

# Chapter 2

## Least Squares Estimation via Plug-In

### 2.1 Regression Estimation

Under i.i.d. random vectors  $(X, Y)$ ,  $(X_1, Y_1)$ ,  $(X_2, Y_2), \dots$  in the regression analysis one is interested in the value of the so called response variable  $Y$  (in  $\mathbb{R}$ ) depending on the value of the observation vector  $X$  (in  $\mathbb{R}^d$ , with distribution  $\mu$ ). To find that, one searches a (measurable) function  $f : \mathbb{R}^d \rightarrow \mathbb{R}$ , such that  $f(X)$  is a "good approximation of  $Y$ ", that is,  $f(X)$  should be "close" to  $Y$ , achieved making the random quantity  $|f(X) - Y|$  "small". For this, assuming square integrability of  $Y$ , we introduce the so-called  $L_2$ -risk or mean squared error of  $f$ :

$$E \{ |f(X) - Y|^2 \}. \quad (2.1)$$

It is well known that the function that minimizes in a certain sense (2.1) is the regression function,  $m(x) := E\{Y|X = x\}$ , unknown if the distribution of  $(X, Y)$  is unknown. In this case, starting from a dataset  $D_n$ , a nonparametric estimator  $m_n$  of the regression function is to construct. There exist different paradigms how to make it, the aim here is to deal with least squares approaches. There, the basic idea is to estimate the unknown mean squared error in (2.1) by approximating the expectation value there appearing via the empirical mean:

$$\frac{1}{n} \sum_{i=1}^n |f(X_i) - Y_i|^2, \quad (2.2)$$

and to choose a function, over a set  $\mathcal{F}_n$  of functions given by the statistician,  $f : \mathbb{R}^d \rightarrow \mathbb{R}$ , that minimizes (2.2).

Examples of possible choices of the set  $\mathcal{F}_n$  are sets of piecewise polynomials with respect to a partition  $\mathcal{P}_n$ . Clearly it doesn't make sense to minimize

(2.2) over all measurable functions  $f$ , because this may lead to a function which interpolates the data and hence is not a reasonable estimate. The least squares estimator of the regression function is defined as:

$$m_n(\cdot) = \arg \min_{f \in \mathcal{F}_n} \frac{1}{n} \sum_{i=1}^n |f(X_i) - Y_i|^2,$$

where the optimal function is not required to be unique. For consistency and rate of convergence of such estimators see [16] and the references cited there.

Sometimes it is possible to observe data from the underlying distribution only with measurement errors. It can for example happen, that the predictor vector  $X$  can be observed only with errors, i.e., instead of  $X_i$  one observes  $W_i = X_i + U_i$  for some random variable  $U_i$  which satisfy  $\mathbf{E}\{U_i|X_i\} = 0$  and the aim is to estimate the regression function from  $\{(W_1, Y_1), \dots, (W_n, Y_n)\}$ . Instead, as in [16], we assume that we can observe the dependent variable  $Y$  only with supplementary, maybe correlated, measurement errors. Since we do not assume that the means of these measurement errors are zero, these kinds of errors are not already included in standard models. Our dataset is

$$\overline{D}_n = \{(X_1, \overline{Y}_{1,n}), \dots, (X_n, \overline{Y}_{n,n})\},$$

where the only assumption on the random variables  $\overline{Y}_{1,n}, \dots, \overline{Y}_{n,n}$  is that the differences between  $Y_i$  and  $\overline{Y}_{i,n}$  are in a certain sense "small". We will therefore assume that the average squared measurement error

$$\frac{1}{n} \sum_{i=1}^n |Y_i - \overline{Y}_{i,n}|^2$$

is small. Set briefly  $\overline{Y}_i := \overline{Y}_{i,n}$ . With the difficulty of additional measurement errors in the dependent variable in our notation the estimator becomes:

$$\overline{m}_n(\cdot)^{(LS)} = \arg \min_{f \in \mathcal{F}_n} \frac{1}{n} \sum_{i=1}^n |f(X_i) - \overline{Y}_i|^2, \quad (2.3)$$

and in this chapter we set  $\overline{m}_n(x) := \overline{m}_n(x)^{(LS)}$ .

Inspired by [16], Corollary 1, we want here to treat consistency in a general case where we require the unknown regression function only to be bounded in absolute value from above by a constant and without continuity assumptions.

