

Functional local linear estimation for the relative regression under left-truncation

Belkais Altendji¹

¹Department of Mathematics
Katholieke University Leuven
Celestijnenlaan 200B bus 02.32

belaltendji@yahoo.com
belkais.altendji@kuleuven.be

ABSTRACT

The focus of this paper is to set up asymptotic properties of the constructed estimator. These results are obtained under some standard conditions in NFDA where we prove the pointwise and the uniform almost complete consistencies of the local linear relative error regressor under left-truncated data.

Keywords: Truncated data; Local linear estimation; Missing data; Censored data; Functional regression; Relative error.

2000 Mathematics Subject Classification: 62G05, 62G08, 62G20, 62G35, 62G07, 62G32, 62G30 secondary 62E20

1 Introduction

The focal point in this present project is the concept of the nonparametric estimation of a scalar response on a functional explanatory variable by considering using a local linear (LL) regression functions. We assume that the covariate X and the response variable Y are linked by:

$$Y = R(X) + \epsilon, \quad (1.1)$$

where the operator $R : \mathcal{F} \rightarrow \mathbb{R}$ and the random error $\epsilon \perp\!\!\!\perp X$.

By combining the ideas of the relative error regression in Altendji et al. (2017) and those of the local linear approach introduced by Berrientos-Marin et al. (2010) we will construct a new estimator of the regression operator.

The local linear method in functional setting has been introduced by Bailo and Gran (2009) where they studied the consistency of the local linear estimate of the regression function when X is functional. The fast version of the functional local linear estimate was proposed by Berrientos-Marrin et al. (2010). For a local linear estimate, the local covariance's inverse of the functional setting was established by Berline et al. (2011). We refer to Zhou and Lin (2016), they gave the asymptotic normality of the functional local linear regression estimate. amongst umpteen papers on one relative error regression and the other, the local linear estimation have been studied in the multivariate case, we refer to A. Chahad, Ait-Hennani. L and Laksaci. A, (2017); Yang and Ye (2013); Hallin et al. (2009); Jones et al. (2008); Masry (1996); Narula and Wellington (1977); e.g., Stone (1977).

The paper is coordinated as follows. Section 2 is devoted to the construction of the model and the formulation of the estimator of R . The needed assumptions and main results are given in Section 3. The uniform version of consistency is studied in section 4. Finally, the proofs of the main results are given in the last section. we fellow decomposition below

2 The model and its estimator

Following the paper of Altendji et al. (2017) we have our relative error estimator:

The relative error estimator (REE). In order to get the estimator of the operator R , the minimization of the following loss function is required

$$\mathbb{E} [(Y - R(X))^2 | X] .$$

The fact that this kind of loss function is sensitive, we go towards the estimation of the operator R . based on the mean squared relative error we get:

$$\mathbb{E} \left[\left(\frac{Y - R(X)}{Y} \right)^2 | X \right] , \quad \text{for } Y > 0. \quad (2.1)$$

The solution of the minimization of (2.1) is:

$$R(x) = \frac{\mathbb{E}[Y^{-1} | X = x]}{\mathbb{E}[Y^{-2} | X = x]} .$$

Notations. For $i = 1, \dots, N$ let (X_i, Y_i) be N independent and identically distributed (i.i.d.) copies of (X, Y) which is valued in $\mathcal{F} \times \mathbb{R}$ with the sample size is deterministic but unknown.

Among N we do not observe all data, we observe only n variables, we suppose that the response observations Y_i is left-truncated by T_i , such that Y and T are observable only if $Y \geq T$. Furthermore, from the i.i.d (X_i, Y_i, T_i) we construct our estimator where $i = 1, \dots, n$. For $\tau > 0$, notice that n is a binomial random variable of parameters:

$$N : \begin{cases} \tau := \mathbb{P}(Y \geq T) & \text{if } n \text{ data are observed,} \\ 0 & \text{if no data can be observed.} \end{cases}$$

The estimation of the distribution functions of F and G . We suppose that f and G are respectively the distribution functions of the response variable Y and the truncating variable T . Regarding to Stute (1993), the estimator of the truncated relative error regressor of R is defined by Lynden-Bell estimator G_n .

$$F^*(y) := \tau^{-1} \int_{-\infty}^y G(v) dF(v) \quad \text{empirically estimated by} \quad F_n^*(y) = n^{-1} \sum_{i=1}^n \mathbb{1}_{\{Y_i \leq y\}}$$

Thus, regarding to Lynden-Bell (1971), the estimator of the distribution function of F is given by

$$F_n(y) = 1 - \prod_{s \leq y} \left[1 - \frac{F_n^*(s)}{R_n(s)} \right]$$

and

$$G^*(t) := \tau^{-1} \int_{-\infty}^{\infty} G(t \wedge v) dF(v) \quad \text{empirically estimated by} \quad G_n^*(y) = n^{-1} \sum_{i=1}^n \mathbb{1}_{\{T_i \leq y\}}$$

and the Lynden-Bell (1971) estimator of the distribution function of G is given by:

$$G_n(y) = 1 - \prod_{s > y} \left[1 - \frac{G_n^*(s)}{R_n(s)} \right].$$

Then, we define

$$R(y) := \tau^{-1} G(y) [1 - F(y)] \quad \text{empirically estimated by} \quad R_n(y) = n^{-1} \sum_{i=1}^n \mathbb{1}_{\{T_i \leq y \leq Y_i\}}.$$

Then, regarding to He and Wang (1998), τ can be estimated by

$$\tau_n := \frac{G_n(y)(1 - F_n(y))}{R_n(y)}$$

we notice that τ_n is independent of y .

The local linear technique. We follow the version proposed by Berrientos-Marin et al. (2010) and we involve the loss function (2.1). Especially, for a fixed point $x \in \mathcal{F}$, we suppose that,

$$\forall x' \text{ in neighborhood of } x : \quad R(x') = a + b\beta(x, x') + o(\beta(x, x'))$$

thus,

$$(\hat{a}, \hat{b}) = \arg \min_{a, b} \sum_{i=1}^n \frac{(Y_i - a - b\beta(X_i, x))^2}{Y_i^2} K(h^{-1}\delta(X_i - x)) \quad (2.2)$$

where $\beta(\cdot, \cdot)$ is a known operator from $\mathcal{F} \times \mathcal{F}$ into $\mathbb{R}$ such that, $\forall \xi \in \mathcal{F}, \beta(\xi, \xi) = 0$, with K a kernel and $h = h_{K,n}$ a bandwidth and $\delta(\cdot, \cdot)$ a function defined on $\mathcal{F} \times \mathcal{F}$ such that $d(\cdot, \cdot) = |\delta(\cdot, \cdot)|$.

We can get easily the solution of Eq. (2.2).

So,

$$\begin{bmatrix} {}^t Q_B \mathbf{M} Q_B \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} - [{}^t Q_B \mathbf{M} Y] = 0$$

which implies:

$$\begin{bmatrix} a \\ b \end{bmatrix} = [{}^t Q_B \mathbf{M} Y] [{}^t Q_B \mathbf{M} Q_B]^{-1}$$

where

$${}^t Q_B = \begin{bmatrix} 1 & \dots & 1 \\ B(X_1, x) & \dots & B(X_n, x) \end{bmatrix}$$

$\mathbf{M} = \text{diag}(Y_1^{-2}K(h^{-1}\delta(x, X_1)), \dots, Y_n^{-2}K(h^{-1}\delta(x, X_n)))$ and ${}^t Y = (Y_1, \dots, Y_n)$

Thus, we get easily :

$$\hat{a} = [{}^t Q_B \mathbf{M} Y] [{}^t Q_B \mathbf{M} Q_B]^{-1} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \hat{b} = [{}^t Q_B \mathbf{M} Y] [{}^t Q_B \mathbf{M} Q_B]^{-1} \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

So,

$$\hat{R}(x) = \hat{a} = \frac{\sum_{i=1}^n \sum_{j=1}^n G_n^{-1}(Y_i) w_{ij}(x) Y_j}{\sum_{i=1}^n \sum_{j=1}^n G_n^{-1}(Y_i) w_{ij}(x)} \quad (2.3)$$

where

$$w_{ij}(x) = \beta(X_i, x)(\beta(X_i, x) - \beta(X_j, x))K(h^{-1}\delta(x, X_i))K(h^{-1}\delta(x, X_j))Y_i^{-2}Y_j^{-2}$$

with the condition $0/0 = 0$.

We can carry out this notation $\beta_i = \beta(X_i, x)$ and $K_i = K(h^{-1}\delta(x, X_i))$

3 Main results of pointwise almost complete convergence

To make reading easier, all proofs are given in the appendix.

3.1 Asymptotic properties

In order to state our main first result, we base on some assumptions.

(H1) $\mathbb{P}(X \in B(x, r)) =: \phi_x(r) > 0$ for all $r > 0$ and $\lim_{r \rightarrow 0} \phi_x(r) = 0$. with

$$B(x, r) := \{x' \in \mathcal{F} : d(x', x) < r\}.$$

(H2) For all $(x_1, x_2) \in \mathcal{N}_x \times \mathcal{N}_x$, we have:

$$|g_\gamma(x_1) - g_\gamma(x_2)| \leq C d^{k_\gamma}(x_1, x_2) \text{ for } k_\gamma > 0,$$

where $\mathcal{N}_x$ denotes a neighborhood of the fixed $x \in \mathcal{F}$.

(H3) The function $\beta(\cdot, \cdot)$ is such that:

$$\forall x' \in \mathcal{F}, \quad C | \delta(x, x') | \leq | \beta(x, x') | \leq C' | \beta(x, x') |.$$

(H4) The kernel K is a measurable function which is supported within $(-1, 1)$ and such that:

$$0 < C \leq K(\cdot) \leq C' < \infty.$$

(H5)

$$h \int_{B(x, h)} \beta(u, x) d\mathbb{P}(u) = o \left(\int_{B(x, h)} \beta(u, x)^2 d\mathbb{P}(u) \right)$$

where $d\mathbb{P}(x)$ is the cumulative distribution of X .

(H6) The bandwidth h satisfies

$$\lim_{n \rightarrow \infty} h = 0 \quad \text{and} \quad \lim_{n \rightarrow \infty} \frac{\log n}{n\phi_x(h)} = 0.$$

(H7) The inverse conditional moments of the response variable verify:

$$\forall m \geq 2, \mathbb{E}[Y^{-m}|X = x] < C < \infty.$$

In this context of nonparametric regression analysis in FDA, we assume that all our asymptotic studies and all the considered conditions.

Specifically, Assumptions (H1) and (H4)-(H6) are the same as that used by Berrientos-Marin et al. (2010) while (H1) which is linked to the functional structure of the functional covariate. The conditions (H2) and (H7) characterize the infinite space of the nonparametric model witch also used by Dmeongeot et al. (2016).

3.2 The result

Theorem 3.1. *Under assumptions (H1)-(H7), we have:*

$$\left| \hat{R}(x) - R(x) \right| = O(h^{b_1}) + O\left(\sqrt{\frac{\log n}{n\phi_x(h)}}\right), \quad a.co \quad (3.1)$$

where $b_1 = k_1 \wedge k_2$

So, to demonstrate the result we base on the following decomposition.

$$\hat{R}(x) - R(x) = \frac{1}{\hat{f}(x)} [\hat{g}(x) - g_1(x)g_2(x)] + \frac{R(x)}{\hat{f}(x)} [g_2^2(x) - \hat{f}(x)] \quad (3.2)$$

where

$$\hat{g}(x) = \frac{1}{n(n-1)\mathbb{E}(w_{12})} \sum_{i \neq j} G_n^{-1}(Y_i) w_{ij} Y_j \quad \text{and} \quad \hat{f}(x) = \frac{1}{n(n-1)\mathbb{E}(w_{12})} \sum_{i \neq j} G_n^{-1}(Y_i) w_{ij}$$

with:

$$w_{12} = \beta_1(\beta_1 - \beta_2) K_1 K_2$$

So, Theorem 3.1's result is a consequence of the following Lemmas,

Lemma 3.2. *Under assumptions (H1) and (H3)-(H7), we get:*

$$\left| \hat{f}(x) - \mathbb{E}[\hat{f}(x)] \right| = O_{a.co} \left(\sqrt{\frac{\log n}{n\phi_x(h)}} \right)$$

and

$$|\hat{g}(x) - \mathbb{E}[\hat{g}(x)]| = O_{a.co} \left(\sqrt{\frac{\log n}{n\phi_x(h)}} \right).$$

Lemma 3.3. *Under assumptions (H1)-(H6), we have:*

$$\left| \mathbb{E}[\hat{f}(x)] - g_2^2(x) \right| = O\left(h^{k_2}\right)$$

and

$$|\mathbb{E}[\hat{g}(x)] - g_2(x)g_1(x)| = O\left(h^{k_1}\right).$$

Corollary 3.4. *Under Theorem 3.1's assumptions, we obtain:*

$$\sum_{n=1}^{\infty} \mathbb{P}\left(\hat{f}(x) < \frac{g_2^2(x)}{2}\right) < \infty$$

4 Uniform almost complete convergence

Now we focus on the study of the uniform almost complete convergence of $\hat{R}(x)$ on some subset $S_{\mathcal{F}}$ of $\mathcal{F}$. Under the following assumptions we get the result.

