



**HAL**  
open science

# Asymptotics for generalized piecewise linear histograms

A. Berlinet, T. Hobza, I. Vajdad

► **To cite this version:**

A. Berlinet, T. Hobza, I. Vajdad. Asymptotics for generalized piecewise linear histograms. *Annales de l'ISUP*, 2002, XXXXVI (3), pp.3-19. <hal-03641556>

**HAL Id: hal-03641556**

**<https://hal.science/hal-03641556v1>**

Submitted on 14 Apr 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

Pub. Inst. Stat. Univ. Paris  
 XXXXVI, fasc. 3, 2002, 3 à 19

## Asymptotics for generalized piecewise linear histograms<sup>1</sup>

A. Berlinet<sup>2</sup>, T. Hobza<sup>3</sup> and I. Vajda<sup>3</sup>

### Abstract

In a recent paper we have extended the concept of piecewise linear histogram (PLH) introduced by Beirlant, Berlinet and Györfi. The disadvantage of the PLH is that, in many models with probability close to 1, it takes on negative values with probability close to one. We have shown that for all models satisfying mild assumptions, the class of our generalized piecewise linear histograms (GPLH's) contains a bona fide density with probability tending to 1 as the sample size  $n$  increases to infinity. In this paper, under the same assumptions about the model as introduced by the above mentioned authors, the mean integrated absolute error of GPLH's is shown to decrease with the same asymptotic rate  $n^{-2/5}$  as the same error of PLH's. We study the optimization of the binwidth for GPLH's and use the established asymptotic properties to compare the GPLH's with several formerly introduced modifications of histograms. Numerical comparisons are included too.

AMS 1991 subject classification: Primary 62H12. Secondary 62G05.

Key Words: Nonparametric density estimation. Histogram. Piecewise linear histogram. Densitogram. Mean integrated absolute error. Rate of convergence. Optimal binwidth.

## 1 Introduction

We consider a random sample  $X_1, \dots, X_n$  with independent components distributed by a probability measure  $\mu$  on  $\mathbb{R}$ . This measure is assumed to be absolutely continuous with a density  $f$  (the dominating Lebesgue measure is omitted in the integrals that follow). The empirical measure associated with the sample is denoted by  $\mu_n$ , i. e.

$$\mu_n(A) = \frac{1}{n} \sum_{i=1}^n I_A(X_i)$$

for any Borel set  $A \subset \mathbb{R}$ . We start with three definitions, all of them illustrated by Figure 1.

Let  $\mathcal{P}_n = \{A_{nj}\}$  be a partition of the real line into naturally ordered disjoint intervals (bins)  $A_{nj}$  of finite lengths (binwidths)  $h_{nj}$ ,  $j = \dots, -1, 0, 1, \dots$ . We consider the histogram  $g_n$  associated with  $\mathcal{P}_n$ , i. e.

$$g_n(x) = \frac{\mu_n(A_{nj})}{h_{nj}} \quad \text{for all } x \in A_{nj}.$$

<sup>1</sup>Supported by the grants IG CVUT 300009704 and GA ĆR grant 102/99/1137.

<sup>2</sup>University of Montpellier II, Place Bataillon, 34095 Montpellier Cedex, France

<sup>3</sup>Institute of Information Theory and Automation, Czech Academy of Sciences, 182 08 Prague, Czech Republic

Further, let  $\tilde{\mathcal{P}}_n = \{A_{nj}^-, A_{nj}^+\}$  be a refinement of  $\mathcal{P}_n$  where  $A_{nj}^-, A_{nj}^+$  are naturally ordered disjoint intervals of respective lengths  $\alpha_{nj} h_{nj}$ ,  $(1 - \alpha_{nj}) h_{nj}$  for some  $0 < \alpha_{nj} < 1$ , with  $A_{nj}^- \cup A_{nj}^+ = A_{nj}$ . By  $\tilde{g}_n$  we denote the *refined histogram* associated with  $\tilde{\mathcal{P}}_n$ , i. e.

$$\tilde{g}_n(x) = \begin{cases} \frac{\mu_n(A_{nj}^-)}{\alpha_{nj} h_{nj}} & \text{if } x \in A_{nj}^- \\ \frac{\mu_n(A_{nj}^+)}{(1 - \alpha_{nj}) h_{nj}} & \text{if } x \in A_{nj}^+. \end{cases}$$

Since there is one-to-one relation between the refinements  $\tilde{\mathcal{P}}_n$  of  $\mathcal{P}_n$  and the collections of parameters  $\{\alpha_{nj}\}$ , we can replace the "histogram associated with  $\tilde{\mathcal{P}}_n$ " by the "histogram associated with  $\mathcal{P}_n$  and the parameters  $\alpha_{nj}$ ".

Similarly, there is one-to-one relation between the refinements  $\tilde{\mathcal{P}}_n$  of  $\mathcal{P}_n$  and the collections of points

$$a_{nj} = \sup A_{nj}^- = \inf A_{nj}^+.$$

Let  $a_{nj}^-, a_{nj}^+$  be the centers of subintervals  $A_{nj}^-$  and  $A_{nj}^+$ . Then it is easy to verify the one-to-one relation

$$a_{nj} = \alpha_{nj} a_{nj}^+ + (1 - \alpha_{nj}) a_{nj}^-$$

between the collections  $\{\alpha_{nj}\}$  and  $\{a_{nj}\}$ . For

$$\gamma_{nj} = \alpha_{nj} \tilde{g}_n(a_{nj}^+) + (1 - \alpha_{nj}) \tilde{g}_n(a_{nj}^-) \quad (1)$$

this relation implies that the point  $(a_{nj}; \gamma_{nj})$  must belong to the straight line  $\{(x; f_n(x) : x \in \mathbb{R})\}$  passing through the points  $(a_{nj}^-; \tilde{g}_n(a_{nj}^-))$  and  $(a_{nj}^+; \tilde{g}_n(a_{nj}^+))$ , see Figure 1. Since  $a_{nj}^+ - a_{nj}^- = h_{nj}/2$ , the linear function  $f_n(x)$  (depending on  $j$ ) must satisfy the relation

$$f_n(x) = \gamma_{nj} + \frac{2}{h_{nj}} (x - a_{nj}) (\tilde{g}_n(a_{nj}^+) - \tilde{g}_n(a_{nj}^-)) \quad \text{for all } x \in A_{nj}. \quad (2)$$

Formula (2) applied to all  $j$  under consideration defines a *generalized piecewise linear histogram* (GPLH)  $f_n$  associated with  $\mathcal{P}_n$  and parameters  $\{\alpha_{nj}\}$  (or points  $\{a_{nj}\}$ ), cf. Figure 1. The GPLH associated with  $\mathcal{P}_n$  and parameters  $\{\alpha_{nj} = 1/2\}$  reduces to the *piecewise linear histogram* (PLH) associated with  $\mathcal{P}_n$  of Beirlant et al (1999).

The above introduced GPLH is closely related to another piecewise linear modification of histogram, the well known frequency polygon. Consider a sequence of partitions  $\mathcal{Q}_n = \{B_{nj} = [b_{nj}^0, b_{n,j+1}^0]\}$  and let  $f_n$  be the GPLH associated with the partitions

$$\mathcal{P}_n = \{A_{nj} = [b_{n,j-1}^0, b_{n,j+1}^0]\}$$

and points  $a_{nj} = b_{nj}^0$ , and  $\tilde{f}_n$  the GPLH associated with the partitions

$$\tilde{\mathcal{P}}_n = \{\tilde{A}_{nj} = [b_{nj}^0, b_{n,j+2}^0]\}$$

and points  $\tilde{a}_{nj} = b_{n,j+1}^0$ . Then the mapping  $\pi_n : \mathbb{R} \rightarrow [0, \infty)$  defined on  $B_{nj}$  by

$$\pi_n(x) = \begin{cases} f_n(x) & \text{if } b_{nj}^0 \leq x < \frac{b_{nj}^0 + b_{n,j+1}^0}{2} \\ \tilde{f}_n(x) & \text{if } \frac{b_{nj}^0 + b_{n,j+1}^0}{2} \leq x < b_{n,j+1}^0 \end{cases}$$

is the *frequency polygon* (FP) associated with  $\mathcal{Q}_n$  (cf. Scott (1985), or Chapter 4 in Scott (1992) and the recent papers by Minnotte (1996) and Beirlant et al (1999)).

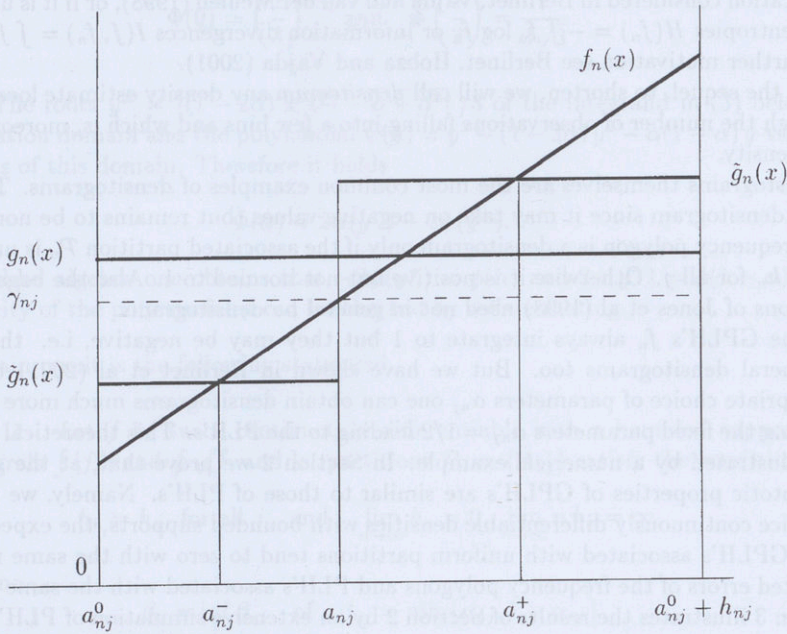


Figure 1: Histograms  $g_n(x)$ ,  $\tilde{g}_n(x)$  and GPLH  $f_n(x)$  on the interval  $A_{n_j} = [a_{n_j}^0, a_{n_j}^0 + h_{n_j}]$ . This interval is divided by the point  $a_{n_j}$  into two subintervals  $A_{n_j}^-$  and  $A_{n_j}^+$  of lengths  $\alpha_{n_j} h_{n_j}$  and  $(1 - \alpha_{n_j}) h_{n_j}$ . The points  $a_{n_j}^-$  and  $a_{n_j}^+$  are centers of subintervals  $A_{n_j}^-$  and  $A_{n_j}^+$ . If  $a_{n_j}$  is the center of  $A_{n_j}$  then the GPLH reduces to the PLH.

In this paper we consider global errors of density estimators  $f_n$ , namely the deviations from the true density  $f$  in the  $L_1$ -space. In other words, by an *error* we mean the  $L_1$ -norm (integral absolute error)

$$\|f - f_n\| = \int |f - f_n|.$$

We consider also the *expected error* (mean integral absolute error, MIAE)  $E\|f - f_n\|$ .

Beirlant et al (1998) proved that for twice continuously differentiable densities  $f$  with a bounded support the expected errors of PLH's and frequency polygons associated with the same uniform partitions tend to zero with the same rate, and that for optimally selected partitions this rate coincides with the best rate  $n^{-2/5}$  achieved in similar situations by the most sophisticated and computationally complicated estimators, such as the kernel estimators.

The numerical simplicity of density estimators is a property important in many applications (cf. e. g. the estimation of distributions characterizing the users of communication networks in Berliant et al (1998), where large amounts of new observations have to be

processed on-line). Of similar practical importance is the need to work with estimates  $f_n$  which are probability densities, i. e. nonnegative and integrating to 1. This requirement can hardly be avoided e. g. if  $f_n$  is used to provide estimates of tail probabilities, as in the application considered in Berlinet, Vajda and van der Meulen (1998), or if it is used to evaluate entropies  $H(f_n) = -\int f_n \log f_n$  or information divergences  $I(f, f_n) = \int f \log(f/f_n)$ . For further motivation see Berlinet, Hobza and Vajda (2001).

In the sequel, to shorten, we will call *densitogram* any density estimate locally defined through the number of observations falling into a few bins and which is, moreover, a bona fide density.

Histograms themselves are the most common examples of densitograms. The PLH is not a densitogram since it may take on negative values (but remains to be normed to 1). The frequency polygon is a densitogram only if the associated partition  $\mathcal{P}_n$  is uniform, i. e.  $h_{nj} = h_n$  for all  $j$ . Otherwise it is positive but not normed to 1. Also the *edge frequency polygons* of Jones et al (1998) need not in general be densitograms.

The GPLH's  $f_n$  always integrate to 1 but they may be negative, i.e. they are not in general densitograms too. But we have shown in Berlinet et al (2001) that by an appropriate choice of parameters  $\alpha_{nj}$  one can obtain densitograms much more often than by using the fixed parameters  $\alpha_{nj} = 1/2$  leading to the PLH's. This theoretical conclusion was illustrated by a numerical example. In Section 2 we prove that, at the same time, asymptotic properties of GPLH's are similar to those of PLH's. Namely, we prove that for twice continuously differentiable densities with bounded supports, the expected errors of all GPLH's associated with uniform partitions tend to zero with the same rate as the expected errors of the frequency polygons and PLH's associated with the same partitions. Section 3 illustrates the results of Section 2 by an extensive simulation of PLH's, GPLH's and frequency polygons. Section 4 contains the proof of the main result of Section 2. Some of the auxiliary results of Section 4 are of their own interest, e. g. the asymptotics of  $L_1$  error for the expected GPLH in Lemma 4, or the asymptotics of the expected  $L_1$  error for the histograms associated with nonuniform partitions in Lemma 7.

## 2 Expected errors of GPLH's

In this section  $f_n$  denotes a GPLH associated with the partitions  $\mathcal{P}_n = \{A_{nj}\}$  and points  $a_{nj} \in A_{nj}$ . We are interested in the expected errors  $E\|f - f_n\|$ . We consider the function

$$\Phi(\alpha) = \int_{-\alpha}^{1-\alpha} |3y^2 - 2(1-2\alpha)y - \alpha(1-\alpha)| dy, \quad 0 < \alpha < 1. \quad (3)$$

Next follow several important properties of this function.

**Lemma 1.** It holds

$$\Phi(\alpha) = \phi(\alpha) + \phi(1-\alpha)$$

where

$$\phi(\alpha) = \frac{2}{27} \left( 1 - 2\alpha + \sqrt{1 - \alpha + \alpha^2} \right) \left[ 1 + 2\alpha(1 - \alpha) + (1 - 2\alpha) \sqrt{1 - \alpha + \alpha^2} \right]$$

is decreasing on  $(0, 1)$  with  $\phi(0) = (2/3)^3$  and  $\phi(1) = 0$ . Consequently  $\Phi(\alpha)$  is symmetric about  $\alpha = 1/2$  and decreasing on  $(0, 1/2)$  with

$$\Phi(0) = \left(\frac{2}{3}\right)^2 \quad \text{and} \quad \Phi\left(\frac{1}{2}\right) = \frac{1}{3\sqrt{3}}. \quad (4)$$

**Proof.** The roots  $y^\pm = [(1 - 2\alpha) \pm \sqrt{1 - \alpha + \alpha^2}] / 3$  of the integrand in (3) belong to the integration domain and the polynomial  $\psi(y) = y^3 - (1 - 2\alpha)y^2 - \alpha(1 - \alpha)y$  vanishes at the ends of this domain. Therefore it holds

$$\Phi(\alpha) = 2\psi(y^-) - 2\psi(y^+).$$

By a standard algebra one obtains that  $-2\psi(y^+) = \phi(\alpha)$  and  $2\psi(y^-) = \phi(1 - \alpha)$ . The monotonicity of the polynomial  $\phi$  can be checked by differentiating. ■

Our main result is the following statement.

**Theorem 1.** Let  $f$  be twice continuously differentiable with a bounded support and finite integrals  $\int |f''|$  and  $\int \sqrt{f}$ , and let partitions  $\mathcal{P}_n = \{A_{nj}\}$  satisfy the conditions

$$h_n = h_{nj} \text{ for all } j \quad \text{and} \quad \lim_{n \rightarrow \infty} h_n = 0, \quad \lim_{n \rightarrow \infty} n h_n = \infty. \quad (5)$$

If the infima

$$\beta_n = \inf_j \beta_{nj} \quad \text{of} \quad \beta_{nj} = \min\{\alpha_{nj}, 1 - \alpha_{nj}\} \quad (6)$$

satisfy the condition

$$\liminf_{n \rightarrow \infty} \beta_n = \beta_* > 0 \quad (7)$$

then  $E\|f_n - f\|$  is larger than

$$\max \left\{ \frac{1}{\sqrt{n} h_n} \frac{2 \int \sqrt{f}}{\sqrt{2\pi(1 - \beta_*)}} + o\left(\frac{1}{\sqrt{n} h_n}\right), \frac{h_n^2}{18\sqrt{3}} \int |f''| + o(h_n^2) \right\} \quad (8)$$

and smaller than

$$\frac{1}{\sqrt{n} h_n} \frac{(1 + \beta_*^2) \int \sqrt{f}}{\beta_* \sqrt{2\pi\beta_*}} + \frac{h_n^2 \Phi(\beta_*)}{6} \int |f''| + o\left(\frac{1}{\sqrt{n} h_n} + h_n^2\right) \quad (9)$$

for  $\Phi$  introduced above.

From here and from Lemma 1 one obtains the following result.

**Corollary 1.** If all assumptions of Theorem 1 hold and (7) takes place with  $\beta_* = 1/2$  then  $E\|f_n - f\|$  is between the bounds

$$\max \left\{ \frac{1}{\sqrt{n} h_n} \frac{2 \int \sqrt{f}}{\sqrt{\pi}} + o\left(\frac{1}{\sqrt{n} h_n}\right), h_n^2 \frac{\int |f''|}{18\sqrt{3}} + o(h_n^2) \right\}$$

and

$$\frac{1}{\sqrt{n} h_n} \frac{5 \int \sqrt{f}}{2\sqrt{\pi}} + h_n^2 \frac{\int |f''|}{18\sqrt{3}} + o\left(\frac{1}{\sqrt{n} h_n} + h_n^2\right).$$

A similar result has been proved by Beirlant et al (1999) in the special case where  $\alpha_{nj} = 1/2$  for all  $n$  and  $j$ , i.e. when  $f_n$  is the PLH. Also the next upper bound is in the case  $\beta_* = 1/2$  close to the bound established in Theorem 5 *ibid* for PLH. Our bound is in fact somewhat weaker since at the place of our  $(1 + \beta_*^2)/(\beta_* \sqrt{2\pi\beta_*})$ , which for  $\beta_* = 1/2$  equals  $5(2\sqrt{\pi}) = 1.41$ , the cited theorem has the slightly less multiplier

$$\sqrt{\frac{2}{\pi}} \left( \frac{\sqrt{5}}{2} + \frac{\ln(2 + \sqrt{5})}{4} \right) = 1.18.$$

**Corollary 2.** If all assumptions of Theorem 1 hold and  $h_n = cn^{-1/5}$  for some  $c > 0$  then  $n^{2/5} E\|f_n - f\|$  is between the bounds

$$\max \left\{ \frac{2 \int \sqrt{f}}{\sqrt{2\pi(1 - \beta_*)c}}, \frac{c^2 \int |f''|}{18\sqrt{3}} \right\} + o(1)$$

and

$$\frac{(1 + \beta_*^2) \int \sqrt{f}}{\beta_* \sqrt{2\pi\beta_*c}} + \frac{c^2 \Phi(\beta_*) \int |f''|}{6} + o(1).$$

The upper bound of the last corollary is minimized by

$$c_{\text{opt}}(\beta_*) = \left( \frac{3(1 + \beta_*^2) \int \sqrt{f}}{2\beta_* \Phi(\beta_*) \sqrt{2\pi\beta_*} \int |f''|} \right)^{2/5} \quad (10)$$

This leads to the following result.

**Corollary 3.** If all assumptions of Theorem 1 hold and  $h_n = c_{\text{opt}}(\beta_*) n^{-1/5}$  then

$$n^{2/5} E\|f_n - f\| \leq 5 \left( \frac{(1 + \beta_*^2) \int \sqrt{f}}{4\beta_* \sqrt{2\pi\beta_*}} \right)^{4/5} \left( \frac{\Phi(\beta_*) \int |f''|}{6} \right)^{1/5} + o(1).$$

In the domain  $0 < \beta_* \leq 1/2$  of possible values of  $\beta_*$ , the bound of Corollary 3 depends strongly on  $1/\beta_*$  and  $\int \sqrt{f}$ , and somewhat less strongly on  $\int |f''|$ . The minimum

$$5 \left( \frac{5 \int \sqrt{f}}{8\sqrt{\pi}} \right)^{4/5} \left( \frac{\int |f''|}{18\sqrt{3}} \right)^{1/5} + o(1)$$

is achieved by the GPLH's which are asymptotically PLH's.

**Remark 1.** The practical meaning of Corollary 3 is that for  $\alpha_{nj}$ 's deviating from  $1/2$  at most by  $\delta_n$  slowly tending to zero, e.g. by  $\delta_n = 1/(2\ln(n+1))$  or  $1/(2\ln\ln(n+1))$ , the asymptotics of the expected  $L_1$  error remains practically the same as in the case of PLH's where all  $\alpha_{nj}$ 's are equal to  $1/2$ . Such deviations extend the possibility of getting a densitogram estimate  $f_n$ .

**Remark 2.** Corollaries 2 and 3 present the same rate of convergence  $n^{-2/5}$  as found by Beirlant et al (1998) for the PLH's, FP's, and also for the edge frequency polygons of Jones et al (1998). (This rate is higher than  $n^{-1/3}$  achieved under the same assumptions by the histogram  $g_n$  associated with  $\mathcal{P}_n$ , and the same as the rate of computationally much more complex density estimators such as the kernel estimators – see Chapter 5 in Devroye and Györfi (1985)). The constant in Corollary 3 is for  $\beta_* \approx 0$  clearly larger than the analogous constants for the mentioned three estimators but for  $\beta_* \approx 1/2$  it is comparable. By inspecting the arguments leading to this asymptotic comparability, one can conclude that it remains to be approximately preserved also if the partitions  $\mathcal{P}_n$  are not exactly uniform. Since even slight nonuniformity of  $\mathcal{P}_n$  sharply decreases the chance to get the FP's or the edge frequency polygons normalized to 1, the relative advantage of GPLH mentioned in Remark 1 can be extended in the case of nonuniform partitions also to the FP's and edge frequency polygons. This was illustrated by an example in Berlinet et al (2001).

### 3 Example: Integrated absolute errors

Let the probability measure  $\mu$  on  $\mathbb{R}$  be described by the exponential distribution function truncated on a nonempty interval  $[0, x_0) \subset \mathbb{R}$ , i. e. by

$$F(x) = \frac{1 - e^{-\theta x}}{1 - e^{-\theta x_0}} \quad \text{for } 0 \leq x \leq x_0, \theta > 0.$$

This truncated exponential model with  $\theta = 2.6$  and  $x_0 = 1.3$  was selected in Berlinet et al (2001) with the aim of providing a simple situation where a partition  $\mathcal{P}_n$  leads to densitogram GPLH's when  $\alpha = 3/4$  and to non-densitogram PLH's when  $\alpha = 1/2$ . In the mentioned work was found a nonuniform partition  $\mathcal{P}_n^0$  of the observation space  $[0, 1.3)$  satisfying this condition. As it could be expected, the MIAE of GPLH slightly exceeded the MIAE of PLH – this is the price paid for the densitogram. The uniform partitions of the interval  $[0, 1.3)$  provide the same chance for GPLH and PLH to be densitograms. Therefore the comparisons of these two estimators with uniform partitions in this model do not seem to be a priori unfair to any of them.

The theory developed in the previous section leads to the uniform partitions which are optimal for the GPLH (more precisely, which are interpreted as optimal). They depend on the true density  $f$  so that their practical applicability is a delicate matter – it depends on what is a priori known about  $f$ , cf. in this respect e.g. Devroye and Györfi (1985) where in Chapter 5 similar optimal partitions were obtained for the histogram and some other estimators, or Beirlant et al (1999) where such partitions were obtained for the PLH and the FP. But, leaving aside the problem of practical applicability, two interesting questions still remain open. (i) How effective this theory is in the sense of actual decrease of MIAE for typical sample sizes  $n$  against some standard and intuitively appealing partitioning, e.g. against the partitioning into the (roughly) equiprobable intervals? (ii) Does this theory, based on the minimization of an upper bound to the MIAE, minimize the MIAE's for typical sample sizes  $n$ ?

This section reports the results of experiments answering both these questions in the above considered model of Berlinet et al (2001) with parameter  $\alpha = 3/4$  of the GPLH for

all three estimators under consideration. As is usual in Statistics, these answers rise a new question: how far one can go with their interpretation beyond the considered model. This issue is out of the scope of the present paper. We will limit ourselves to the mathematical results given in Sections 2 and 4 and to the facts observed in our experiments.

In the mentioned model we evaluated the optimum binwidths

$$h_n(\text{GPLH}) = \frac{1.3}{m_n(\text{GPLH})}, \quad h_n(\text{PLH}) = \frac{1.3}{m_n(\text{PLH})}, \quad h_n(\text{FP}) = \frac{1.3}{m_n(\text{FP})}$$

of the uniform partitions  $\mathcal{P}_n = \{A_{n1}, \dots, A_{nm_n}\}$  for the estimator GPLH with

$$\alpha_{nj} = 3/4 \quad \text{for all } 1 \leq j \leq m_n,$$

and for PLH and FP where, by definition,

$$\alpha_{nj} = 1/2 \quad \text{for all } 1 \leq j \leq m_n.$$

For selected values  $50 \leq n \leq 10\,000$  and the corresponding values of  $m_n \in \{m_n(\text{GPLH}), m_n(\text{PLH}), m_n(\text{FP})\}$ , we simulated 1000 realizations of the sample  $X_1, \dots, X_n$  and calculated the estimates GPLH, PLH and FP, and their integrated absolute errors

$$e_n(\text{GPLH}), \quad e_n(\text{PLH}), \quad e_n(\text{FP}).$$

The sample means and sample standard deviations of these three errors can be found in Tables 1a–1c.

The optimal partition sizes  $m_n(\text{GPLH})$ ,  $m_n(\text{PLH})$ ,  $m_n(\text{FP})$  have been evaluated on the basis of Corollary 3 in Section 4 and in Beirlant et al (1999), by rounding off the values

$$\frac{1.3n^{1/5}}{c_{\text{opt}}(\text{GPLH})}, \quad \frac{1.3n^{1/5}}{c_{\text{opt}}(\text{PLH})}, \quad \frac{1.3n^{1/5}}{c_{\text{opt}}(\text{FP})}$$

where  $c_{\text{opt}}(\text{GPLH})$  is given by (10) for  $\beta_* = 2/5$  and  $c_{\text{opt}}(\text{PLH})$ ,  $c_{\text{opt}}(\text{FP})$  are given after Theorem 7 in Beirlant et al (1999). In these formulas we substitute

$$\int \sqrt{f} = \sqrt{\frac{\theta}{1 - e^{-\theta x_0}}} \int_0^{x_0} e^{-\theta x/2} dx = \frac{2}{\sqrt{\theta}} \sqrt{\frac{1 - e^{-\theta x_0/2}}{1 + e^{-\theta x_0/2}}} \doteq 1.029,$$

$$\int |f''| = \frac{\theta^3}{1 - e^{-\theta x_0}} \int_0^{x_0} e^{-\theta x} dx = \theta^2 = 6.76,$$

and

$$\Phi(2/5) = \int_{-2/5}^{3/5} |3y^2 - 2y/5 - 6/25| dy \doteq 0.196.$$

Therefore

$$c_{\text{opt}}(\text{GPLH}) = 1.353$$

$$c_{\text{opt}}(\text{PLH}) = \left[ \frac{9}{4} \sqrt{\frac{6}{\pi}} \left( \sqrt{5} + \frac{1}{2} \ln(2 + \sqrt{5}) \right) \frac{\int \sqrt{f}}{\int |f''|} \right]^{2/5} \doteq 1.144,$$

and

$$c_{\text{opt}}(\text{FP}) = \left[ \frac{1}{\sqrt{\pi}} \left( \sqrt{2} + \ln(1 + \sqrt{2}) \right) \frac{\int \sqrt{f}}{\int |f''|} \right]^{2/5} \doteq 0.523.$$

The partition sizes  $m_n(\text{GPLH})$ ,  $m_n(\text{PLH})$  and  $m_n(\text{FP})$  obtained in this manner for the selected sample sizes  $50 \leq n \leq 10000$  are in the first columns of Tables 1a-1c.

In Tables 1a-1c one can observe some pleasant and some unpleasant facts. A pleasant fact is that for all sample sizes  $50 \leq n \leq 10000$ , the GPLH with the optimally adapted partitions achieved the absolute minima of the "means"  $\doteq$  MIAE's presented in these tables (the deviations from this rule are due to FP for  $100 \leq n \leq 500$  and for the partitions optimally adapted to the PLH in Table 1b; quantitatively they are of the negligible order of  $10^{-2}$ ). Another pleasant fact partly observable in these tables is that the partitions optimal for the GPLH's in sense of the theory in Section 2, really minimize the expected values of IAE's (to this end compare the "means" of GPLH in Table 1a with those in Table 1b for  $n = 100$  and  $500 \leq n \leq 10000$ , and with those in Table 1c for all  $50 \leq n \leq 10000$ ).

An unexpected fact clearly observable in these tables is that the partitions optimal for PLH and FP in sense of the theory in Beirlant et al (1999) do not minimize the expected IAE's: the "means" of PLH in Table 1a are uniformly less than (or equal to) those in Table 1b and the "means" of FP in Table 1b are considerably less than those in Table 1c. In other words, the partition sizes  $m_n$  in Table 1b (for PLH) are overestimated slightly and in Table 1c (for FP) quite considerably. The partition sizes of Table 1a (for GPLH) practically optimize all three estimators simultaneously.

The unexpected fact mentioned above can be explained by taking into account that the optimal partitions in fact optimize the upper bound of the error (see (10) and Corollary 3). In the case of GPLH the upper bound is probably more tight than in the case of PLH and FP.

$n$	$m_n(\text{GPLH})$	GPLH		PLH		FP	
		mean	stdev	mean	stdev	mean	stdev
50	2	.184	0.64	.192	0.67	.186	0.54
100	2	.148	.041	.151	.045	.159	.037
200	3	.120	.038	.122	.040	.113	.031
500	3	.082	.021	.083	.022	.087	.018
1000	4	.066	.017	.067	.017	.065	.014
2000	4	.049	.011	.049	.011	.053	.010
5000	5	.034	.007	.035	.007	.037	.006
10000	6	.026	.006	.026	.006	.028	.006

**Table 1a.** Sample means and standard deviations for the optimal uniform partition sizes  $m_n$  of GPLH.

$n$	$m_n(\text{PLH})$	GPLH		PLH		FP	
		mean	stdev	mean	stdev	mean	stdev
50	2	.184	.064	.192	.067	.186	.054
100	3	.163	.056	.167	.057	.144	.046
200	3	.120	.038	.122	.040	.113	.031
500	4	.090	.025	.090	.026	.081	.021
1000	5	.072	.018	.072	.019	.063	.015
2000	5	.052	.013	.052	.013	.049	.010
5000	6	.036	.008	.036	.008	.035	.007
10000	7	.028	.006	.028	.006	.027	.005

**Table 1b.** The same as in Table 1a for the optimal uniform partition sizes  $m_n$  of PLH.

$n$	$m_n(\text{FP})$	GPLH		PLH		FP	
		mean	stdev	mean	stdev	mean	stdev
50	6	.345	.084	.347	.084	.269	.067
100	7	.267	.063	.268	.062	.207	.050
200	8	.205	.044	.205	.045	.160	.037
500	9	.138	.028	.137	.028	.107	.022
1000	11	.109	.018	.109	.019	.085	.015
2000	12	.081	.014	.080	.014	.063	.011
5000	15	.057	.009	.057	.009	.044	.007
10000	17	.043	.006	.043	.006	.034	.005

**Table 1c.** The same as in Table 1a for the optimal uniform partitions sizes  $m_n$  of FP.

## 4 Proof of Theorem 1

Throughout this section  $f_n$  denotes the GPLH and  $\tilde{g}_n$  the refined histogram, both associated with the partitions  $\mathcal{P}_n = \{A_{n_j}\}$  and collections of points  $a_{n_j}$  (or the corresponding parameters  $\alpha_{n_j}$ ).

**Lemma 2.** Let  $f(x)$  be twice continuously differentiable in an open set containing  $A_{n_j}$ . If  $h_{n_j} \rightarrow 0$  then it holds uniformly for all  $x \in A_{n_j}$

$$E f_n(x) = f(a_{n_j}) + f'(a_{n_j})(x - a_{n_j}) + \frac{f''(a_{n_j})(1 - 2\alpha_{n_j})}{3} h_{n_j}(x - a_{n_j})$$

$$+ \frac{f''(a_{n_j}) \alpha_{n_j} (1 - \alpha_{n_j})}{6} h_{n_j}^2 + o(h_{n_j}^2).$$

**Proof.** By the definition of  $a^+ = a_{n_j}^+$ ,  $A^+ = A_{n_j}^+$  and  $\tilde{g}_n(x)$ ,

$$E \tilde{g}_n(a^+) = \frac{1}{(1-\alpha)h} \int_{A^+} f = \frac{1}{(1-\alpha)h} \int_a^{a+(1-\alpha)h} f(x).$$

Using the Taylor expansion at  $a = a_{n_j}$ ,

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2}(x-a)^2 + o(h^2), \quad (11)$$

which is valid uniformly for all  $x \in A$  when  $h \rightarrow 0$ , we obtain

$$\begin{aligned} E \tilde{g}_n(a^+) &= \frac{1}{(1-\alpha)h} \int_0^{(1-\alpha)h} \left[ f(a) + f'(a)x + \frac{f''(a)}{2}x^2 + o(h^2) \right] \\ &= f(a) + \frac{f'(a)}{2}(1-\alpha)h + \frac{f''(a)}{6}(1-\alpha)^2 h^2 + o(h^2). \end{aligned}$$

Similarly we obtain from (11)

$$\begin{aligned} E \tilde{g}_n(a^-) &= \frac{1}{\alpha h} \int_{a-\alpha h}^a f(x) \\ &= f(a) - \frac{f'(a)}{2}\alpha h + \frac{f''(a)}{6}\alpha^2 h^2 + o(h^2). \end{aligned}$$

Therefore

$$E [\tilde{g}_n(a^+) - \tilde{g}_n(a^-)] = \frac{f'(a)}{2}h + \frac{f''(a)}{6}(1-2\alpha)h^2 + o(h^2)$$

and, by (1),  $\gamma_n = \gamma_{n_j}$  satisfies the relation

$$\begin{aligned} E \gamma_n &= E [\alpha \tilde{g}_n(a^+) + (1-\alpha) \tilde{g}_n(a^-)] \\ &= f(a) + \frac{f''(a)}{6} [\alpha(1-\alpha)^2 + (1-\alpha)\alpha^2] h^2 + o(h^2) \\ &= f(a) + \frac{f''(a)}{6} \alpha(1-\alpha) h^2 + o(h^2). \end{aligned}$$

Thus it follows from (2) that, uniformly for  $x \in A$ ,

$$\begin{aligned} E f_n(x) &= E \gamma_{n_j} + \frac{2}{h}(x-a) E [\tilde{g}_n(a^+) - \tilde{g}_n(a^-)] \\ &= f(a) + \frac{f''(a)}{6} \alpha(1-\alpha) h^2 + f'(a)(x-a) + \frac{f''(a)(1-2\alpha)}{3} h(x-a) + o(h^2), \end{aligned}$$

which implies the desired asymptotic formula. ■

By combining the result of Lemma 2 with (11) one obtains the asymptotic relation

$$\begin{aligned} f(x) - E f_n(x) &= \frac{f''(a_{n_j})}{6} [3(x-a_{n_j})^2 - 2(1-2\alpha_{n_j})h_{n_j}(x-a_{n_j}) \\ &\quad - \alpha_{n_j}(1-\alpha_{n_j})h_{n_j}^2] + o(h_{n_j}^2) \end{aligned} \quad (12)$$

valid uniformly for all  $x \in A_{n_j}$  when  $h_{n_j} \rightarrow 0$ .

**Lemma 3.** Under the assumptions of Lemma 2 it holds asymptotically for  $h_{nj} \rightarrow 0$

$$\int_{A_{nj}} |f - E f_n| = \frac{|f''(a_{nj})| h_{nj}^3}{6} \Phi(\alpha_{nj}) + o(h_{nj}^3),$$

where  $\Phi(\alpha)$  is the function defined by (3) and characterized by Lemma 1.

**Proof.** Since  $A_{nj} = [a_{nj} - \alpha_{nj} h_{nj}, a_{nj} + (1 - \alpha_{nj}) h_{nj}]$ , the desired formula follows from (3) by substituting  $y = (x - a_{nj})/h_{nj}$  in (12). ■

**Lemma 4.** Let the partitions  $\mathcal{P}_n$  be uniform with binwidths  $h_n$  tending to zero as  $n \rightarrow \infty$ . If  $f$  is continuously differentiable with a bounded support and with  $\int |f''| < \infty$  then (7) implies that asymptotically for  $n \rightarrow \infty$

$$\frac{h_n^2}{18\sqrt{3}} \int |f''| + o(h_n^2) \leq \|f - E f_n\| \leq \frac{h_n^2 \Phi(\beta_*)}{6} \int |f''| + o(h_n^2).$$

**Proof.** Since

$$\|f - E f_n\| = \sum_j \int_{A_{nj}} |f - E f_n|,$$

Lemma 3 and the assumptions about  $f$  imply

$$\|f - E f_n\| = \frac{h_n^2 \xi_n}{6} \int |f''| + o(h_n^2),$$

where  $\xi_n$  is between the minimum and maximum of  $\Phi(\alpha_{nj})$  taken over all  $j$  positively contributing to the sum above. By the monotonicity of  $\Phi(\alpha)$  in Lemma 1 and by (4),

$$\frac{1}{3\sqrt{3}} \leq \xi_n \leq \Phi(\beta_n)$$

for  $\beta_n$  considered in (7), which proves the desired asymptotic relation with  $\Phi(\beta_n)$  at the place of  $\Phi(\beta_*)$ . Since  $\int |f''|$  is assumed to be finite and the monotonicity of  $\Phi$  implies

$$\limsup_{n \rightarrow \infty} \Phi(\beta_n) = \Phi(\beta_*),$$

the proved relation remains valid with  $\Phi(\beta_n)$  replaced by  $\Phi(\beta_*)$ . ■

**Lemma 5.** It holds  $\|f - E f_n\| \leq E \|f - f_n\|$  and  $\|f - f_n\| \geq |\tilde{g}_n - E \tilde{g}_n|$ .

**Proof.** The first inequality follows from the convexity of the  $L_1$  norm and from the Jensen inequality. To prove the second inequality take into account that, by the definition of histograms  $g_n, \tilde{g}_n$  and by (1) and (2), it holds for every  $A \in \{A_{nj}^-, A_{nj}^+\}$  and every  $x \in A$

$$\tilde{g}_n(x) - E \tilde{g}_n(x) = [\mu_n(A) - E \mu_n(A)]/\lambda(A) \quad \text{and} \quad \int_A f_n = \mu_n(A), \quad (13)$$

where  $E \mu_n(A) = \int_A f$  and

$$\lambda(A) = \begin{cases} \alpha_{nj} h_{nj} & \text{if } A = A_{nj}^-, \\ (1 - \alpha_{nj}) h_{nj} & \text{if } A = A_{nj}^+ \end{cases}$$

is the Lebesgue measure of  $A$ . Therefore

$$\begin{aligned} \int_A |\tilde{g}_n - E \tilde{g}_n| &= |\mu_n(A) - E \mu_n(A)| \\ &= \left| \int_A f_n - \int_A f \right| \leq \int_A |f_n - f|. \end{aligned}$$

Since  $A_{nj}^-$  and  $A_{nj}^+$  are disjoint and their union is  $A_{nj}$ , it follows that

$$\int_{A_{nj}} |g_n - E g_n| \leq \int_{A_{nj}} |f_n - f|.$$

The desired inequality between the norms follows by summing up both sides of this inequality over all  $A_{nj}$  from  $\mathcal{P}_n$ . ■

**Lemma 6.** For  $\beta_n$  defined by (6)

$$\|f_n - E f_n\| \leq \frac{1 + \beta_n^2}{2\beta_n} \|\tilde{g}_n - E \tilde{g}_n\|$$

and the equality is attainable in the class of all  $f_n$  under consideration.

**Proof.** Consider for simplicity  $A_{nj} = [0, h_{nj})$  and put

$$h = h_{nj}, \quad \alpha = \alpha_{nj}, \quad A^- = A_{nj}^- = [0, \alpha h) \quad \text{and} \quad A^+ = A_{nj}^+ = [\alpha h, h).$$

Then it follows from (1) and (2) that for every  $x \in A_{nj}$

$$f_n(x) = \frac{\mu_n(A^-)}{\alpha h} \left( 1 + \alpha - \frac{2x}{h} \right) + \frac{\mu_n(A^+)}{(1 - \alpha)h} \left( \frac{2x}{h} - \alpha \right)$$

and, consequently,

$$E f_n(x) = \frac{E \mu_n(A^-)}{\alpha h} \left( 1 - \alpha - \frac{2x}{h} \right) + \frac{E \mu_n(A^+)}{(1 - \alpha)h} \left( \frac{2x}{h} - \alpha \right)$$

where

$$E \mu_n(A^-) = \int_{A^-} f \quad \text{and} \quad E \mu_n(A^+) = \int_{A^+} f.$$

Define independent pair of random variables

$$Y^\pm = \mu_n(A^\pm) - E \mu_n(A^\pm) = \mu_n(A^\pm) - \int_{A^\pm} f.$$

Then for every  $x \in A_{nj} = [0, h)$

$$f_n(x) - E f_n(x) = \frac{Y^-}{\alpha h} \left(1 + \alpha - \frac{2x}{h}\right) + \frac{Y^+}{(1-\alpha)h} \left(\frac{2x}{h} - \alpha\right)$$

and, according to (13),

$$\tilde{g}_n(x) - E \tilde{g}_n(x) = \frac{Y^-}{\alpha h} I_{A^-}(x) + \frac{Y^+}{(1-\alpha)h} I_{A^+}(x).$$

Therefore

$$\int_{A_{nj}} |\tilde{g}_n - E \tilde{g}_n| = |Y^-| + |Y^+|$$

and

$$\begin{aligned} \int_{A_{nj}} |f_n - E f_n| &\leq \frac{|Y^-|}{\alpha h} \int_0^h \left|1 + \alpha - \frac{2x}{h}\right| + \frac{|Y^+|}{(1-\alpha)h} \int_0^h \left|\frac{2x}{h} - \alpha\right| \\ &\leq \max \left\{ \frac{1}{\alpha h} \int_0^h \left|1 + \alpha - \frac{2x}{h}\right|, \frac{1}{(1-\alpha)h} \int_0^h \left|\frac{2x}{h} - \alpha\right| \right\} (|Y^-| + |Y^+|) \\ &= \max \left\{ \frac{1 + \alpha^2}{2\alpha}, \frac{1 + (1-\alpha)^2}{2(1-\alpha)} \right\} \int_{A_{nj}} |\tilde{g}_n - E \tilde{g}_n| \\ &= \frac{1 + \beta_{nj}^2}{2\beta_{nj}} \int_{A_{nj}} |\tilde{g}_n - E \tilde{g}_n| \end{aligned}$$

for  $\beta_{nj}$  introduced in (6). If  $\beta_{nj} = \alpha$  then both inequalities reduce to the equalities when  $Y^+ = 0$  and if  $\beta_{nj} = 1 - \alpha$  then the same takes place when  $Y^- = 0$ . The desired statement follows directly from here. ■

**Lemma 7.** Let  $f$  be continuous with a compact support and  $\int \sqrt{f} < \infty$ , and let the binwidths  $h_{nj}$  satisfy (5). Then under (7) it holds asymptotically for  $n \rightarrow \infty$

$$\frac{1}{\sqrt{nh_n}} \sqrt{\frac{2}{\pi(1-\beta_*)}} \int \sqrt{f} + o\left(\frac{1}{\sqrt{nh_n}}\right) \leq E \|\tilde{g}_n - E \tilde{g}_n\| \leq \frac{1}{\sqrt{nh_n}} \sqrt{\frac{2}{\pi\beta_*}} \int \sqrt{f} + o\left(\frac{1}{\sqrt{nh_n}}\right). \quad (14)$$

**Proof.** Let  $A$  be the set introduced in the proof of Lemma 5 and put  $p = p(A) = \int_A f$ . By (13),

$$E \int_A |\tilde{g}_n - E \tilde{g}_n| = E |\mu_n(A) - p|.$$

If  $p = 0$  then  $E |\mu_n(A) - p| = 0$ . For  $p \neq 0$  the Berry-Esseen inequality implies

$$\left| E \left| \frac{n\mu_n(A) - np}{\sqrt{np(1-p)}} \right| - E|Y| \right| \leq \frac{C}{\sqrt{n}},$$

where  $Y$  is the standard normal random variable so that  $E|Y| = \sqrt{2/\pi}$ , and  $C$  is a universal constant. Therefore for all  $A$  under consideration

$$\left| E \int_A |\tilde{g}_n - E \tilde{g}_n| - \sqrt{\frac{2p(1-p)}{\pi n}} \right| \leq \frac{C\sqrt{p(1-p)}}{n}.$$

Consequently

$$\left| E \|\tilde{g}_n - E \tilde{g}_n\| - \sqrt{\frac{2}{\pi}} \sum_A \sqrt{\frac{p(1-p)}{n}} \right| \leq \frac{C}{n} \sum_A \sqrt{p(1-p)},$$

where the sum extends over  $A \in \{A_{nj}^-, A_{nj}^+\}$  and over all  $j$ . In the next step we make use of the Lebesgue measure  $\lambda(A)$  introduced in the proof of Lemma 5, and of the inequalities

$$h_n \beta_n \leq h_n \beta_{nj} \leq \lambda(A) \leq h_n(1 - \beta_{nj}) \leq h_n(1 - \beta_n)$$

obviously valid for  $\beta_{nj}$  and  $\beta_n$  figuring in (6). Let us define

$$p_{nj}^\pm = \int_{A_{nj}^\pm} f$$

and

$$\begin{aligned} c_{nj} &= \sum_{A \in \{A_{nj}^-, A_{nj}^+\}} \sqrt{\lambda(A) p(1-p)} \\ &= \sqrt{h_n \alpha_{nj} p_{nj}^-(1 - p_{nj}^-)} + \sqrt{h_n (1 - \alpha_{nj}) p_{nj}^+(1 - p_{nj}^+)}. \end{aligned}$$

Then there exists

$$\beta_n < \gamma_n < 1 - \beta_n \tag{15}$$

such that

$$\sum_{A \in \{A_{nj}^-, A_{nj}^+\}} \sqrt{p(1-p)} = \frac{c_{nj}}{\sqrt{h_n \gamma_n}}$$

and consequently also such that

$$\sum_A \sqrt{p(1-p)} = \frac{1}{\sqrt{h_n \gamma_n}} \sum_j c_{nj}.$$

Now, by taking into account the inequalities

$$\begin{aligned} \left| \sqrt{p_{nj}^-(1 - p_{nj}^-)} - \sqrt{h_n \alpha_{nj} f(a_{nj}^-)} \right| &\leq p_{nj}^- + \left| \sqrt{p_{nj}^-} - \sqrt{h_n \alpha_{nj} f(a_{nj}^-)} \right| \\ &\leq p_{nj}^- + \sqrt{h_n \alpha_{nj}} \sqrt{\left| \frac{p_{nj}^-}{h_n \alpha_{nj}} - f(a_{nj}^-) \right|} \end{aligned}$$

and

$$\left| \sqrt{p_{nj}^+(1 - p_{nj}^+)} - \sqrt{h_n (1 - \alpha_{nj}) f(a_{nj}^+)} \right| \leq p_{nj}^+ + \sqrt{h_n (1 - \alpha_{nj})} \sqrt{\left| \frac{p_{nj}^+}{h_n (1 - \alpha_{nj})} - f(a_{nj}^+) \right|}$$

for the points  $a_{nj}^- \in A_{nj}^-$  and  $a_{nj}^+ \in A_{nj}^+$  defined in Section 1, and the continuity of  $f$  on its compact support, we can argue that

$$\left| \sum_j c_{nj} - \sum_j \left[ \sqrt{h_n \alpha_{nj}} f(a_{nj}^-) + \sqrt{h_n (1 - \alpha_{nj})} f(a_{nj}^+) \right] \right| = O(\sqrt{h_n}).$$

Taking into account moreover the finiteness of  $\int \sqrt{f}$ , it follows from here

$$\frac{1}{\sqrt{nh_n}} \left| \sum_j c_{nj} - \int \sqrt{f} \right| = o\left(\frac{1}{\sqrt{nh_n}}\right).$$

In view of what has been said above, this implies

$$\left| \frac{1}{\sqrt{n}} \sqrt{\frac{2}{\pi}} \sum_A \sqrt{p(1-p)} - \frac{1}{\sqrt{nh_n \gamma_n}} \sqrt{\frac{2}{\pi}} \int \sqrt{f} \right| = o\left(\frac{1}{\sqrt{nh_n \gamma_n}}\right)$$

and

$$\left| E\|\tilde{g}_n - E\tilde{g}_n\| - \frac{1}{\sqrt{nh_n \gamma_n}} \sqrt{\frac{2}{\pi}} \int \sqrt{f} \right| = o\left(\frac{1}{\sqrt{nh_n \gamma_n}}\right)$$

for  $\gamma_n$  satisfying (15). By using similar argument as at the end of proof of Lemma 4, we obtain from here under (7) the desired relation (14). ■

**Proof of Theorem 1.** By the triangle inequality and Lemma 6,

$$\begin{aligned} E\|f_n - f\| &\leq E\|f_n - E f_n\| + \|f - E f_n\| \\ &\leq \frac{1 + \beta_n^2}{2\beta_n} E\|\tilde{g}_n - E\tilde{g}_n\| + \|f - E f_n\| \end{aligned}$$

so that the upper bound in Theorem 2 follows from the upper bounds in Lemmas 4 and 7. Further, by Lemma 5,

$$E\|f_n - f\| \geq \min\{E\|\tilde{g}_n - E\tilde{g}_n\|, \|f - E f_n\|\}$$

so that the lower bound in Theorem 2 follows from the lower bounds in Lemmas 4 and 7. ■

## References

- [1] Beirlant, J., Berlines, A. and Györfi, L., "On piecewise linear density estimates". *Statistica Neerlandica* 53 (1999), 287-308.
- [2] Berlines, A., Vajda, I. and van der Meulen, E. C., "Asymptotic rate of approximation of Barron density estimates". *Transactions IEEE on Information Theory* 44 (1998), 999-1009.

- [3] Berlinet, A., Hobza, T. and Vajda, I., "Generalized piecewise linear histograms". (2001), submitted.
- [4] Devroye, L. and Györfi, L., *Nonparametric Density Estimation: The  $L_1$  View*. Wiley, New York 1985.
- [5] Hobza, T., *Simulations of Statistical Estimates* (in Czech, MSc theses). Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University Prague, 1998.
- [6] Jones, M. C., Samiuddin, M., Al-Harbey, A. H. and Maatouk, T. A. H., "The edge frequency polygon". *Biometrika* 85 (1998), 235-239.
- [7] Massart, P., "The tight constant in the Dvoretzky-Kiefer-Wolfowitz inequality". *Annals of Probability* 18 (1990), 1264-1283.
- [8] Minotte, M., "The bias-optimized frequency polygon". *Journal of Computational Statistics* 11 (1996), 35-48.
- [9] Scott, D. W., "Frequency polygons: Theory and application". *Journal of the American Statistical Association* 80 (1985), 348-354.
- [10] Scott, D. W., *Multivariate Density Estimation*. Wiley, New York 1992.

**Keywords.** Local Time, Occupation Density, Semimartingales, Density Estimation, Gaussian Processes, Irregular Trajectories, Numerical Approximation.

**Résumé.** On donne un tour d'horizon des concepts de Temps Local de semimartingales et de Densités d'Occupation, afin d'établir une cartographie de l'utilisation de chaque statistique, au regard de l'estimation recherchée dans des processus à temps continu. Des exemples du lien privilégié entre ces deux statistiques, comme de leurs différences, introduisent leur exploitation dans le cadre de l'estimation de régression.

**Mots Clés.** Temps Local, Densité d'Occupation, Semimartingales, Estimation de densité, Processus Gaussiens, Trajectoires Irrégulières, Approximation Numérique.