Theorem 2.**

If \(r\) and \(K_n\) satisfy the condition C and if

\[
n^{-1} \ln n i(n) \int K_n^2(x, \cdot) d\mu \to 0\]

then for every fixed point \(x\) in \(E\)

\[
r_n(x) \to P C o r(x)
\]

**Remarks**

When the process \((X_i, Y_i)_{i \in N}\), is m-dependent we can choose the sequence with \(i_\phi\) constant in which case the conditions (5.2) and (5.4) become

\[
n^{-1} \int K_n^2(x, \cdot) d\mu \to 0 \quad \text{and} \quad n^{-1} \ln(n) \int K_n^2(x, \cdot) d\mu \to 0.
\]

When \((X_i, Y_i)_{i \in N}\), is Markovian and stationary, Rosenblatt (1971) show that if in addition the process satisfies Doeblin's condition or the usual condition

\[L^p \text{ for } p = 1 \text{ or } p = \infty, \text{ then } i(n) \geq \ln(n).\]

6. Proofs of Convergences of the Estimator of WTP

6.1. A Preliminary Result

We show a preliminary result, which is an extension of Bernstein's inequality to \(\phi\)-mixing random variables (Renyi, 1956), established by Collomb (1984).

**Lemma 6.1.**

For all \(n \in N\), let \(\Delta_i = \Delta_{ni}, i \in N\) be a \(\phi\)
-mixing sequence of standardized random variables satisfying:

\[ P\left(\left|\Delta_i\right| \leq d\right) = 1, \quad E\left(\Delta_i\right) \leq \delta, \quad \text{and} \quad E\left(\Delta_i^2\right) \leq D \]  \hspace{1cm} (6.1)

Writing

\[ \phi_m = \sum_{i=1}^{m} \phi_i \quad \forall m \in N \]  \hspace{1cm} (6.2)

We then have, for all \( \varepsilon > 0 \) and every integer \( n \),

\[ P\left(\sum_{i=1}^{m} \Delta_i > \varepsilon\right) \leq C_0 \exp(-\alpha \varepsilon + \alpha^2 n C_1) \]  \hspace{1cm} (6.3)

Where

\[ C_1 = 6(D + 4\delta d \phi_m) \]  \hspace{1cm} and \( C_0 = 2 \exp\left(3\sqrt{e n \phi_m / m}\right) \]  \hspace{1cm} (6.4)

And \( \alpha \) and \( m \) are respectively a positive (real) number and an integer less than \( n \) satisfying

\[ \alpha m \leq \frac{1}{4} \]  \hspace{1cm} (6.5)

(The numbers \( \alpha \) and \( m \), as well as \( d \), \( \alpha \), \( \delta \) and \( D \), may depend on \( n \)).

6.2 Proof of Results

In order to simplify the notation, we write \( r_n \) as defined in (2.2):

\[ r_n(x) = \frac{R_{1,n}(x)}{R_{0,n}(x)} \]

Where, for \( j=0,1 \),

\[ R_{j,n}(x) = n^{-1} \sum_{i=1}^{n} Y_i^j K_n(x, X_i) \]

We can write

\[ r_n(x) - r(x) = \frac{\left(R_{1,n}(x) - ER_{1,n}(x)\right) - r(x)\left(R_{0,n}(x) - ER_{0,n}(x)\right) + G}{R_{0,n}(x)} \]

Where \( G = ER_{1,n}(x) - r(x)ER_{0,n}(x) \)

Theorem 1 and Theorem 2 are then an immediate consequence of the following Lemmas.

Results (6.7) (and (6.8)) of Lemma 6.3 imply convergence in probability (and almost complete, respectively) of the two first terms of the sum, towards 0.

The asymptotic lower bound in probability (and almost complete) of the random variable \( R_{0,n}(x) \) by a positive non-zero number is given by Lemma 6.4.

Convergence towards zero of terms of the sum is given by Lemma 6.5.

**Lemma 6.2**

For all \( \varepsilon > 0 \), and for \( j = 0,1 \), we have:

\[ P\left(\left|R_{j,n}(x) - ER_{j,n}(x)\right| > \varepsilon\right) \leq a \exp(H) \]  \hspace{1cm} (6.6)

where

\[ H = \frac{-bn}{i(n)\int K_n^2(x,z) \mu(dz)} \]

where \( n \) is a sufficiently large integer, where \( a \) and \( b \) are positive constants, independent of \( n \) and \( i(n) \).