4.1 Asymptotic properties

(U1) There exists a differentiable function $\phi(\cdot)$, such that

$$\forall x \in S_{\mathcal{F}}, \quad 0 < C\phi(h) \leq \phi_x(h) \leq C'\phi(h) < \infty \quad \text{and} \quad \exists \nu_0 > 0, \forall \nu < \nu_0, \quad \phi'(\nu) < C,$$

where ϕ' denotes the first derivative of ϕ .

(U2) $\exists \nu > 0$, such that: $\forall x, x' \in S_{\mathcal{F}}^{\nu}, \quad |g_{\gamma}(x) - g_{\gamma}(x')| \leq C d^{k_{\gamma}}(x, x')$,

where

$$S_{\mathcal{F}}^{\nu} = \{x \in \mathcal{F}, \exists x' \in S_{\mathcal{F}}, d(x, x') \leq \nu\}.$$

(U3) The function $\beta(\cdot, \cdot)$ is such that:

$$\forall x' \in \mathcal{F}, \quad Cd(x, x') \leq |\beta(x, x')| \leq C' |\beta(x, x')|$$

and

$$\forall (x_1, x_2) \in S_{\mathcal{F}}^2, \quad |\beta(x_1, x') - \beta(x_2, x')| \leq Cd(x_1, x_2).$$

(U4) The kernel K satisfies (H4) and, Lipschitz condition as follow:

$$|K(x) - K(y)| \leq C |x - y|.$$

(U5) For all $x \in S_{\mathcal{F}}$, the condition (H5) is verified.

(U6) The subset $S_{\mathcal{F}}$ is such that:

$$S_{\mathcal{F}} \subset \bigcup_{k=1}^{d_n} B(x_k, r_n),$$

where $x_k \in \mathcal{F}$, $r_n = O(\frac{\log n}{n})$ and the sequence d_n satisfies

$$\frac{(\log n)^2}{n\phi(h)} < \log d_n < \frac{n\phi(h)}{\log n} \quad \text{and} \quad \sum_{n=1}^{\infty} d_n^{1-\beta} < \infty \quad \text{for some } \beta > 1.$$

(U7) The inverse conditional moments of the response variable verify:

$$\forall m \geq 2, \mathbb{E}[Y^{-m}|X = x] < C < \infty, \quad \forall x \in S_{\mathcal{F}} \quad \text{and} \quad \inf_{x \in S_{\mathcal{F}}} g_2(x) \geq C' > 0.$$

4.2 The result

Theorem 4.1. Assume that assumptions (U1)-(U7) hold, then, we have:

$$\sup_{x \in S_{\mathcal{F}}} |\hat{R}(x) - R(x)| = O(h^{b_1}) + O\left(\sqrt{\frac{\log d_n}{n\phi(h)}}\right), \quad a.co$$

Clearly, to proof it we can base on the decomposition 3.2. Therefore, the Theorem 4.1 is a consequence of the convergence rate of Lemmas 3.2 and 3.3 and the corollary 3.4 and the following results.

Lemma 4.2. Under the hypotheses (U1) and (U3)-(U7), we have:

$$\sup_{x \in S_{\mathcal{F}}} |\hat{f}(x) - \mathbb{E}[\hat{f}(x)]| = O_{a.co}\left(\sqrt{\frac{\log d_n}{n\phi_x(h)}}\right)$$

and

$$\sup_{x \in S_{\mathcal{F}}} |\hat{g}(x) - \mathbb{E}[\hat{g}(x)]| = O_{a.co}\left(\sqrt{\frac{\log d_n}{n\phi_x(h)}}\right).$$

Lemma 4.3. Under the hypothesis (U1)-(U6), we obtain:

$$\sup_{x \in S_{\mathcal{F}}} |\mathbb{E}[\hat{f}(x)] - g_2^2(x)| = O(h^{k_2})$$

and

$$\sup_{x \in S_{\mathcal{F}}} |\mathbb{E}[\hat{g}(x)] - g_2(x)g_1(x)| = O(h^{k_1}).$$

Corollary 4.4. Under Theorem 4.1's hypotheses, we obtain:

$$\exists C > 0, \quad \sum_{n=1}^{\infty} \mathbb{P}\left(\inf_{x \in S_{\mathcal{F}}} \hat{f}(x) < C\right) < \infty$$

Acknowledgments

I am the author, Belkais ALTENDJI greatly thank an associate editor and an anonymous referee for a careful reading of the article.

References

- [1] Altendji. B. J. Demongeot, A. Laksaci., and M. Rachdi. (2017). Functional data analysis: Estimation of the relative error in functional regression under random left-truncation model. *Journal of Nonparametric Statistics*. **30**: 472–490.
- [2] Aneiros, G. Bongiorno, E. G. Cao, R. and Vieu, P. (2017). *Functional Statistics and Related Fields*, Springer International Publishing, Cham, Switzerland.
- [3] Baïllo, A and A. Gran. (2009). Local linear regression for functional predictor and scalar response. *Journal of Multivariate Analysis*. **100**:102–11.
- [4] Barrientos-Marrin, J., F. Ferraty., and P. Vieu. (2010). Locally modelled regression and functional data. *Journal of Nonparametric Statistics*. **22**:617–32.
- [5] Berline, A., A. Elamine., and A. Mas. (2011). Local linear regression for functional data. *Annals of the Institute of Statistical Mathematics*. **63**:104–75.
- [6] Benhenni. K., Ferraty. F., Rachdi. M., and Vieu. P. (2007). Nonparametric regression for functional data: Optional local bandwidth choice. *Computational Statistics*. **22**:353–369.
- [7] Boj.F., Delicado. P., and Fortiana. J. (2008). Local linear based on weighted distance-based regression. in *Functional and Operatorial Statistics*, eds. S. Dabo-Niang and F. Ferraty, Heidelberg: Physica-Verlang, pp. 57–64.
- [8] Burba. F., Ferraty. F., and Vieu. P. (2009). k –Nearest neighbor method in functional non-parametric regression. *Journal of Nonparametric Statistics*. **21**:453–469.
- [9] Chen. J., and Zhang. L. (2009). Asymptotic properties of nonparametric M –estimation for mixing functional data. *Journal of Statistics Planning and Inference*, **139**:533–46.
- [10] Chadad. A., Ait-Hennani. L., and Laksaci. A. (2017). Functional local linear estimate for functional relative-error regression. *J. Statistical Theory and Practice*. **11**:771–789.
- [11] Cuevas, A. (2014). A partial overview of the theory of statistics with functional data. *J. Statist. Plann. Inference*, **147**:1–23.
- [12] Demongeot, J., Hamie, A., Laksaci, A., and Rachdi, M. (2016). Relative-error prediction in nonparametric functional statistics: theory and practice. *J. Multivariate Anal.* **46**:261–268.
- [13] Fan. J., Yao. Q., and Tong. H. (1996). Estimation of conditional densities and sensitivity measures in nonlinear dynamical systems. *Biometrika*. **83**:189–206.