**Theorem 2.1.** *Assume that  $Y - m(X)$  is sub-Gaussian in the sense that*

$$K^2 \mathbf{E} \left\{ e^{(Y-m(X))^2/K^2} - 1 | X \right\} \leq \sigma_0^2 \quad \text{almost surely,}$$

for some  $K, \sigma_0 > 0$ . Let  $L \geq 1$  and assume that the regression function is bounded in absolute value by  $L$  ( $\Rightarrow m \in L_2(\mu)$ ). We define  $\mathcal{F}_n$  as a subset of a linear space, consisting of real-valued functions on  $\mathbb{R}^d$ , with dimension  $D_n \in \mathbb{N}$  and with the property  $|f| \leq L$  for  $f \in \mathcal{F}_n$ , where  $\mathcal{F}_n \uparrow, D_n \rightarrow \infty$  for  $n \rightarrow \infty$ , but  $\frac{D_n}{n} \rightarrow 0$ . Furthermore  $\cup_n \mathcal{F}_n$  is required to be dense in the subspace of  $L_2(\mu)$ , consisting of the functions in  $L_2(\mu)$  absolutely bounded by  $L$ . In addition it shall hold

$$\frac{1}{n} \sum_{i=1}^n |Y_i - \bar{Y}_i|^2 \xrightarrow{P} 0. \quad (2.4)$$

Then, we have

$$\int |\bar{m}_n(x) - m(x)|^2 \mu(dx) \xrightarrow{P} 0.$$

(Consistency of the least squares estimator of the regression function with additional measurements error in the response variable)

For the proof of Theorem 2.1 we use the following lemma:

**Lemma 2.1.**  *$\{U_n\}$  and  $\{V_n\}$  are nonnegative real random sequences. Assume  $V_n \xrightarrow{P} 0$  and  $\mathbf{P}\{U_n > V_n\} \rightarrow 0$  ( $n \rightarrow \infty$ ). Then  $U_n \xrightarrow{P} 0$ .*

*Proof.* For each  $\varepsilon > 0$  and each  $\delta > 0$  there exists an  $n_0$  such that, for every  $n \geq n_0$ :

$$\mathbf{P} \underbrace{\{V_n > \varepsilon\}}_{\Omega'_n} \leq \frac{\delta}{2}$$

and

$$\mathbf{P} \underbrace{\{U_n > V_n\}}_{\Omega''_n} \leq \frac{\delta}{2},$$

thus

$$\begin{aligned} \mathbf{P}(\{U_n \leq \varepsilon\}^c) &\leq \mathbf{P}(\{\{V_n \leq \varepsilon\} \cap \{U_n \leq V_n\}\}^c) \\ &= \mathbf{P}(\{\Omega_n'^c \cap \Omega_n''^c\}^c) \stackrel{\text{De Morgan}}{=} \mathbf{P}(\{\Omega_n' \cup \Omega_n''\}) \\ &\leq \mathbf{P}\{\Omega_n'\} + \mathbf{P}\{\Omega_n''\} \leq \frac{\delta}{2} + \frac{\delta}{2} = \delta. \end{aligned}$$

□

We are now ready for the following:

*Proof of Theorem 2.1.* According to [16], Corollary 1, there is a positive constant  $c$  depending only on  $L$ ,  $\sigma_0$ ,  $K$  with the following property:

$$\begin{aligned} & P \left\{ \int |\overline{m}_n(x) - m(x)|^2 \mu(dx) \right. \\ & \left. > c \left( \frac{1}{n} \sum_{i=1}^n |Y_i - \overline{Y}_i|^2 + \frac{D_n}{n} + \inf_{f \in \mathcal{F}_n} \int |f(x) - m(x)|^2 \mu(dx) \right) \right\} \rightarrow 0. \end{aligned}$$

By Lemma 2.1, it is enough to show

$$\left( \frac{1}{n} \sum_{i=1}^n |Y_i - \overline{Y}_i|^2 + \frac{D_n}{n} + \inf_{f \in \mathcal{F}_n} \int |f(x) - m(x)|^2 \mu(dx) \right) \xrightarrow{P} 0 \quad (2.5)$$

But (2.5) holds, because of (2.4),  $\frac{D_n}{n} \rightarrow 0$  and  $\inf_{f \in \mathcal{F}_n} \int |f(x) - m(x)|^2 \mu(dx) \rightarrow 0$  (due to  $\mathcal{F}_n \uparrow$  and the density of  $\cup_n \mathcal{F}_n$  in the mentioned subspace of  $L_2(\mu)$ ).

## 2.2 Local Variance Estimation with Additional Measurement Errors

The quality of the regression function  $m$  in view of small mean squared error is globally given by  $\mathbf{E}\{(Y - m(X))^2\}$  and locally by

$$\sigma^2(x) := \mathbf{E}\{(Y - m(X))^2 | X = x\} = \mathbf{E}\{Y^2 | X = x\} - m^2(x). \quad (2.6)$$

$\sigma^2(x)$  is the so called local variance. We define a new variable

$$Z := Y^2 - m^2(X) \quad (2.7)$$

and consequently its observations (in the case of known  $m$ ):

$$Z_i := Y_i^2 - m^2(X_i);$$

finally the observations with additional errors:

$$\overline{Z}_i := \overline{Y}_i^2 - \overline{m}_n^2(X_i),$$

with  $\overline{m}_n = \overline{m}_n^{(LS)}$  according to (2.4). Combining (2.6) and (2.7) allows us to say that the local variance is a regression on  $(X, Z)$ .

We can therefore define the least squares estimator of the local variance, analogously to the estimator (2.3) as

$$\bar{\sigma}_n^2(\cdot)^{(LS)} = \arg \min_{g \in \mathcal{G}_n} \frac{1}{n} \sum_{i=1}^n |g(X_i) - \bar{Z}_i|^2, \quad (2.8)$$

where  $g : \mathbb{R}^d \rightarrow \mathbb{R} \in \mathcal{G}_n$ , with suitable function space  $\mathcal{G}_n$ . Briefly, define  $\bar{\sigma}_n^2(x) := \bar{\sigma}_n^2(x)^{(LS)}$ .

For consistency and convergence rate under Lipschitz conditions see [16]. We want here to treat consistency in a general case where no smoothness conditions on the  $m$  and  $\sigma^2$  are required.

**Theorem 2.2.** *Assume that  $Y^2 - m^2(X)$  is sub-Gaussian in the sense that*

$$K^2 \mathbf{E} \left\{ e^{(Y^2 - m^2(X))^2 / K^2} - 1 | X \right\} \leq \sigma_0^2 \quad \text{almost surely}$$

for some  $K, \sigma_0 > 0$ . Let

$$\frac{1}{n} \sum_{i=1}^n |Y_i^p - \bar{Y}_i^p|^2 \xrightarrow{P} 0, \quad p = 1, 2. \quad (2.9)$$

It is assumed that  $L^* > 0$  and  $L > 0$  exist such that  $\sigma^2 \leq L^*$  and  $|m| \leq L$ . Let  $\mathcal{G}_n$  be defined as a subset of a linear space, consisting of nonnegative real-valued functions on  $\mathbb{R}^d$  bounded by  $L^*$ , with dimension  $D_n \in \mathbb{N}$ , with the properties  $\mathcal{G}_n \uparrow, D_n \rightarrow \infty$  for  $n \rightarrow \infty$  but  $\frac{D_n}{n} \rightarrow 0$ . Furthermore  $\cup_n \mathcal{G}_n$  is required to be dense in the subspace of  $L_2(\mu)$  consisting of the nonnegative functions in  $L_2(\mu)$  bounded by  $L^*$ . Let also  $\mathcal{F}_n$  be defined as a subset of a linear space of real-valued functions on  $\mathbb{R}^d$  absolutely bounded by  $L$ , with dimension  $D'_n \in \mathbb{N}$ , with the properties  $\mathcal{F}_n \uparrow, D'_n \rightarrow \infty$  for  $n \rightarrow \infty$  but  $\frac{D'_n}{n} \rightarrow 0$ . Furthermore  $\cup_n \mathcal{F}_n$  is required to be a dense subset of  $C_{0,L}^0(\mathbb{R}^d)$  (with respect to the max norm), where  $C_{0,L}^0(\mathbb{R}^d)$  denotes the space of continuous real valued functions on  $\mathbb{R}^d$  absolutely bounded by  $L$ , with compact support. Then

$$\int |\bar{\sigma}_n^2(x) - \sigma^2(x)|^2 \mu(dx) \xrightarrow{P} 0.$$

(Consistency of the least squares estimator of the local variance with additional measurements error in the response variable)

*Remark 2.1.* In case of boundedness of  $Y_i$  and  $\bar{Y}_i$  (2.9) for  $p = 1$  implies (2.9) for  $p = 2$  (because of  $Y_i^2 - \bar{Y}_i^2 = (Y_i - \bar{Y}_i)(Y_i + \bar{Y}_i)$ ).

