A STRONG LAW OF THE EMPIRICAL
DENSITY FUNCTION

by
P. RÉVÉSZ (Budapest)

Summary
Let $X_1, X_2, \ldots$ be a sequence of i.i.d. r.v.'s with a twice differentiable density function $f$. Suppose that $f$ is vanishing outside of the interval $[0, 1]$, strictly positive inside and $|f''|$ is bounded. Further let $\lambda$ be an arbitrary density function satisfying some regularity conditions and define the empirical density function $f_n$ as

$$f_n(x) = (nh_n)^{-1} \sum_{k=1}^{n} \lambda((x - X_k)h_n^{-1})$$

where $\{h_n\}$ is a decreasing sequence of positive numbers tending to 0 and satisfying some further restrictions. Then for any $\varepsilon(>0)$ we have

$$\lim_{n \to \infty} \left( \frac{nh_n}{2A^2 \log h_n^{-1}} \right)^{1/2} \sup_{\varepsilon < x < 1-\varepsilon} \frac{|f_n(x) - f(x)|}{f^2(x)} = 1 \text{ a.s.}$$

where $A^2 = \int_{-\infty}^{+\infty} \lambda^2(x) dx$.

1. Introduction

Let $X_1, X_2, \ldots$ be a sequence of i.i.d. r.v.'s with density function $f$. Suppose that

1.a. $f$ is vanishing outside the interval $[0, 1]$,
1.b. $f$ is twice differentiable in $(0, 1)$ and $|f''| \leq C$,
1.c. $f$ is strictly positive in $(0, 1)$ and $f \geq \alpha > 0$.

Let $\lambda$ be a density function for which

2.a. $\lambda \leq C$,
2.b. $\lambda(-x) = \lambda(x)$,
2.c. $\lim_{x \to \pm \infty} x^n \lambda(x) = 0$,

2.d. $\lambda$ is twice differentiable in an interval $-\infty < -a < x < a < +\infty$, vanishing outside and $|\lambda''| \leq C$ in $(-a, +a)$.

Let $\{h_n\}$ be a sequence of positive numbers satisfying the conditions:

AMS (MOS) subject classifications (1970). Primary 60F15; Secondary 60G05.
Key words and phrases. Empirical density function, Gaussian Processes.
3.a. \( n^2 \log h_n \rightarrow 0 \), \( nh_n \rightarrow \infty \),

3.b. \( \frac{\log^4 h_n}{nh_n \log h_n^{-1}} \rightarrow 0 \), \( \frac{nh_n^3}{\log h_n^{-1}} \rightarrow 0 \).

Define the empirical density function \( f_n \) as

\[
(1) \quad f_n(x) = (nh_n)^{-1} \sum_{k=1}^{n} \lambda((x - X_k)h_n^{-1}) = h_n^{-1} \int_0^1 \lambda((x - y)h_n^{-1})dF_n(y)
\]

where

\[
F_n(y) = \frac{1}{n} \sum_{k=1}^{n} I_{(-\infty, y]}(X_k)
\]

is the empirical distribution function based on the sample \( X_1, X_2, \ldots, X_n \) and \( I_{(-\infty, y]} \) is the indicator of \((-\infty, y]\).

Several theorems state that \( f_n \) is a good approximation of \( f \) (see e.g. [1], [5]). Our main goal is to prove

**Theorem 1.** Suppose that Conditions 1.–3. are satisfied. Then for any \( \varepsilon(>0) \) we have

\[
(2) \quad \lim_{n \rightarrow \infty} \left( \frac{nh_n}{2A^2 \log h_n^{-1}} \right)^{1/2} \sup_{|x|<1-\varepsilon} \left| \frac{f_n(x) - f(x)}{f_1^{1/2}(x)} \right| = 1 \quad a.s.
\]

where \( A^2 = \int_{-\infty}^{\infty} \lambda^2(x)dx \).

Taking stochastic convergence in (2) instead of almost sure convergence we get a simple and well-known statement (see [1]). The problems of almost sure convergence were investigated by REISS [4] who obtained somewhat weaker results.

The proof of our Theorem 1 is based on some results regarding for Gaussian Processes. Let us recall the most important definitions.

**Definition of the Wiener Process in \( \mathbb{R}^1 \) and \( \mathbb{R}^2 \).** A separable Gaussian Process \( W(x) (x \geq 0) \) resp. \( W(x, y) (x, y \geq 0) \) is called a **Wiener Process** if \( EW(x) = 0 \) and \( EW(x_1)W(x_2) = x_1 \wedge x_2 \), resp. \( EW(x, y) = 0 \),

\[
EW(x_1, y_1)W(x_2, y_2) = (x_1 \wedge x_2)(y_1 \wedge y_2).
\]

**Definition of the Brownian Bridge.** Let \( W(x) \) be a Wiener Process then the process \( B(x) = W(x) - xW(1) \) \((0 \leq x \leq 1)\) is called a **Brownian Bridge**.

**Definition of the Kiefer Process.** Let \( W(x, y) \) be a Wiener Process then the process \( K(x, y) = W(x, y) - xW(1, y) \) \((0 \leq x \leq 1, 0 \leq y)\) is called a **Kiefer Process**.