- [14] Ferraty, F. and Vieu, P. (2006) *Nonparametric Functional Data Analysis. Theory and Practice*, Springer-Verlag, New York.
- [15] Hallin. M., Lu. Z., and Yu. K. (2009). Local linear spatial quantile regression. *Bernoulli*. **15**:659–86.
- [16] He, S. and Yang, G. (1998). Estimation of the truncation probability in the random truncation model. *Annals of Statistics*, **26**:1011–1027.
- [17] Hsing. T., and Eubank. R. (2015). *Theoretical foundations of functional data analysis, with an introduction to linear operators*. Wiley Series in Probability and Statistics. Chichester, UK: John Wiley & Sons.
- [18] Horrigue, W. and Ould-Said, E. (2011). Strong uniform consistency of a nonparametric estimator of a conditional quantile for censored dependent data and functional regressors. *Random Oper. Stoch. Equ.*, **19**: 131–156.
- [19] Jones. M. C., Park. H., Shin. S., Vines. K., and Jeong. O. (2008). Relative error prediction via kernel regression smoothers. *Journal of Statistics Planning and Inference*. **138**:2889–98.
- [20] Masry. E. (1996). Multivariate local polynomial regression for time series: Uniform strong consistency and rates. *Journal of Time Series Analysis*. **17**:571–99.
- [21] Ramsay. J, and Silverman. B. W. (2002). *Applied functional data analysis: Methods and Case Studies*. New York: Springer-Verlag.
- [22] Ramsay. J, and Silverman. B. W. (2005). *Functional data analysis (2nd ed)*. New York: Springer-Verlag.
- [23] Sarda. P., and Vieu. P. (2000). Kernel regression. In *Smoothing and regression: Approaches, computation and application*, ed. Schimek. M. G, 43–70. Wiley Series in Probability and Statistics. New York. Chichester: Wiley.
- [24] Zhang, J. (2014). *Analysis of variance for functional data*. Monographs on Statistics and Applied Probability, 127. CRC Press, Boca Raton, FL.
- [25] Zhou. Z., and Lin. (2016). Asymptotic normality of locally modelled regression estimator for functional data. *Journal of Nonparametric Statistics*. **28**:116–131.
- [26] Watson. G. S. (1964). Smooth regression analysis. *Sankhya Series A*. **26**:359–372.

5 Appendix. Proofs of the main results

For any $x \in \mathcal{F}$, and $\forall i = 1, \dots, n$ we put:

$$K_i(x) = K(h^{-1}\delta(x, X_i)), \quad \delta_i(x) = \delta(X_i, x) \quad \text{and} \quad \beta_i(x) = \beta(X_i, x).$$

Proof. (Lemma 3.2)

We observe that:

$$\begin{aligned} \hat{f}(x) = & \underbrace{\frac{n^2 h^2 \phi_x^2(h)}{n(n-1)\mathbb{E}[w_{12}]}}_A \left[\underbrace{\left(\frac{1}{n} \sum_{j=1}^n \frac{K_j(x) Y_j^{-2}}{\phi_x(h)} \right)}_{B_1} \underbrace{\left(\frac{1}{n} \sum_{i=1}^n \frac{G_n^{-1}(Y_i) K_i(x) \beta_i^2(x) Y_i^{-2}}{h^2 \phi_x(h)} \right)}_{B_2} \right. \\ & \left. - \underbrace{\left(\frac{1}{n} \sum_{j=1}^n \frac{K_j(x) \beta_j(x) Y_j^{-2}}{h \phi_x(h)} \right)}_{B_3} \underbrace{\left(\frac{1}{n} \sum_{i=1}^n \frac{G_n^{-1}(Y_i) K_i(x) \beta_i(x) Y_i^{-2}}{h \phi_x(h)} \right)}_{B_4} \right] \end{aligned}$$

which allow us to write:

$$\hat{f}(x) - \mathbb{E}(\hat{f}(x)) = A \left[B_1 B_2 - \mathbb{E}[B_1 B_2] - (B_3 B_4 - \mathbb{E}[B_3 B_4]) \right].$$

Moreover, observe that, $\forall i, j = 1, \dots, 4$,

$$\begin{aligned} B_i B_j - \mathbb{E}[B_i B_j] &= (B_i - \mathbb{E}[B_i])(B_j - \mathbb{E}[B_j]) + (B_j - \mathbb{E}[B_j])\mathbb{E}[B_i] \\ &\quad + (B_i - \mathbb{E}[B_i])\mathbb{E}[B_j] + \mathbb{E}[B_i]\mathbb{E}[B_j] - \mathbb{E}[B_i B_j]. \end{aligned}$$

We have to study in each term of the above decomposition. Thus, the following statements lead to the results:

$$\sum_{i=1}^n \mathbb{P} \left\{ |B_i - \mathbb{E}[B_i]| > \nu \sqrt{\frac{\log n}{n \phi_x(h)}} \right\} < \infty \quad \text{for } i = 1, \dots, 4, \quad (5.1)$$

$$A = O(1), \quad \mathbb{E}[B_i] = O(1). \quad (5.2)$$

$$\text{Cov}(B_i, B_j) = o \left(\sqrt{\frac{\log n}{n \phi_x(h)}} \right). \quad (5.3)$$

