Reports of Enlarged Session of the
Seminar of I. Vekua Institute
of Applied Mathematics
Volume 29, 2015

ON THE STATISTICAL STRUCTURE DEFINED BY THE LAW OF THE ITERATED LOGARITHM

Kiria T.

Abstract. It is shown that an estimator of an unknown average quadratic deviation defined by the law of the iterated logarithm is not defined for "almost every" infinite sample in R^N. It is constructed a certain modification of this estimator which employs the strong objectivity property.

Keywords and phrases: A strong objective infinite-sample well-founded estimate, shy set, prevalence, Haar ambivalent, the law of the iterated logarithm.

AMS subject classification: 91B02, 91B06, 91B70.

1. Introduction. The notion of a Haar ambivalent introduced by Balka, