Certainty bands for the conditional cumulative distribution function and applications

Myriam Maumy-Bertrand, A Muller-Gueudin

CERTAINTY BANDS FOR THE CONDITIONAL CUMULATIVE DISTRIBUTION FUNCTION AND APPLICATIONS

BY M. MAUMY-BERTRAND[1] AND A. MULLER-GUEUDIN[2]

ABSTRACT. In this paper, we establish uniform asymptotic certainty bands for the conditional cumulative distribution function. To this aim, we give exact rate of strong uniform consistency for the local linear estimator of this function. The corollaries of this result are the asymptotic certainty bands for the quantiles and the regression function. We illustrate our results with simulations and an application on fetopathologic data.

[1]IRMA (UMR 7501), Université de Strasbourg, France.[1], [2]IECL (UMR 7502), Nancy-Université, CNRS, INRIA, members of the BIGS (BIology, Genetics and Statistics) team at INRIA, France[2]

Keywords: Conditional cumulative distribution function, local polynomial estimator, uniform asymptotic certainty bands, regression function, quantiles.

1. Introduction

1.1. **Motivations.** Consider (X, Y), a random vector defined in $\mathbb{R} \times \mathbb{R}$. Throughout, we work with a sample $\{(X_i, Y_i)_{1 \leqslant i \leqslant n}\}$ of independent and identically replica of (X, Y). We will assume that (X, Y) [resp. X] has a density function $f_{X,Y}$ [resp. f_X] with respect to the Lebesgue measure. In this paper, we will mostly focus on the conditional cumulative distribution function (*cond-cdf*) of Y given $X = x$, defined by:

$$\forall t \in \mathbb{R}, \quad F(t|x) = \mathbb{E}\left(\mathbb{1}_{\{Y \leqslant t\}} | X = x\right) = \mathbb{P}\left(Y \leqslant t | X = x\right). \tag{1}$$

Saying that, we are implicitly assuming the existence of a regular version for the conditional distribution of Y given X.

In this article, we study the conditional cumulative distribution function and a nonparametric estimator associated to this function.

The present paper is organized as follows. First, we introduce the local linear estimator of the *cond-cdf*, with the main notations and assumptions needed for our task. Then we establish an uniform law of the logarithm for the local linear estimator of the *cond-cdf* in Section 2. In Section 3, we show that limit laws of the logarithm are useful in the construction of uniform asymptotic certainty bands for the *cond-cdf*, the regression function and the conditional quantile function. Such certainty bands are obtained from simulations in Section 4 and from fetopathologic data in Section 5.

1.2. **Notations and assumptions.** Let $(X_1, Y_1), (X_2, Y_2), \ldots,$ be independent and identically distributed replica of (X, Y) in $\mathbb{R} \times \mathbb{R}$. Let $I = [a, b], J = [a', b'] \supsetneq I$, two fixed compacts of $\mathbb{R}$.

First, we impose the following set of assumptions upon the distribution of (X, Y):

(F.1) $f_{X,Y}$ is continuous on $J \times \mathbb{R}$ and f_X is continuous and strictly positive on J;
(F.2) $Y\mathbb{1}_{\{X \in J\}}$ is bounded on $\mathbb{R}$.

[1]mmaumy@math.unistra.fr
[2]aurelie.gueudin@univ-lorraine.fr

Remark 1.

(1) Under (F.1-2), the *cond-cdf* is well defined.
(2) The assumption (F.2) is very useful for the proof of our results. This boundedness assumption is common in non-parametric estimation. It ensures the existence of several moments of the *cond-cdf*.

K denotes a positive-valued kernel function defined on $\mathbb{R}$, fulfilling the conditions:

(K.1) K is right-continuous function with bounded variation on $\mathbb{R}$;
(K.2) K is compactly supported and $\int_{\mathbb{R}} K(u)\mathrm{d}u = 1$;
(K.3) $\int_{\mathbb{R}} uK(u)\mathrm{d}u = 0$ and $\int_{\mathbb{R}} u^2K(u)\mathrm{d}u \neq 0$.

We note: $||K||_2^2 = \int_{\mathbb{R}} K^2(u)du$.

Further, introduce the following assumptions on the non-random sequence $(h_n)_{n\geqslant 1}$:

(H.0) for all n, $0 < h_n < 1$;
(H.1) $h_n \to 0$, as $n \to +\infty$;
(H.2) $nh_n/\log n \to +\infty$, as $n \to +\infty$;
(H.3) $h_n \searrow 0$ and $nh_n \nearrow +\infty$, as $n \to +\infty$;
(H.4) $\log(h_n^{-1})/\log\log n \to +\infty$, as $n \to +\infty$.

Remark 2.

(1) The assumption (H.0) is necessary to define $\sqrt{\log(h_n^{-1})}^{-1}$ (see later in our Theorem 2.1).
(2) The assumptions (H.0-2) are necessary and sufficient for our uniform convergence in probability (see Theorem 2.1).
(3) In order to have almost surely convergence results, we need the assumptions (H.3-4) (see Blondin [3]).
(4) The assumptions (H.0, H.2-4) are called the Csörgö-Révész-Stute assumptions.

Our aim will be to establish the strong uniform consistency of the local linear estimator of the conditional cumulative distribution function, defined by:

$$\widehat{F}_n^{(1)}(t, h_n|x) = \frac{\widehat{f}_{n,2}(x,h_n)\widehat{r}_{n,0}(x,t,h_n) - \widehat{f}_{n,1}(x,h_n)\widehat{r}_{n,1}(x,t,h_n)}{\widehat{f}_{n,0}(x,h_n)\widehat{f}_{n,2}(x,h_n) - \left(\widehat{f}_{n,1}(x,h_n)\right)^2} \tag{2}$$

where $^{(1)}$ denotes the order 1 of the local polynomial estimator, and

$$\widehat{f}_{n,j}(x,h_n) = \frac{1}{nh_n}\sum_{i=1}^{n}\left(\frac{x-X_i}{h_n}\right)^j K\left(\frac{x-X_i}{h_n}\right), \text{ for } j = 0,1,2, \tag{3}$$

$$\widehat{r}_{n,j}(x,t,h_n) = \frac{1}{nh_n}\sum_{i=1}^{n}\mathbb{1}_{\{Y_i\leqslant t\}}\left(\frac{x-X_i}{h_n}\right)^j K\left(\frac{x-X_i}{h_n}\right), \text{ for } j = 0,1. \tag{4}$$

Remark 3.

(1) The Nadaraya-Watson estimator $\widehat{F}_n^{(0)}(t, h_n|x)$ can be also written with the functions $\widehat{f}_{n,j}$ and $\widehat{r}_{n,j}$ as

$$\widehat{F}_n^{(0)}(t, h_n|x) = \frac{\widehat{r}_{n,0}(x,t,h_n)}{\widehat{f}_{n,0}(x,h_n)}.$$

It is the local polynomial estimator of order 0 of the conditional cumulative distribution function.

(2) The estimator $\widehat{F}_n^{(1)}(t,h_n|x)$ is better than the Nadaraya-Watson estimator when the design is random and has the favorable property to reproduce polynomial of order 1. Precisely, the local linear estimator has a high minimax efficiency among all possible estimators, including nonlinear smoothers (see Fan and Gijbels [8]).

(3) We have state in the beginning of this Section that we restrict ourselves to the local polynomial estimator of order 1. The local polynomial estimator can be generalized to the orders $p \geqslant 2$, but the equations become more complicated. We show briefly the form of the local polynomial estimator of order 2:

$$\widehat{F}_n^{(2)}(t,h_n|x) = \frac{a_1\widehat{r}_{n,0}(x,t,h_n) + a_2\widehat{r}_{n,1}(x,t,h_n) + a_3\widehat{r}_{n,2}(x,t,h_n)}{a_1\widehat{f}_{n,0}(x,h_n) + a_2\widehat{f}_{n,1}(x,h_n) + a_3\widehat{f}_{n,2}(x,h_n)}$$

$$\text{where } \begin{cases} a_1 &= \widehat{f}_{n,2}(x,h_n)\widehat{f}_{n,4}(x,h_n) - \left(\widehat{f}_{n,3}(x,h_n)\right)^2 \\ a_2 &= \widehat{f}_{n,2}(x,h_n)\widehat{f}_{n,3}(x,h_n) - \widehat{f}_{n,1}(x,h_n)\widehat{f}_{n,4}(x,h_n) \\ a_3 &= \widehat{f}_{n,1}(x,h_n)\widehat{f}_{n,3}(x,h_n) - \left(\widehat{f}_{n,2}(x,h_n)\right)^2 \end{cases}$$

and $\widehat{f}_{n,3}$, $\widehat{f}_{n,4}$ and $\widehat{r}_{n,2}$ are the direct extensions of the definitions given in the Equations (3) and (4). Note also that, it is not very interesting to study $p \geqslant 3$, see Fan and Gijbels [8], pp. 20-22 and 77-80. The argument is that the mean square error increases with p.

Now, we study the consistency of the estimator $\widehat{F}_n^{(1)}(t,h_n|x)$ via the following decomposition:

$$\widehat{F}_n^{(1)}(t,h_n|x) - F(t|x) = \underbrace{\widehat{F}_n^{(1)}(t,h_n|x) - \widehat{\mathbb{E}}\left(\widehat{F}_n^{(1)}(t,h_n|x)\right)}_{(1)} + \underbrace{\widehat{\mathbb{E}}\left(\widehat{F}_n^{(1)}(t,h_n|x)\right) - F(t|x)}_{(2)}$$

where, following the ideas of Deheuvels and Mason (see [7]), the centering term is defined by:

$$\widehat{\mathbb{E}}\left(\widehat{F}_n^{(1)}(t,h_n|x)\right) = \frac{f_{n,2}(x,h_n)r_{n,0}(x,t,h_n) - f_{n,1}(x,h_n)r_{n,1}(x,t,h_n)}{f_{n,0}(x,h_n)f_{n,2}(x,h_n) - f_{n,1}^2(x,h_n)}$$

where $f_{n,j}(x,h_n) = \mathbb{E}\left(\widehat{f}_{n,j}(x,h_n)\right)$ for $j = 0,1,2$ and $r_{n,j}(x,t,h_n) = \mathbb{E}\left(\widehat{r}_{n,j}(x,h_n)\right)$ for $j = 0,1$.

The *random part* (1) is the object of our theorem given in the following Section. Under (F.1-2), (H.1) and (K.1-3), the *deterministic term* (2), so-called bias, converges uniformly to 0 over $(x,t) \in I \times \mathbb{R}$.

2. **Uniform consistency of the local linear estimator**

We have now all the ingredients to state our main results. The uniform law of the logarithm concerning the local linear estimator of the *cond-cdf*, is given in Theorem 2.1 below.

Theorem 2.1. *Under (F.1-2), (H.0-2) and (K.1-3), we have:*

$$\sup_{x\in I}\sqrt{\frac{nh_n}{\log(h_n^{-1})}}\left|\widehat{F}_n^{(1)}(t,h_n|x) - \widehat{\mathbb{E}}\left(\widehat{F}_n^{(1)}(t,h_n|x)\right)\right| \xrightarrow[n\to+\infty]{\mathbb{P}} \sigma_{F,t}(I) \tag{5}$$

where $\sigma_{F,t}^2(I) = 2||K||_2^2 \sup_{x\in I}\left(\frac{F(t|x)(1-F(t|x))}{f_X(x)}\right)$.

Moreover, we have:

$$\sup_{t\in\mathbb{R}}\sup_{x\in I}\sqrt{\frac{nh_n}{\log(h_n^{-1})}}\left|\widehat{F}_n^{(1)}(t,h_n|x) - \widehat{\mathbb{E}}\left(\widehat{F}_n^{(1)}(t,h_n|x)\right)\right| \xrightarrow[n\to+\infty]{\mathbb{P}} \sigma_F(I) \tag{6}$$

where

$$\sigma_F^2(I) = 2||K||_2^2 \sup_{t\in\mathbb{R}} \sup_{x\in I} \left(\frac{F(t|x)(1-F(t|x))}{f_X(x)} \right) = \frac{||K||_2^2}{2 \inf_{x\in I} f_X(x)}.$$

Remark 4.

(1) The matching almost surely convergence result can also be obtained by assuming (H.2-4) instead of (H.0-2).
(2) The terms $\sigma_{F,t}(I)$ and $\sigma_F(I)$ depend upon the unknown density f_X. But it is a minor problem in practice, because, as shown in Deheuvels [5], and Deheuvels and Mason [7], an application of Slutsky's Lemma allows us to replace, without loss of generality, this quantity by $\widehat{f}_{n,0}(x,h_n)$ (or by any other estimator of $f_X(x)$ which is uniformly consistent on I). Indeed, under (F.1-2), (H.0-2), (K.1-3) we have $\sup_{x\in I} \left| \frac{\widehat{f}_{n,0}(x,h_n)}{f_X(x)} - 1 \right| \xrightarrow[n\to+\infty]{\mathbb{P}} 0$.

This last remark yields to the following corollary.

Corollary 1. *Under (F.1-2), (H.0-2), (K.1-3), we have:*

$$\sup_{t\in\mathbb{R}} \sup_{x\in I} \sqrt{\frac{2nh_n}{\|K\|_2^2 \log(h_n^{-1})} \widehat{f}_{n,0}(x,h_n)} \left| \widehat{F}_n^{(1)}(t,h_n|x) - \widehat{\mathbb{E}} \left(\widehat{F}_n^{(1)}(t,h_n|x) \right) \right| \xrightarrow[n\to+\infty]{\mathbb{P}} 1. \tag{7}$$

We introduce the following quantity $L_n(x) := \sqrt{\frac{2nh_n}{\|K\|_2^2 \log(h_n^{-1})} \widehat{f}_{n,0}(x,h_n)}^{-1}$. We have noted at the end of the Section 1 that the bias part can be neglected, then we have the following proposition.

Proposition 1. *Under (F.1-2), (H.0-2) and (K.1-3), and if h_n is such that the bias term $\sup_{t\in\mathbb{R}} \sup_{x\in I} \{L_n(x)\}^{-1} \left| F(t|x) - \widehat{\mathbb{E}} \left(\widehat{F}_n^{(1)}(t,h_n|x) \right) \right| \xrightarrow[n\to+\infty]{} 0$ then we have:*

$$\sup_{t\in\mathbb{R}} \sup_{x\in I} \{L_n(x)\}^{-1} \left| \widehat{F}_n^{(1)}(t,h_n|x) - F(t|x) \right| \xrightarrow[n\to+\infty]{\mathbb{P}} 1. \tag{8}$$

Remark 5.

(1) The matching almost surely convergence result can also be obtained by assuming (H.2-4) instead of (H.0-2).
(2) For our applications in Sections 4 and 5, a reference choice for h_n is given by minimizing the weighted Mean Integrated Square Error (MISE) criteria (see for instance Berlinet [2], Deheuvels [4] or Deheuvels and Mason [7]). A detailed discussion about the theoretical choice of this bandwidth is given in Ferrigno [9]. The asymptotically optimal constant bandwidth is given by:
$$h_n = C(K,F,f_X) n^{-\frac{1}{5}}.$$
(3) The choice of the kernel K is not important in practice. The most common used kernels are the Gaussian, the indicator function over $[-\frac{1}{2},\frac{1}{2}]$, and the Epanechnikov kernels (see for instance Deheuvels [4]). Note that the Gaussian kernel is not compactly supported, but our results can be extended to this case.

3. Uniform asymptotic certainty bands

3.1. **Application to the** *cond-cdf*. We show now how the Proposition 1 can be used to construct uniform asymptotic certainty bands for $F(t|x)$, in the following sense. Under the assumptions of the Proposition 1, we have, for each $0 < \varepsilon < 1$, and as $n \to +\infty$:

$$\mathbb{P}\left\{F(t|x) \in \left[\widehat{F}_n^{(1)}(t, h_n|x) \pm (1+\varepsilon)L_n(x)\right], \text{ for all } (x,t) \in I \times \mathbb{R}\right\} \to 1 \tag{9}$$

and

$$\mathbb{P}\left\{F(t|x) \in \left[\widehat{F}_n^{(1)}(t, h_n|x) \pm (1-\varepsilon)L_n(x)\right], \text{ for all } (x,t) \in I \times \mathbb{R}\right\} \to 0. \tag{10}$$

Whenever (9) and (10) hold jointly for each $0 < \varepsilon < 1$, we have the following corollary:

Corollary 2. *Under (F.1-2), (H.0-2) and (K.1-3), and if h_n is such that the bias term $\sup_{t\in\mathbb{R}} \sup_{x\in I}\{L_n(x)\}^{-1}|F(t, h_n|x) - \widehat{\mathbb{E}}\left(\widehat{F}_n^{(1)}(t, h_n|x)\right)| \xrightarrow[n\to+\infty]{} 0$ then the interval*

$$\left[\widehat{F}_n^{(1)}(t, h_n|x) \pm L_n(x)\right] \tag{11}$$

provides uniform asymptotic certainty bands (at an asymptotic confidence level of 100%) for the cond-cdf $F(t|x)$, uniformly in $(x,t) \in I \times \mathbb{R}$.

Remark 6.

(1) Probability convergence is sufficient for forming certainty bands, and requires less restrictive hypotheses on the bandwidth h_n than the almost surely convergence results. That is why we use only the probability convergence result of the Proposition 1.
(2) Following a suggestion of Deheuvels and Derzko [6], we use, for these upper and lower bounds for $F(t|x)$, the qualification of certainty bands, rather that of confidence bands, because there is no preassigned confidence level $\alpha \in (0,1)$. Some authors (see for instance Deheuvels and Mason [7], or Blondin [3]) have used the term confidence bands.

3.2. **Application to the regression function.** Let $m(x) = \mathbb{E}(Y|X=x)$ the regression function and $\widehat{m}_n^{(1)}(x) = \int y\widehat{F}_n^{(1)}(dy, h_n|x)$ its local linear estimator. The Proposition 1 has the following corollary for the regression function.

Corollary 3. *Under (F.1-2), (H.0-2) and (K.1-3), and if h_n is such that the bias term $\sup_{t\in\mathbb{R}} \sup_{x\in I}\{L_n(x)\}^{-1}\left|F(t|x) - \widehat{\mathbb{E}}\left(\widehat{F}_n^{(1)}(t, h_n|x)\right)\right| \xrightarrow[n\to+\infty]{} 0$ and the variable Y lives in the real interval $[\alpha, \beta]$, then the interval*

$$\left[\widehat{m}_n^{(1)}(x) \pm (\beta - \alpha)L_n(x)\right] \tag{12}$$

provides uniform asymptotic certainty bands (at an asymptotic confidence level of 100%) for the conditional regression function $m(x)$, uniformly in $x \in I$.

3.3. **Application to the conditional quantiles.** Let $0 < \alpha < 1$. We define the conditional α-quantile of the *cond-cdf* by:

$$q_\alpha(x) = \inf\{t \in \mathbb{R} : F(t|x) \geqslant \alpha\}, \quad \text{for all } \alpha \in (0,1).$$

The local linear estimator of the conditional α-quantile is defined by:

$$\widehat{q}_{\alpha,n}^{(1)}(x) = \inf\{t \in \mathbb{R} : \widehat{F}_n^{(1)}(t, h_n|x) \geqslant \alpha\}, \quad \text{for all } \alpha \in (0,1).$$

The Proposition 1 has the following corollary for the conditional quantiles.

Corollary 4. *Under (F.1-2), (H.0-2) and (K.1-3), if h_n is such that the bias term* $\sup_{t\in\mathbb{R}} \sup_{x\in I}\{L_n(x)\}^{-1} \left|F(t|x) - \widehat{\mathbb{E}}\left(\widehat{F}_n^{(1)}(t,h_n|x)\right)\right| \xrightarrow[n\to+\infty]{} 0$ *and if the function* $x \mapsto f_{X,Y}\left(x, q_\alpha(x)\right) \neq 0$ *for all* $x \in I$*, then the interval*

$$\left[\widehat{q}_{\alpha,n}^{(1)}(x) \pm \frac{2L_n(x)f_X(x)}{f_{X,Y}\left(x, q_\alpha(x)\right)}\right] \tag{13}$$

provides uniform asymptotic certainty bands (at an asymptotic confidence level of 100%) for the conditional α-quantile $q_\alpha(x)$, uniformly in $x \in I$.

Remark 7.

(1) The form of these certainty bands is not very useful in practice since the bounds depend upon the unknown conditional density $f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)}$. Nevertheless, this gives the order of the deviation $\left|\widehat{q}_{\alpha,n}^{(1)}(x) - q(x)\right|$.

(2) To give a more practical result, the idea is to replace the conditional density $f_{Y|X}(q_\alpha(x)|x)$ by an estimator $\widehat{f}_{Y|X}\left(\widehat{q}_{\alpha,n}^{(1)}(x)|x\right)$ such that $\sup_{x\in I}\left|\frac{\widehat{f}_{Y|X}\left(\widehat{q}_{\alpha,n}^{(1)}(x)|x\right)}{f_{Y|X}(q_\alpha(x)|x)} - 1\right| \xrightarrow[n\to+\infty]{\mathbb{P}} 0$. This is not the object of the present article, and will be presented in a future work. A review of kernel estimators for the conditional density is given for instance in [15, 14]. We can cite here the kernel estimator of Parzen-Rosenblatt [11, 12].

4. **A simulation study**

In this paragraph, the *cond-cdf* and the certainty bands introduced in Corollary 2 are constructed on simulated data. We considered the case: $X \sim \mathcal{N}(0,1)$ where $\mathcal{N}(0,1)$ denotes the Gaussian distribution with mean 0 and standard deviation 1. We present two models:

(M_1) $Y|X = x$ follows a Beta(a, b) distribution with shape parameters $a = 1$ and $b = 1 + x^2$.
(M_2) $Y|X = x$ follows an Uniform distribution between $-|x|$ and $|x|$.

We worked with the sample sizes $n = 100$ and $n = 500$. For the kernel K, we opted for the Epanechnikov kernel. For the bandwidth, we selected $h_n = n^{-1/5}$. The Figure 1 illustrates the results for the models (M_1) and (M_2) defined above. For each model, we give the graph of a sample $(X_i, Y_i)_{i=1,\dots,n}$, and the *cond-cdf*: the true function $F(t|x)$ is in full line, whereas the estimated conditional distribution $\widehat{F}_n^{(1)}(t, h_n|x)$ is in black dashed line, and certainty bands in grey line, for $x = 0$ and 1.

The confidence bands appear to be adequate. The fact that the true function does not belong to our certainty bands for some points was expected: it is due to the ε term in Equations (9) and (10). For $n = 500$, the results are better than for $n = 100$.

5. **Application in study of the fetal growth**

The study is based on 3606 fetuses autopsied in fetopathologic units of the "Service de foetopathologie et de placentologie" of the Maternité Régionale Universitaire (CHU Nancy, France) between 1996 and 2013. From this dataset, 694 fetuses were carefully selected by exclusion of multiple pregnancies, malformed, macerated or serious ill fetuses, or those with chromosomal abnormalities.

The naive idea, classically used by the fetopathologists or the echographists (see for instance [1], [13]), is to fit a parametric regression model $Y_i = \beta_0 + \beta_1 X_i + \beta_2 X_i^2 + \epsilon_i$ with the assumptions that ϵ_i, for $i = 1, \dots, n$ are independent and follow the Gaussian distribution

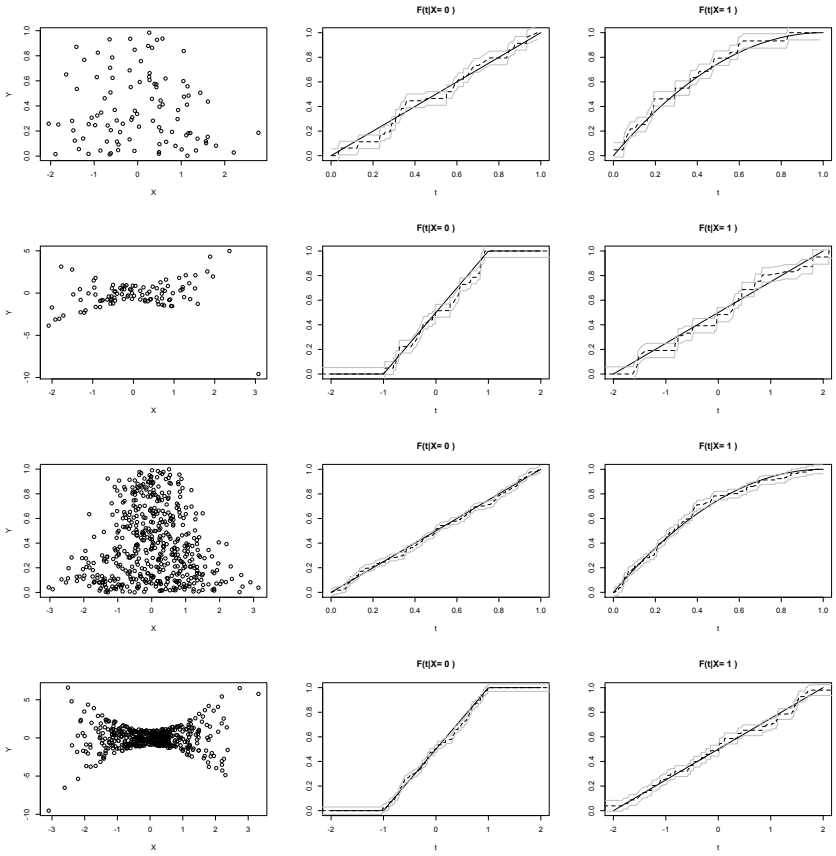

FIGURE 1. From top to bottom: models (M_1) and (M_2) for $n = 100$, and (M_1) and (M_2) for $n = 500$.

$\mathcal{N}(0, \sigma)$. The parameters $\beta_0, \beta_1, \beta_2, \sigma$ are estimated by the least squares method. We use the `R 2.15.1` function `lm`.

The result is shown on the left graph of the Figure 2. This method yields to several problems:

- We obtain heteroscedastic and non-Gaussian errors.
- Moreover, regarding the confidence intervals of the previsions, they show that the prevision uncertainty is not growing with the gestational week: this is not consistent with the medical intuition.
- Another problem is that the global polynomial estimation can not enhance some changes in the growing curve of the fetal weight. For the fetopathologists, such changes are important as they correspond to delicate periods during the intrauterine growth. These change points can not been observed by a global estimation.

For these reasons, the local polynomial estimation is then a non-parametric alternative to the global parametric regression model.

We can conclude, by the observation of the right graph of the Figure 2:

- Our method gives the mean, the confidence intervals and the median weight. Satisfactorily, the confidence intervals show the growing of the prevision uncertainty with the gestational week.
- We observe for instance a change point between the 20th and 25th gestational week on the 0.975 percentile curve. This change point corresponds to the viability date of