For Eq.(5.1), let write

$$\Delta_i^{l,k} = \frac{1}{n h_n^l} \sum_{i=1}^n G_n^{-1}(Y_i) \phi_x^{-1}(h) \left[K_i(x) Y_i^{-k} \beta_i^l - \mathbb{E}[K_i(x) Y_i^{-k} \beta_i^l(x)] \right] \quad \text{for } l = 0, 1, 2 \quad \text{and} \quad k = 1, 2.$$

In order to apply an exponential equality, we focalize on the absolute m^{th} moment of Δ_{1i} :

$$\begin{aligned}
 \mathbb{E} \left| \Delta_{1i}^m \right| &= \mathbb{E} \left| \left(G_n^{-1}(Y_i) \phi_x^{-1}(h) \left[K_i(x) Y_i^{-k} \beta_i^l(x) - \mathbb{E}[K_i(x) Y_i^{-k} \beta_i^l(x)] \right] \right)^m \right| \\
 &= \phi_x^{-m}(h) \mathbb{E} \left| \left(G_n^{-1}(Y_i) \left[K_i(x) Y_i^{-k} \beta_i^l(x) - \mathbb{E}[K_i(x) Y_i^{-k} \beta_i^l(x)] \right] \right)^m \right| \\
 &= \phi_x^{-m}(h) \left| \int \left(G_n^{-1}(Y_i) \left[K_i(x) Y_i^{-k} \beta_i^l(x) - \mathbb{E}[K_i(x) Y_i^{-k} \beta_i^l(x)] \right] f^*(y|u) dy d\mathbb{P}_x^X(x) \right)^m \right| \\
 &= \phi_x^{-m}(h) \left| \int \left(\left[K_i(x) Y_i^{-k} \beta_i^l(x) - \mathbb{E}[K_i(x) Y_i^{-k} \beta_i^l(x)] \right] f(y|u) dy d\mathbb{P}_x^X(x) \right)^m \right| \\
 &= \phi_x^{-m}(h) \mathbb{E} \left| \left(\left[K_i(x) Y_i^{-k} \beta_i^l(x) - \mathbb{E}[K_i(x) Y_i^{-k} \beta_i^l(x)] \right] \right)^m \right| \\
 &= \phi_x^{-m}(h) \mathbb{E} \left| \sum_{d=0}^m C_{d,m} \left(K_i(x) Y_i^{-k} \beta_i^l(x) \right)^d \left(\mathbb{E}[K_i(x) Y_i^{-k} \beta_i^l(x)] \right)^{m-d} (-1)^{m-d} \right|
 \end{aligned}$$

with: $C_{k,m} = m!/(k!(m-k)!)$.

Conditioning by X_1 we get:

$$\mathbb{E} \left| \Delta_{1i}^m | X_1 \right| \leq \phi_x^{-m}(h) \sum_{d=0}^m C_{d,m} \mathbb{E} \left[\beta_i^{dl}(x) K_i^d(x) \mathbb{E}[Y_i^{-dk} | X_1] \right] \left| \mathbb{E} \left[\beta_i^l(x) K_i(x) \mathbb{E}[Y_i^{-k} | X_1] \right] \right|^{m-d}$$

Under condition (H7) we have :

$$\mathbb{E}[Y_1^{-kd} | X] = O(1), \quad \text{for all } d \leq m$$

from Lemma A.1's Barrientos-Marin and Ferraty & P.Vieu (2010) we get with $l = 0$,

$$\begin{aligned}
 \mathbb{E} \left| \Delta_{1i}^m \right| &= O \left(\phi_x^{-m}(h) \sum_{d=0}^m \mathbb{E} \left[\beta_i^{dl}(x) K_i^d(x) \right] \left[\mathbb{E}[\beta_i^l(x) K_i(x)] \right]^{m-d} \right) \\
 &= O \left(\max_{d \in \{0, \dots, m\}} \phi_x^{-d+1}(h) \right) \\
 &= O \left(\phi_x^{-m+1}(h) \right).
 \end{aligned}$$

It is sufficient to apply Corrally A.8 in Ferraty and Vieu (2006) with $a_n^2 = (\phi_x)^{-1}$ to get

$$\sum_{i=1}^n \mathbb{P} \left\{ \left| B_i - \mathbb{E}[B_i] \right| > v \sqrt{\frac{\log n}{n \phi_x(h)}} \right\} < \infty \quad \text{for } i = 1, \dots, 4.$$

We will follow same demonstration as in Barrientos-Marin et al. (2010) to demonstrate A in Eq.(5.2). It suffices now, to show that $\mathbb{E}[B_i] = O(1)$, we do examine

$$\begin{aligned}\mathbb{E}[B_1] &= \phi_x^{-1}(h) \mathbb{E} \left[G_n^{-1}(Y_i) K_1(x) \beta_1(x) Y_1^{-l} \right] \\ &= \phi_x^{-1}(h) \int G_n^{-1}(Y_i) K_1(x) \beta_1(x) Y_1^{-l} f^*(y|u) dy d\mathbb{P}_x^X(x) \\ &= \phi_x^{-1}(h) \int K_1(x) \beta_1(x) Y_1^{-l} f(y|u) dy d\mathbb{P}_x^X(x) \\ &= \phi_x^{-1}(h) \mathbb{E} \left[K_1(x) \beta_1(x) Y_1^{-l} \right]\end{aligned}$$

for $l = 0, 1, 2$ and $k = 1, 2$, by conditioning on X_1 , we have:

$$\mathbb{E} \left[K_i(x) \beta_i(x) Y_i^{-l} \right] = O(h^l \phi_x(h)).$$

Therefore,

$$\mathbb{E}[B_i] = O(1), \quad \text{for } i = 1, 2, 3, 4.$$

It is not difficult to prove Eq.(5.3), for that we propose that for $i, j = 1, 2, 3, 4$ all $Cov(B_i, B_j)$ are of order

$$\frac{1}{nh^l} \phi_x^2(h) \mathbb{E} \left[K_i(x) \beta_i^l(x) Y_i^{-k} \right], \quad \text{for } l = 0, 1, 2, 3, 4 \quad \text{and} \quad k = 1, 2, 3, 4.$$