*Proof of Theorem 2.2.* As in the proof of Theorem 2.1 we obtain that there exists a generic positive constant  $c$  depending only from  $L, \sigma_0, K$  with the

following property:

$$\begin{aligned} & \mathbf{P} \left\{ \int |\bar{\sigma}_n^2(x) - \sigma^2(x)|^2 \mu(dx) > c \cdot \right. \\ & \left. \left( \frac{1}{n} \sum_{i=1}^n |Z_i - \bar{Z}|^2 + \frac{D_n}{n} + \inf_{g \in \mathcal{G}_n} \int |g(x) - \sigma^2(x)|^2 \mu(dx) \right) \right\} \rightarrow 0. \end{aligned} \quad (2.10)$$

We notice

$$\begin{aligned} & \frac{1}{n} \sum_{i=1}^n |Z_i - \bar{Z}|^2 \\ & \leq \frac{2}{n} \sum_{i=1}^n |\bar{m}_n^2(X_i) - m^2(X_i)|^2 + \frac{2}{n} \sum_{i=1}^n |Y_i^2 - \bar{Y}_i^2|^2 \\ & \leq \frac{8}{n} L^2 \sum_{i=1}^n |\bar{m}_n(X_i) - m(X_i)|^2 + \underbrace{\frac{2}{n} \sum_{i=1}^n |Y_i^2 - \bar{Y}_i^2|^2}_{\xrightarrow{P} 0 \text{ assumption (2.9)}}. \end{aligned} \quad (2.11)$$

It remains to prove

$$\frac{1}{n} \sum_{i=1}^n |\bar{m}_n(X_i) - m(X_i)|^2 \xrightarrow{P} 0.$$

Via conditioning with respect to  $(X_1, \dots, X_n)$ , by [16], Lemma 3, we obtain

$$\begin{aligned} & \mathbf{P} \left\{ \frac{1}{n} \sum_{i=1}^n |\bar{m}_n(X_i) - m(X_i)|^2 \right. \\ & > c \left( \frac{1}{n} \sum_{i=1}^n |Y_i - \bar{Y}_i|^2 + \frac{D'_n}{n} \right. \\ & \quad \left. \left. + \min_{f \in \mathcal{F}_n} \frac{1}{n} \sum_{i=1}^n |f(X_i) - m(X_i)|^2 \right) \middle| X_1 = x_1, \dots, X_n = x_n \right\} \rightarrow 0 \end{aligned} \quad (2.12)$$

where  $\frac{D'_n}{n} \rightarrow 0$  and  $\frac{1}{n} \sum_{i=1}^n |Y_i - \bar{Y}_i|^2 \xrightarrow{P} 0$  (assumption (2.9)).

Regarding the last term in the round brackets in (2.12), for an arbitrary

we choose  $\varepsilon' > 0$  a continuous function with compact support  $\tilde{m}$  such that  $\mathbf{E}|\tilde{m}(X) - m(X)|^2 \leq \varepsilon'$ .

We observe

$$\begin{aligned}
 & \min_{f \in \mathcal{F}_n} \frac{1}{n} \sum_{i=1}^n |f(X_i) - m(X_i)|^2 \leq \\
 & \leq 2 \underbrace{\min_{f \in \mathcal{F}_n} \frac{1}{n} \sum_{i=1}^n |f(X_i) - \tilde{m}(X_i)|^2}_{\rightarrow 0 \text{ a.s.}} \\
 & + 2 \underbrace{\frac{1}{n} \sum_{i=1}^n |\tilde{m}(X_i) - m(X_i)|^2}_{\text{a.s. } \rightarrow \mathbf{E}\{|\tilde{m}(X) - m(X)|^2\} \leq \varepsilon' \text{ (Strong Law of Large Numbers)}}
 \end{aligned} \tag{2.13}$$

where (2.13) follows from

$$\begin{aligned}
 & \min_{f \in \mathcal{F}_n} \frac{1}{n} \sum_{i=1}^n |f(X_i) - \tilde{m}(X_i)|^2 \\
 & \leq \min_{f \in \mathcal{F}_n} \left[ \sup_x |f(x) - \tilde{m}(x)|^2 \right] \rightarrow 0.
 \end{aligned}$$

By Lemma 2.1 the assertion follows.

