EMPIRICAL DISTRIBUTION FUNCTIONS:
THE CONVERGENCE RATE WITH RESPECT
TO PSEUDOMOMENTS

E. Gordienko, J. Ruiz de Chavez
1,2Department of Mathematics
UAM-Iztapalapa
Av. San Rafael Atlixco No. 186
Col. Vicentina, 09340, Iztapalapa, CDMX, MEXICO

Abstract: Let \( \hat{F}_n(x) = \frac{1}{n} \sum_{k=1}^{n} \mathbb{I}_{X_k \leq x}, \ x \in \mathbb{R} \) be the empirical distribution function over a sample \( \{X_k, k \geq 1\} \) of i.i.d random variables with a common distribution function \( F \). We examine the conditions under which the expected \( p \)-th difference pseudomoments
\[
E \kappa_p(F, F_n) := E \left[ p \int_{-\infty}^{\infty} |x|^{p-1} |F(x) - \hat{F}_n(x)| dx \right]
\]
are of order \( \frac{1}{\sqrt{n}} \) as \( n \to \infty \).

AMS Subject Classification: 60K25, 90B22
Key Words: empirical distribution, pseudomoments, rate of convergence, M\( |GI|\)\(^{\infty}\) queue

1. Motivation

Given \( p \geq 1 \), the \( p - th \) difference pseudomoment
\[
\kappa_p(F_X, F_Y) := p \int_{-\infty}^{\infty} |x|^{p-1} |F_X(x) - F_Y(x)| dx
\]
is finite for any pair of random variables \( X, Y \in L_p = L_p(\Omega, \mathcal{F}, P) \). In (1), \( F_X \) and \( F_Y \) stand for the corresponding distribution functions.

It is well-known (see e.g [4]) that \( \kappa_p \) defines a metric on the space of distribution functions of random variables from \( L_p \), and that this metric is widely used in limits theorems and inequalities expressing the quantitative continuity of mathematical models (see, for instance [3], [2], and [4]).

Received: December 2, 2018
Revised: December 17, 2018
Published: December 18, 2018

Correspondence author
To deal with the problems of continuity (stability) of the model, often one needs to approximate an unknown distribution function $F$ by its empirical counter-part, for instance, by the empirical distribution function:

$$
\hat{F}_n(x) := \frac{1}{n} \sum_{k=1}^{n} \mathbb{I}_{\{X_k \leq x\}}, \quad x \in \mathbb{R}.
$$

Here $X_1, X_2, \ldots$ are i.i.d copies ("observations") of random variable $X$ with the distribution function $F$. In this respect, one would require to estimate the vanishing rate of $E\kappa_p(F, \hat{F}_n)$ as $n \to \infty$.

For $p = 1$, $\kappa_1$ in (1) is a well-known representation of the Kantorovich metric. In the preprint [1] it is proven that for integrable random variables $E\kappa_1(F, \hat{F}_n) \to 0$ as $n \to \infty$, and that the finiteness of the integral $\int_{-\infty}^{\infty} \sqrt{F(x)(1-F(x))} \, dx$ is the necessary and sufficient condition for $E\kappa_1(F, \hat{F}_n) = O\left(\frac{1}{\sqrt{n}}\right)$.

It seems an open problem to show for $p > 1$ that $\int_{-\infty}^{\infty} |x|^p dF(x) < \infty$ implies $E\kappa_p(F, \hat{F}_n) \to 0$ as $n \to \infty$.

The focus of the present note is to show that $E\kappa_p(F, \hat{F}_n)$ cannot converge to zero faster than $\frac{\text{const}}{\sqrt{n}}$, and to give the conditions sufficient for

$$E\kappa_p(F, \hat{F}_n) = O\left(\frac{1}{\sqrt{n}}\right) \text{ as } n \to \infty.
$$

2. The results

Let $X \in L_p$ be a random variable with the distribution function $F$. Then, as it is easy to verify,

$$I_{p,F} := p \int_{-\infty}^{\infty} |x|^{p-1}|F(x)(1-F(x))| \, dx < \infty.
$$

**Proposition 1.**

$$E\kappa_p(F, \hat{F}_n) \geq \frac{1}{\sqrt{n}} \frac{1}{\sqrt{2}} I_{p,F}, \quad n = 1, 2, \ldots \tag{4}$$

**Proof.** In view of (1),
EMPIRICAL DISTRIBUTION FUNCTIONS...

\[ E \kappa_p \left( F; \hat{F}_n \right) = E \int_{-\infty}^{\infty} p|x|^{p-1}|F(x) - \hat{F}_n(x)| \, dx. \]

Applying the Fubini theorem to the right-hand side of this equality, we obtain:

\[ E \kappa_p \left( F; \hat{F}_n \right) = \int_{-\infty}^{\infty} p|x|^{p-1} E \left( |F(x) - \hat{F}_n(x)| \right) \, dx. \] (5)

Then, using Lemma 3.4 in [1] and the arguments given in the proof of Theorem 3.1 in [1], we can show that

\[ E|F(x) - \hat{F}_n(x)| \geq \frac{1}{\sqrt{2n}} F(x) (1 - F(x)). \] (6)

The combination of (3), (5) and (6) provides inequality (4).