**Lemma 6.3**

If assumptions (5.2) and (5.4), are satisfied, then we have, respectively:

\[ R_{j,n}(x) - ER_{j,n}(x) \xrightarrow{P} 0 \]  \hspace{1cm} (6.7)

And

\[ R_{j,n}(x) - ER_{j,n}(x) \xrightarrow{P.C} 0 \]  \hspace{1cm} (6.8)

**Lemma 6.4**

Under the hypotheses of Theorem 1, and of Theorem 2, we have respectively:

\[ \exists \delta > 0 : P\left(R_{0,n}(x) \leq \delta\right) \xrightarrow{} 0 \]

And

\[ \exists \delta > 0 : \sum_{n=1}^{\infty} P\left(R_{0,n}(x) \leq \delta\right) < +\infty \]
Lemma 6.5
If condition C is satisfied for $\psi() = |r() - r(x)|$
then we have
$$\lim_{n \to +\infty} ER_{1,n}(x) - r(x)ER_{0,n}(x) = 0$$

7. Other Applications

To illustrate the interest of extending the results in this way, we give some examples from applied statistics, for which these methods can be used; for example to estimate the regression of a real random variable $Y$ with $X$:

7.1 In Meteorology

Suppose that we seek to predict the next day's rainfall $y$ based on one's position $x$ expressed by the latitude and longitude from the center of the Atlantic anticyclone, given that we have a sample of pairs $(X_i, Y_i), i = 1, \cdots, n$ with

\[
\begin{align*}
X_i & : \text{the pair (latitude, longitude) from the center} \\
Y_i & : \text{the quantity of rain falling on the next day}
\end{align*}
\]

Since the days $i = 1, \cdots, n$ are in general consecutive, the pairs $(X_i, Y_i), i = 1, \cdots, n$ are not independent. Nevertheless it is reasonable to suppose that for sufficiently separated days there is "almost independence", which is expressed by the $\phi$-mixing condition. In addition we suppose that the pairs $(X_i, Y_i), i = 1, \cdots, n$ are identically distributed and writing $(X, Y)$ for a pair with the same law, it is natural to predict $y$ from $x$ by

$$y = r(x) = E(Y / X = x)$$

We are therefore estimating the regression of a real random variable $Y$ with another random variable $X$ taking values in the space $E$ which we can easily consider to be metric. $E$ is the unit sphere with the distance

$$d(x, z) = \inf \{arc(x, z), 2\pi - arc(x, z)\}$$

where $arc(x, z)$ is the arc of the greatest circle passing through $x$ and $z$.

In the same domain (i.e. meteorology), $X$ could also be the pair (wind-force, wind-direction).

$X$ takes its values on the surface of a cylinder, with unit base.

In this case, we can use

$$d(x, z) = \sup \{\|\theta_2 - \theta_1\|, \|h_2 - h_1\|\}$$

for the metric.

7.2 In Medicine

$X$ is a temperature graph (which we suppose is observed "continuously") for a patient during a time interval $T$; $E$ is the set of real functions defined on $T$. In this case we can take one of the functions

$$d(x, z) = \int_T |x(t) - z(t)|^2 dt \text{ or } d(x, z) = \sup_{t \in T} |x(t) - z(t)|$$

as metrics. We can easily use the estimation method described in (2.2) for the examples cited above. For example we could take the sequences $(K_n)_{n \in \mathbb{N}}$ defined by

$$K_n(x, z) = 1\{d(x, z) \leq h_n\}$$

where $(h_n)_{n \in \mathbb{N}}$ is a real and strictly positive sequence.

8. Conclusion

For a nonparametric and mixing estimator of the WTP, we have shown almost complete convergence and convergence in probability.

From the analysis of a survey of 300 individuals residing in Seine Maritime administrative region, we found that the WTP does not follow a known distribution. Nonparametric estimation methods allow us to avoid this problem for large databases and
therefore in the particular case studied here.

The applications cited as examples are not exhaustive - the methods could be used in many other domains such as geophysics, bioclimatology, agroforestry, and so on. The smoothness parameter can be estimated with cross validation technique (Taïbi-Hassani and Youndjé 2003).

Our results can easily be extended to situations with independence, and to other nonparametric estimators such as those obtained using the k-nearest neighbours method.