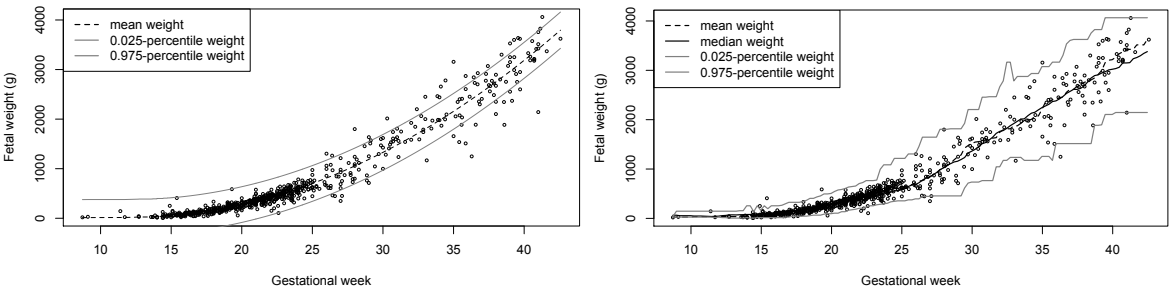

FIGURE 2. Fetal weight during the pregnancy: estimation of mean and quantiles with the second order polynomial regression (left), and with the linear local method (right).

the fetus. We can also remark a decrease of the growing speed around the 35th week. This has also been remarked in the medical article [10], where it is explained that this time corresponds to the regression (in the medical sense) of the placenta. More precise statistical tests to detect the change points of the fetal growth will be presented in a future work.

ACKNOWLEDGEMENT

We thank to Doctors Bernard Foliguet, Jean-Pierre Masutti and Alain Miton of the Service de foetopathologie et de placentologie of the Maternité Régionale Universitaire (CHU Nancy, France), for the fetal data.

REFERENCES

[1] D.G. Altman, L.S. Chitty, *Charts of fetal size: 1. Methodology.*, Br. J. Obstet. Gynaecol., 101 (1994), pp. 29-34.

[2] A. Berlinet, L. Devroye, *A comparison of kernel density estimates*, Publ. Inst. Statist. Univ. Paris, 38 (1994), pp. 3-79.

[3] D. Blondin, *Lois limites uniformes et estimation non paramétrique de la régression*, PhD thesis, Université de Paris 6, France, 2004.

[4] P. Deheuvels, *Estimation non-paramétrique de la densité par histogramme généralisé*, La Revue de Statistique Appliquée, 35 (1977), pp. 5-42.

[5] P. Deheuvels, *Limit laws for kernel density estimators for kernels with unbounded supports*, In Asymptotics in Statistics and Probability, (Ed., M. L. Puri), V.S.P., Amsterdam, 2000.

[6] P. Deheuvels, G. Derzko, *Asymptotic certainty bands for kernel density estimators based upon a bootstrap resampling scheme*, Statistical Models and Methods for Biomedical and Technical Systems, 3 (2008), pp. 171-186.

[7] P. Deheuvels, D.M. Mason, *General asymptotic confidence bands based on kernel-type function estimators*, Statist. Infer. Stochastic Process, 7(3) (2004), pp. 225-277.

[8] J. Fan, I. Gijbels, *Local polynomial modeling and its applications.* Monographs on Statistics and Applied Probability, Chapman and Hall, Vol. 66, 1996.

[9] S. Ferrigno, *Un test d'adéquation global pour la fonction de répartition conditionnelle*, PhD thesis, Université de Montpellier 2, France, 2004.

[10] A.M. Guihard-Costa, *Les variations des vitesses de croissance au cours de la vie foetale*, Bulletins et Mémoires de la Société d'anthropologie de Paris, Nouvelle Série, 5(1-2) (1993), pp.11-20.

[11] E. Parzen, *On estimation of a probability density function and mode*, Ann. Math. Statist., 33 (1962), pp. 1065-1076.

[12] M. Rosenblatt, *Remarks on some nonparametric estimates of a density function*, Ann. Math. Statist., 27 (1956), pp. 832-837.

[13] P. Royston, E.M. Wright, *How to construct "normal ranges" for fetal variables*, Ultrasound Obstet. Gynecol., 11 (1998), pp. 30-8.

[14] E. Youndje, *Convergence properties of the kernel estimator of conditional density*, Rev. Roumaine Math. Pures Appl., 41(7-8), 1996.

[15] E. Youndje, *Contribution à l'estimation non-paramétrique par la méthode du noyau*, Habilitation à diriger des recherches, Université de Rouen et du Havre, France, 2011.