**Proposition 2.**

(a) \[ E \kappa_p \left( F; \hat{F}_n \right) \leq \frac{1}{\sqrt{n}} J_{p,F}, \quad n = 1, 2, \ldots \] (8)

(b) If for some \( \alpha > 0 \),

\[ H := E|X|^{2p+\alpha} < \infty, \quad \text{then} \]

\[ E \kappa_p \left( F; \hat{F}_n \right) \leq \frac{1}{\sqrt{n}} \left[ 1 + \frac{4H^{\frac{1}{2p}}}{\alpha} \right], \quad n = 1, 2, \ldots \] (9)

**Proof.** (a) To prove (8) we use the arguments similar to those given in the proof of Theorem 3.2 in [1]. By (5),

\[ E \kappa_p \left( F; \hat{F}_n \right) \leq \int_{-\infty}^{\infty} p|x|^{p-1} \left( E \left[ F(x) - \hat{F}_n(x) \right]^2 \right)^{\frac{1}{2}} \, dx. \] (10)

In view of (2),

\[ E \left[ F(x) - \hat{F}_n(x) \right]^2 = \frac{1}{n} V ar \left( I_{\{X_1 \leq x\}} \right) = \frac{1}{n} F(x) \left( 1 - F(x) \right). \] (11)

Now (7), (10) and (11) yield inequality (8).
(b) Rewriting (7) as follows

\[ j_{p,F} = \int_{-1}^{1} p|x|^{p-1} [F(x)(1 - F(x))]^{\frac{2}{p}} dx + \int_{1}^{\infty} p|x|^{p-1} [F(x)(1 - F(x))]^{\frac{2}{p}} dx \]

\[ + \int_{-\infty}^{-1} p|x|^{p-1} [F(x)(1 - F(x))]^{\frac{2}{p}} dx =: j_1 + j_2 + j_3 \tag{12} \]

and observing that \( F(1 - F) \leq \frac{1}{4} \), we find that in (12),

\[ j_1 \leq 1. \tag{13} \]

For \( x \geq 1 \) by the Markov inequality,

\[ F(x)(1 - F(x)) \leq 1 - F(x) \leq \frac{H}{x^{2p+\alpha}}. \]

Hence in (12)

\[ j_2 \leq pH^{\frac{1}{2}} \int_{1}^{\infty} x^{p-1} \frac{1}{x^p x^{\frac{\alpha}{2}}} dx = \frac{2pH^{\frac{1}{2}}}{\alpha}. \tag{14} \]

For \( x \leq -1 \), we have:

\[ F(x)(1 - F(x)) \leq F(x) \leq P(|X| \geq |x|) \leq \frac{H}{|x|^{2p+\alpha}}, \]

and similarly to (14), in (12)

\[ j_3 \leq \frac{2pH^{\frac{1}{2}}}{\alpha}. \tag{15} \]

Finally, joining (8), (10), (12), (14) and (15) gives inequality (9).

For non negative random variables with light-tailed distributions inequality (9) can be improved as follows.

**Corollary 3.** Let for some \( \lambda > 0 \), \( M := E e^{\lambda x} < \infty \), and \( X \geq 0 \). Then,

\[ E\kappa_1(F,F_n) \leq \frac{1}{\sqrt{n}} \frac{2pM^\frac{p}{2}}{\lambda^p} \Gamma(p), \quad n = 1, 2, \ldots \tag{16} \]
Indeed,
\[
\int_0^\infty px^{p-1}P(X > x)^{1/2}dx \leq pM^{1/2} \int_0^\infty x^{p-1}e^{-\lambda x}dx = pM^{1/2} \Gamma(p).
\]

The following example shows that the condition \( X \in \mathbb{L}_2p \) does not guarantee the finiteness of \( J_{p,F} \) in (7) and (8).

**Example 4.** Let, for instance, \( p = 2 \), and
\[
F(x) = \begin{cases} 
1 - \frac{e}{x^4 \log^2(x)}, & x \geq 2, \\
0, & x < 2.
\end{cases}
\]

Then \( X \in \mathbb{L}_4 \), but \( J_{2,F} = \infty \). Indeed, choosing \( x_0 > 2 \) such that \( F(x_0) \geq \frac{1}{2} \) we get that
\[
J_{2,F} \geq \frac{2}{\sqrt{2}} \int_{x_0}^\infty x \frac{\sqrt{e}}{x^2 \log(x)}dx = \infty.
\]

**Example 5.** Consider two \( M|GI|1|\infty \) queues: \( Q \) named “the original one”, and \( \tilde{Q} \) interpreted as an approximation to \( Q \). We assume that both queues \( Q \) and \( \tilde{Q} \) have the same Poisson input flows, and the service times in \( Q \) have an unknown distribution function \( F \). Moreover, given \( n \geq 1 \) we admit that the distribution of services times in the queue \( \tilde{Q} \) is given by the empirical distribution function \( \hat{F}_n \).

Let \( W \) and \( \tilde{W}_n \) stand for the stationary waiting times in the queues \( Q \) and \( \tilde{Q} \), respectively. Also, let \( d_{TV} \) denote the total variation distance between distributions of random variables.

Under certain conditions, Theorem 1 in [3] provides the following inequality, that could be understood as “the stability inequality”:

\[
d_{TV}(W, \tilde{W}_n) \leq c \cdot \max \left\{ \frac{1}{2} \kappa_1(F, \hat{F}_n), \frac{1}{6} \kappa_3(F, \hat{F}_n) \right\}.
\]  

(17)

If we assume that for some \( \alpha > 0 \), \( \int_0^\infty x^{6+\alpha}dF(x) < \infty \), then we obtain by (9), (17) that \( Ed_{TV}(W, \tilde{W}_n) \) is not greater than a constant times \( \frac{1}{\sqrt{n}} \).

Particularly, this allows to evaluate (in terms of the values of \( n \)) the accuracy of the approximation of the distribution of \( W \) by means of the distribution of \( \tilde{W}_n \).
References