## 2.3 Rate of Convergence

In this section we investigate the rate of convergence of the least squares estimator of the local variance function with additional measurement errors in the dependent variable. In the special case that there are no additional measurement errors and that  $d = 1$  Kohler's Corollary 3 [16] investigates the rate of convergence of the estimator.

In this section we choose as suitable function space for the minimization problem in (2.8) the space of B-spline functions, as Mathe did [22].

We recall now briefly the definitions of B-splines and the B-splines space. For a deeper discussion of splines with proofs we refer the reader to the classical reference here [3].

**Definition 2.1.** Let  $K := (K_i)$  be a nondecreasing sequence. The  $i$ -th univariate (normalized) B-spline of order  $M$  for the knot sequence  $K$  is denoted by the rule

$$B_{i,K,M}(x) := (K_{i+M} - K_i)[K_i, \dots, K_{i+M}](K - x)_+^{M-1} \quad \text{for } x \in \mathbb{R}.$$

Notice that B-splines consists of nonnegative functions which sum up to 1, i.e.,  $B_{i,K,M}$  provides a partition of unity. Further explanations can be found again in [3].

**Definition 2.2.** For  $i = (i_1, \dots, i_d) \in \mathbb{Z}^d$ , the multivariate B-splines of order  $m$  are denoted by

$$B_{i,K,M}^d(x_1, \dots, x_d) := B_{i_1,K,M}(x_1) \cdot \dots \cdot B_{i_d,K,M}(x_d)$$

**Definition 2.3.** A spline function of order  $M$  with knot sequence  $K$  is any linear combination of B-splines of order  $M$  for the knot sequence  $K$ . The collection of all such functions is denoted by  $S_{K,M}$ . In symbols,

$$S_{K,M}([0, 1]^d) := \left\{ \sum_i \alpha_i B_{i,K,M}^d : \alpha_i \text{ real, for all } i \right\}.$$

Notice that the functions from  $S_{K,M}$  are multivariate polynomials of degree smaller or equal to  $M$  and for  $M > 0$  they are  $(M - 1)$ -times continuously differentiable.

Because of the bound of the regression function and the local variance function it makes sense to bound also the functions of the spline space. Therefore, we bound the estimate introducing the following two modifications of the spline space  $S_{K,M}$

$$S_{K,M}^{L+1}([0, 1]^d) := \left\{ \sum_i \alpha_i B_{i,K,M}^d : 0 \leq \alpha_j \leq L + 1 \right. \\ \left. (j \in \{1, \dots, K_n + M\}^d) \right\}$$

and

$$S_{K,M}^{4L^2+1}([0, 1]^d) := \left\{ \sum_i \alpha_i B_{i,K,M}^d : 0 \leq \alpha_j \leq 4L^2 + 1 \right. \\ \left. (j \in \{1, \dots, K'_n + M\}^d) \right\}.$$

Because of the properties of the B-splines to be positive and to sum up to one, the functions from the space  $S_{K,M}^{L+1}([0, 1]^d)$  are nonnegative and bounded by  $L + 1$ . Analogously, the functions from  $S_{K,M}^{4L^2+1}([0, 1]^d)$  are bounded by



$4L^2 + 1$ . The following theorem deals with the rate of convergence of the estimator of the local variance.

**Theorem 2.3.** *Let  $L \geq 1$ ,  $C > 0$  and  $p = k + \beta$  for some  $k \in \mathbb{N}_0$  and  $\beta \in (0, 1]$ . Assume that  $X \in [0, 1]^d$  almost surely. Assume also that  $|Y_i| \leq L$ ,  $|\bar{Y}_i| \leq L$  and*

$$\frac{1}{n} \sum_{i=1}^n |Y_i - \bar{Y}_i|^2 = O_P \left( n^{-\frac{2p}{2p+d}} \right). \quad (2.14)$$

Moreover, let  $\Gamma > 0$ ,  $\Lambda > 0$  and assume that  $m$  and  $\sigma^2$  are  $(p, \Gamma)$  and  $(p, \Lambda)$ -smooth, respectively, that is, for every  $\alpha = (\alpha_1, \dots, \alpha_d)$ ,  $\alpha_j \in \mathbb{N}_0$ ,  $\sum_{j=1}^d \alpha_j = k$

$$\left| \frac{\partial^k m}{\partial x_1^{\alpha_1}, \dots, \partial x_d^{\alpha_d}}(x) - \frac{\partial^k m}{\partial x_1^{\alpha_1}, \dots, \partial x_d^{\alpha_d}}(z) \right| \leq \Gamma \|x - z\|^\beta \quad x, z \in \mathbb{R}^d$$

and

$$\left| \frac{\partial^k \sigma^2}{\partial x_1^{\alpha_1}, \dots, \partial x_d^{\alpha_d}}(x) - \frac{\partial^k \sigma^2}{\partial x_1^{\alpha_1}, \dots, \partial x_d^{\alpha_d}}(z) \right| \leq \Lambda \|x - z\|^\beta \quad x, z \in \mathbb{R}^d$$

( $\|\cdot\|$  denoting the Euclidean norm).