References

[1] B. Alvarez-Farizo, J.M. Gil, and B.J. Howard, Impacts from restoration strategies: Assessment through valuation workshops, Ecological Economics 68 (2009), pp. 787-797.

[2] O. Beaumais, D. Laroutis, and R. Chakir, Conservation versus conversion des zones humides : une analyse comparative appliquée à l'estuaire de la Seine, Revue d'Economie Régionale et Urbaine 4 (2008), pp. 565-590.

[3] W. Bekele, and L. Drake, Soil and water conservation decision behaviour of substance farmers in the Eastern Highlands of Ethiopia: a case study of the Hunde-Lafto area, Ecological Economics 46 (2003), pp. 437-451.

[4] P. Billingsley, Convergence of probability measures. Wiley (1968).

[5] P. Calkins, B. Larue, and M. Vezina, Willingness to Pay for Drinking Water in the Sahara: The Case of Douentza in Mali, Cahiers d'économie et sociologie rurales 64 (2002), pp.38-56.

[6] M. Christie, N. Hanley, J. Warren, K. Murphy, R. Wright, and T. Hyde, Valuing the diversity of biodiversity, Ecological Economics 58 (2006), pp.304-317.

[7] G. Collomb, Estimation non paramétrique de la régression : Revue bibliographique, International Statistical Review 49 (1981), pp.75-93.

[8] G. Collomb, Propriétés de convergence presque complète du prédicteur à noyau, ZWahrscheinlichkeitsstheorie und V.Gebiete 66 (1984), pp.442-460.

[9] G. Collomb, Non parametric regression: an up to date bibliography. Mathematische Operationssonforschung un Statistik, series Statistics, 2 (1985b), pp. 297-307.

[10] G. Collomb, W. Härdle, and S. Hassani, A note on prediction via estimation of the conditional mode function, Journal of Statistical Planning and Inference 15 (1986a), pp. 227-236.

[11] G. Collomb, S. Hassani, P. Sarda, and P. Vieu, Approche non paramétrique en théorie de la fiabilité : revue bibliographique. Revue de Statistique Appliquée, vol 35, 4 (1986b), pp. 27-41.

[12] J.R. Crooker, and J.A. Herriges, Parametric and Semi-Nonparametric Estimation of Willingness-to-Pay in the Dichotomous Choice Contingent Valuation Framework, Environmental and Resource Economics 27 (2004), pp. 451-480.

[13] F. Ferraty, A. Goia, and P. Vieu, Régession non-paramétrique pour des variables aléatoires fonctionnelles mélangeantes : Nonparametric regression for mixing functional random variables in C. R. Acad. Sci. Paris, Volume 334, Issue 3 (2002), pp. 217-220.

[14] G. Collomb, W. Härdle, and S. Hassani, A note on prediction via estimation of the conditional mode function, Journal of Statistical Planning and Inference 15 (1986a), pp. 227-236.

[15] G. Collomb, W. Härdle, and S. Hassani, A note on prediction via estimation of the conditional mode function, Journal of Statistical Planning and Inference 15 (1986a), pp. 227-236.
Pointwise Convergence of a Nonparametric Estimator of Regression in a Measurable Space Used in Contingent Valuation Method

[22] A. Renyi, Probability Theory, American Elsevier Publishing Company, Inc., New York (1970).

[23] M. Rosenblatt, Markov processes- Structure and asymptotic behavior, Springer, Berlin (1971).

[24] C.L. Spash, K. Urama, R. Burton, W. Kenyon, P. Shannon, and G. Hill, Motives behind willingness to pay for improving biodiversity in a water ecosystem: Economics, ethics and social psychology, Ecological Economics 68 (2009), pp. 955-964.

[25] S. Taibi-Hassani, and E. Youndjé, Validation croisée pour l'estimateur lissé de la fonction de hasard : cas des données censurées, Revue de Statistique Appliquée 51 (2003), pp. 73-86

[26] X. Wang, J. Bennet, C. Xie, Z. Zhang, and D. Liang, Estimating non-market environmental benefits of the Conversion of Cropland to Forest and Grassland Program: A choice modelling approach, Ecological Economics 63 (2007), pp. 114-125

[27] W. Yang, J. Chang, B. Xu, C. Peng, and Y. Ge, Ecosystem service value assessment for constructed wetlands: A case study in Hanqzhou, China, Ecological Economics 68 (2008), pp. 116-125.