By the same reasoning we use (H3) to get:

$$\begin{aligned}\mathbb{E}[B_i B_j] &= \frac{1}{nh^l} \phi_x^2(h) \mathbb{E} \left[G_n^{-1}(Y_i) K_i(x) \beta_i^l(x) Y_i^{-k} \right] \\ &= \frac{1}{nh^l} \phi_x^2(h) \mathbb{E} \left[K_i(x) \beta_i^l(x) Y_i^{-k} \right] \\ &= \frac{1}{nh^l} \phi_x^2(h) \mathbb{E} \left[K_i(x) Y_i^{-k} \right] \quad \text{for } l = 0 \quad (\text{Lemme A1}) \\ &= \frac{1}{nh^l} \phi_x^2(h) O(\phi_x(h)) \\ &= O((n\phi_x(h))^{-1})\end{aligned}$$

which is negligible with respect to $\sqrt{\frac{\log n}{n\phi_x(h)}}$,

therefore:

$$Cov(B_i, B_j) = o\left(\sqrt{\frac{\log n}{n\phi_x(h)}}\right).$$

With the same way we work on $\hat{g}$ such that for $i, j = 1, \dots, 6$ we have:

$$\begin{aligned} \hat{g}(x) = & \underbrace{\frac{n^2 h^2 \phi_x^2(h)}{n(n-1)\mathbb{E}[w_{12}]}}_A \left[\underbrace{\left(\frac{1}{n} \sum_{i=1}^n \frac{G_n^{-1}(Y_i) K_i(x) \beta_i^2(x) Y_i^{-2}}{h^2 \phi_x(h)} \right)}_{B_2} \underbrace{\left(\frac{1}{n} \sum_{j=1}^n \frac{K_j(x) Y_j^{-1}}{\phi_x(h)} \right)}_{B_5} \right. \\ & \left. - \underbrace{\left(\frac{1}{n} \sum_{i=1}^n \frac{G_n^{-1}(Y_i) K_i(x) \beta_i(x) Y_i^{-2}}{h \phi_x(h)} \right)}_{B_4} \underbrace{\left(\frac{1}{n} \sum_{j=1}^n \frac{K_j(x) \beta_j(x) Y_j^{-1}}{h \phi_x(h)} \right)}_{B_6} \right] \end{aligned}$$

□

Proof. (Lemma 3.3)

Notice that $(X_i, Y_i)_{i=1, \dots, n}$ are an i.i.d observations, so:

$$\begin{aligned} \mathbb{E}[\hat{f}(x)] - g_2^2(x) &= \frac{1}{\mathbb{E}[w_{12}]} \mathbb{E} \left[(K_1(x) Y_1^{-2}) (G_n^{-1}(Y_1) K_1(x) \beta_1^2(x) Y_1^{-2}) \right. \\ &\quad \left. - (K_1(x) \beta_1(x) Y_1^{-2}) (G_n^{-1}(Y_1) K_1(x) \beta_1(x) Y_1^{-2}) - g_2^2(x) \mathbb{E}[w_{12}] \right] \\ &= \frac{1}{\mathbb{E}[w_{12}]} \left[\mathbb{E} [K_1(x) Y_1^{-2}] \int G_n^{-1}(Y_1) K_1(x) \beta_1^2(x) Y_1^{-2} f^*(x|u) dy d\mathbb{P}_x^X(x) \right. \\ &\quad \left. - \mathbb{E} [K_1(x) \beta_1(x) Y_1^{-2}] \int G_n^{-1}(Y_1) K_1(x) \beta_1(x) Y_1^{-2} f^*(x|u) dy d\mathbb{P}_x^X(x) - g_2^2(x) \mathbb{E}[w_{12}] \right] \\ &= \frac{1}{\mathbb{E}[w_{12}]} \left[\mathbb{E} [K_1(x) Y_1^{-2}] \int K_1(x) \beta_1^2(x) Y_1^{-2} f(x|u) dy d\mathbb{P}_x^X(x) \right. \\ &\quad \left. - \mathbb{E} [K_1(x) \beta_1(x) Y_1^{-2}] \int K_1(x) \beta_1(x) Y_1^{-2} f(x|u) dy d\mathbb{P}_x^X(x) - g_2^2(x) \mathbb{E}[w_{12}] \right] \\ &= \frac{1}{\mathbb{E}[w_{12}]} \left[\mathbb{E} [K_1(x) Y_1^{-2}] \mathbb{E} [K_1(x) \beta_1^2(x) Y_1^{-2}] - \mathbb{E}^2 [K_1(x) \beta_1(x) Y_1^{-2}] - g_2^2(x) \mathbb{E}[w_{12}] \right] \\ &\leq \frac{1}{\mathbb{E}[w_{12}]} \left[\mathbb{E} [K_1(x) Y_1^{-2}] \left| \mathbb{E} [\beta_1^2(x) K_1(x) Y_1^{-2}] - g_2(x) \mathbb{E} [\beta_1^2(x) K_1(x)] \right| \right. \\ &\quad \left. + g_2(x) \mathbb{E} [\beta_1^2(x) K_1(x)] \left| \mathbb{E} [K_1(x) Y_1^{-2}(x)] - g_2(x) \mathbb{E} [K_1(x)] \right| \right. \\ &\quad \left. + \mathbb{E} [K_1(x) \beta_1(x) Y_1^{-2}] \left| \mathbb{E} [K_1(x) \beta_1(x) Y_1^{-2}] - g_2(x) \mathbb{E} [\beta_1(x) K_1(x)] \right| \right. \\ &\quad \left. + g_2(x) \mathbb{E} [\beta_1(x) K_1(x)] \left| \mathbb{E} [\beta_1(x) K_1(x) Y_1^{-2}(x)] - g_2(x) \mathbb{E} [\beta_1(x) K_1(x)] \right| \right]. \end{aligned}$$

From assumptions (H2), (H3) and (H4) we get:

$$\left| \mathbb{E} [\beta_1^2(x) K_1(x) Y_1^{-2}] - g_2(x) \mathbb{E} [\beta_1^2(x) K_1(x)] \right| \leq C \mathbb{E} [\beta_1^2(x) K_1(x)] h^{k_2}$$

and

$$\left| \mathbb{E} [\beta_1(x) K_1(x) Y_1^{-2}(x)] - g_2(x) \mathbb{E} [\beta_1(x) K_1(x)] \right| \leq C \mathbb{E} [\beta_1(x) K_1(x)] h^{k_2}.$$

Then, for $k, i = 0, 1, 2$,

$$\mathbb{E} [\beta_1^l(x) K_1(x) Y_1^{-k}(x)] = O(h^l \phi_x(h))$$

we get

$$\left| \mathbb{E}[\hat{f}(x)] - g_2^2(x) \right| \leq \frac{Ch^2 \phi_x^2(h)}{\mathbb{E}[w_{12}]} h^{k_2} \leq C' h^{k_2} = O(h^{k_2}).$$

By using same arguments as previously, we show that

$$\left| \mathbb{E}[\hat{f}(x)] - g_2(x) g_1(x) \right| \leq Ch^{k_1} = O(h^{k_1})$$

□

Proof. (Corollary 3.4) It easy to see that:

$$|\hat{f}(x)| \leq \frac{g_2^2(x)}{2} \implies |g_2^2(x) - \hat{f}(x)| \geq \frac{g_2^2(x)}{2}.$$

Thus

$$\mathbb{P} \left(|\hat{f}(x)| \leq \frac{g_2^2(x)}{2} \right) \leq \mathbb{P} \left(|g_2^2(x) - \hat{f}(x)| \geq \frac{g_2^2(x)}{2} \right)$$

by consequence:

$$\sum_{n=1}^{\infty} \mathbb{P} \left(|\hat{f}(x)| \leq \frac{g_2^2(x)}{2} \right) < \infty.$$

□

Proof. (Lemma 4.2)

For that, we use the same decomposition as for the proof of Lemma 3.2, and then we can show the uniform almost complete convergence of Eqs. (5.1), (5.2) and (5.3).

Note that, by the hypothesis (U1) we get Eq. (5.2) and (5.3) then the progress is obtained in Lemma (3.2). For $k = 1, 2, \dots, 6$, to show Eq.(5.1) we follow the decomposition below:

$$\begin{aligned} \sup_{x \in S_{\mathcal{F}}} |\Gamma_k(x) - \mathbb{E}[\Gamma_k(x)]| &= \underbrace{\sup_{x \in S_{\mathcal{F}}} |\Gamma_k(x) - \Gamma_k(x_{j(x)})|}_{W_1} + \underbrace{\sup_{x \in S_{\mathcal{F}}} |\Gamma_k(x_{j(x)}) - \mathbb{E}[\Gamma_k(x_{j(x)})]|}_{W_2} \\ &\quad + \underbrace{\sup_{x \in S_{\mathcal{F}}} |\mathbb{E}[\Gamma_k(x_{j(x)})] - \mathbb{E}[\Gamma_k(x)]|}_{W_3}. \end{aligned}$$

So, for terms W_1 and W_3 , since we have K is supported within $[-1, 1]$, we have:

$$\begin{aligned}
 W_1 &\leq \frac{1}{nh^l \phi_x(h)} \sup_{x \in S_{\mathcal{F}}} \sum_{i=1}^n \left| G_n^{-1}(Y_i) K_i(x) Y_i^{-k} \beta_i^l(x) \mathbb{I}_{B(x,h)}(X_i) - G_n^{-1}(Y_i) K_i(x_{j(x)}) Y_i^{-k} \beta_i^l(x_{j(x)}) \mathbb{I}_{B(x_{j(x)},h)}(X_i) \right| \\
 &\leq \frac{C}{nh^l \phi_x(h)} \sup_{x \in S_{\mathcal{F}}} \sum_{i=1}^n \underbrace{G_n^{-1}(Y_i) K_i(x) Y_i^{-k} \mathbb{I}_{B(x,h)}(X_i) \left| \beta_i^l(x) - \beta_i^l(x_{j(x)}) \mathbb{I}_{B(x_{j(x)},h)}(X_i) \right|}_{(1)} \\
 &\quad + \frac{1}{nh^l \phi_x(h)} \sup_{x \in S_{\mathcal{F}}} \sum_{i=1}^n \underbrace{G_n^{-1}(Y_i) Y_i^{-k} \beta_i^l(x_{j(x)}) \mathbb{I}_{B(x_{j(x)},h)}(X_i) \left| K_i(x) \mathbb{I}_{B(x,h)}(X_i) - K_i(x_{j(x)}) \right|}_{(2)}.
 \end{aligned}$$

The Lipschitz condition on β and K allows us to get:

$$\begin{aligned}
 (1) &= \mathbb{I}_{B(x,h)}(X_i) \left| \beta_i(x) - \beta_i(x_{j(x)}) \mathbb{I}_{B(x_{j(x)},h)}(X_i) \right| \leq \varepsilon \mathbb{I}_{B(x,h) \cap B(x_{j(x)},h)}(X_i) + h \mathbb{I}_{B(x,h) \cap \overline{B(x_{j(x)},h)}}(X_i) \\
 (2) &= \mathbb{I}_{B(x_{j(x)},h)}(X_i) \left| K_i(x) \mathbb{I}_{B(x,h)}(X_i) - K_i(x_{j(x)}) \right| \leq C \varepsilon \mathbb{I}_{B(x,h) \cap B(x_{j(x)},h)}(X_i) + C \mathbb{I}_{B(x_{j(x)},h) \cap \overline{B(x,h)}}(X_i).
 \end{aligned}$$

Thus,

$$W_1 \leq C \sup_{x \in S_{\mathcal{F}}} (W_{11} + W_{12} + W_{13} + W_{14}),$$

with:

$$\begin{aligned}
 W_{11} &= \frac{C}{n\phi(h)} \sum_{i=1}^n G_n^{-1}(Y_i) Y_i^{-k} \mathbb{I}_{B(x,h) \cap \overline{B(x_{j(x)},h)}}(X_i) \\
 W_{12} &= \frac{C\varepsilon}{n\phi(h)} \sum_{i=1}^n G_n^{-1}(Y_i) Y_i^{-k} \mathbb{I}_{B(x,h) \cap B(x_{j(x)},h)}(X_i) \\
 W_{13} &= \frac{C\varepsilon}{n\phi(h)} \sum_{i=1}^n G_n^{-1}(Y_i) Y_i^{-k} \mathbb{I}_{B(x,h) \cap B(x_{j(x)},h)}(X_i) \\
 W_{14} &= \frac{C}{n\phi(h)} \sum_{i=1}^n G_n^{-1}(Y_i) Y_i^{-k} \mathbb{I}_{B(x_{j(x)},h) \cap \overline{B(x,h)}}(X_i).
 \end{aligned}$$

To demonstrate the almost complete limits of W_{11} , W_{12} , W_{13} and W_{14} we apply Bernstein's inequality on:

$$\Pi_i = \begin{cases} \frac{G_n^{-1}(Y_i) Y_i^{-k}}{\phi(h)} \sup_{x \in S_{\mathcal{F}}} \left[\mathbb{I}_{B(x,h) \cap \overline{B(x_{j(x)},h)}}(X_i) \right] & \text{for } W_{11} \\ \frac{\varepsilon G_n^{-1}(Y_i) Y_i^{-k}}{h\phi(h)} \sup_{x \in S_{\mathcal{F}}} \left[\mathbb{I}_{B(x,h) \cap B(x_{j(x)},h)}(X_i) \right] & \text{for } W_{12}, W_{13} \\ \frac{G_n^{-1}(Y_i) Y_i^{-k}}{\phi(h)} \sup_{x \in S_{\mathcal{F}}} \left[\mathbb{I}_{B(x_{j(x)},h) \cap \overline{B(x,h)}}(X_i) \right] & \text{for } W_{14}. \end{cases}$$

Under (U1), (U6) and (U7) we get for W_{11} and W_{14} :

$$\mathbb{E}[\Pi_1] = O(\varepsilon(\phi(h))^{-1}) \quad \text{and} \quad \mathbb{E}|\Pi_1|^m = O(\varepsilon(\phi^m(h))^{-1})$$

then for W_{12} and W_{13} :

$$\mathbb{E}[\Pi_1] = O(\varepsilon h^{-1}) \quad \text{and} \quad \mathbb{E}|\Pi_1|^m = O(\varepsilon^m (h^m \phi^{m-1}(h))^{-1})$$

Which implies that:

$$\begin{aligned} \mathbb{E}[W_{11}] &= O(\varepsilon \phi(h)^{-1}) + O_{\text{a.co}} \left(\sqrt{\frac{\varepsilon \log n}{n \phi(h)^2}} \right) \\ \mathbb{E}[W_{12}] &= O_{\text{a.co}} \left(\sqrt{\frac{\log d_n}{n \phi(h)}} \right) \\ \mathbb{E}[W_{13}] &= O_{\text{a.co}} \left(\sqrt{\frac{\log d_n}{n \phi(h)}} \right) \\ \mathbb{E}[W_{14}] &= O(\varepsilon \phi(h)^{-1}) + O_{\text{a.co}} \left(\sqrt{\frac{\varepsilon \log n}{n \phi(h)^2}} \right). \end{aligned}$$

Hence, under (U6) we get:

$$W_1 = O_{\text{a.co}} \left(\sqrt{\frac{\log d_n}{n \phi(h)}} \right).$$

Moreover, since

$$W_3 \leq \mathbb{E} \left[\sup_{x \in S_{\mathcal{F}}} |\Gamma_k(x) - \Gamma_k(x_{j(x)})| \right]$$

we get:

$$W_3 = O_{\text{a.co}} \left(\sqrt{\frac{\log d_n}{n \phi(h)}} \right)$$

Next, for $W_2^k, \forall \nu > 0$ so:

$$\begin{aligned} \mathbb{P} \left(W_2 > \nu \sqrt{\frac{\log d_n}{n \phi(h)}} \right) &= \mathbb{P} \left(\max_{j \in \{1, \dots, d_n\}} \left| \Gamma_k(x_{j(x)}) - \mathbb{E}[\Gamma_k(x_{j(x)})] \right| > \nu \sqrt{\frac{\log d_n}{n \phi(h)}} \right) \\ &\leq d_n \max_{j \in \{1, \dots, d_n\}} \mathbb{P} \left(\left| \Gamma_k(x_{j(x)}) - \mathbb{E}[\Gamma_k(x_{j(x)})] \right| > \nu \sqrt{\frac{\log d_n}{n \phi(h)}} \right) \end{aligned}$$

we put

$$\Pi_{ki} = \frac{1}{nh^l \phi(h)} \left(G_n^{-1}(Y_i) K_i(x_k) Y_i^{-k} \beta_i^l(x_k) - \mathbb{E}[G_n^{-1}(Y_i) K_i(x_k) Y_i^{-k} \beta_i^l(x_k)] \right).$$

We use same ideas in proof of Lemma 3.2 to show that:

$$\mathbb{E}|\Pi_{ki}|^m = O(\phi(h)^{-m+1}).$$

Once again we apply a Bernstein-type inequality to obtain:

$$\begin{aligned} \mathbb{P} \left(\left| \Gamma_k(x_{j(x)}) - \mathbb{E}[\Gamma_k(x_{j(x)})] \right| > \nu \sqrt{\frac{\log d_n}{n \phi(h)}} \right) &= \mathbb{P} \left(\left| \sum_{i=1}^n \Pi_{ki} \right| > \nu \sqrt{\frac{\log d_n}{n \phi(h)}} \right) \\ &\leq 2 \exp(-C\nu^2 \log d_n) \end{aligned}$$

We put $\beta = C\nu^2$ we get:

$$\max_{k \in \{1, \dots, d_n\}} \mathbb{P} \left(\left| \Gamma_k(x_{j(x)}) - \mathbb{E}[\Gamma_k(x_{j(x)})] \right| > \nu \sqrt{\frac{\log d_n}{n\phi(h)}} \right) \leq C' d_n^{1-\beta}.$$

Then since $\sum_{n=1}^{\infty} d_n^{1-\beta} < \infty$, we get:

$$W_2 = O_{\text{a.co}} \left(\sqrt{\frac{\log d_n}{n\phi(h)}} \right).$$

□

To prove Lemma 4.3, it suffices to combine between the proofs of Lemma 3.3 and conditions of Lipschitz, uniformly on x in $S_{\mathcal{F}}$.

Proof. (Corollary 4.4)

It's obvious that:

$$\inf_{x \in S_{\mathcal{F}}} |\hat{f}(x)| \leq \inf_{x \in S_{\mathcal{F}}} \frac{g_2^2(x)}{2}$$

which implies that:

$$\sup_{x \in S_{\mathcal{F}}} |\hat{f}(x) - g_2^2(x)| \geq \inf_{x \in S_{\mathcal{F}}} \frac{g_2^2(x)}{2}$$

so,

$$\begin{aligned} \sum_{i=1}^n \mathbb{P} \left(\inf_{x \in S_{\mathcal{F}}} |\hat{f}(x)| \leq \inf_{x \in S_{\mathcal{F}}} \frac{g_2^2(x)}{2} \right) &\leq \sum_{i=1}^n \mathbb{P} \left(\sup_{x \in S_{\mathcal{F}}} |\hat{f}(x) - g_2^2(x)| \geq \inf_{x \in S_{\mathcal{F}}} \frac{g_2^2(x)}{2} \right) \\ &< \infty. \end{aligned}$$

□