Identify  $\mathcal{F}_n$  and  $\mathcal{G}_n$  with  $S_{K'_n, M}^{L+1}([0, 1]^d)$  and  $S_{K_n, M}^{4L^2+1}([0, 1]^d)$ , respectively, with respect to an equidistant partition of  $[0, 1]^d$  into

$$K'_n = \lceil \Gamma^{\frac{2}{2p+d}} n^{\frac{1}{2p+d}} \rceil$$

for  $\mathcal{F}_n$  and

$$K_n = \lceil \Lambda^{\frac{2}{2p+d}} n^{\frac{1}{2p+d}} \rceil,$$

for  $\mathcal{G}_n$ , respectively. Then

$$\int |\bar{\sigma}_n^2(x) - \sigma^2(x)|^2 \mu(dx) = O_P \left( n^{-\frac{2p}{2p+d}} \right).$$

(Rate of convergence of the least squares estimator of the local variance with additional measurements error in the response variable)

*Proof.* We use (2.10). Because of the dimension  $D_n = c \cdot K_n$  of  $\mathcal{G}_n$  it follows

$$\frac{D_n}{n} \leq O\left(n^{-\frac{2p}{2p+d}}\right). \quad (2.15)$$

From the  $(p, \Gamma)$ -smoothness of  $\sigma^2$  and the definition of  $\mathcal{G}_n$  we can conclude (cf. [22], p. 66)

$$\inf_{g \in \mathcal{G}_n} \int |g(x) - \sigma^2(x)|^2 \mu(dx) \leq O\left(n^{-\frac{2p}{2p+d}}\right) \quad (2.16)$$

In view of the assertion it remains to show

$$\frac{1}{n} \sum_{i=1}^n |Z_i - \bar{Z}_i|^2 = O_P\left(\Lambda^{\frac{1}{2p+d}} n^{-\frac{p}{2p+d}}\right).$$

Now we use (2.11). It holds

$$\frac{1}{n} \sum_{i=1}^n |Y_i^2 - \bar{Y}_i^2|^2 = O_P\left(\Lambda^{\frac{1}{2p+d}} n^{-\frac{p}{2p+d}}\right)$$

because of

$$\frac{1}{n} \sum_{i=1}^n |Y_i^2 - \bar{Y}_i^2|^2 \leq 4L^2 \cdot \frac{1}{n} \sum_{i=1}^n |Y_i - \bar{Y}_i|^2$$

by uniform boundedness of the sequence  $\frac{1}{n} \left( \sum_{i=1}^n |Y_i - \bar{Y}_i|^2 \right)^{1/2}$ , and (2.14).

Thus it remains to show

$$\frac{1}{n} \sum_{i=1}^n |\bar{m}_n(X_i) - m(X_i)|^2 = O_P\left(\Lambda^{\frac{2}{2p+d}} n^{-\frac{2p}{2p+d}}\right).$$

We work now conditionally on  $(X_1, \dots, X_n)$  and observe that

$$\begin{aligned} & P \left\{ \frac{1}{n} \sum_{i=1}^n |\bar{m}_n(X_i) - m(X_i)|^2 \right. \\ & > c \left( \frac{1}{n} \sum_{i=1}^n |Y_i - \bar{Y}_i|^2 + \frac{D'_n}{n} \right. \\ & \quad \left. \left. + \min_{f \in \mathcal{F}_n} \frac{1}{n} \sum_{i=1}^n |f(X_i) - m(X_i)|^2 \right) \middle| X_1 = x_1, \dots, X_n = x_n \right\} \rightarrow 0 \end{aligned} \quad (2.17)$$

Local Variance Estimation for Uncensored and  
Censored Observations

Ferrario, P.G.

2013, XVII, 130 p. 3 illus., Softcover

ISBN: 978-3-658-02313-3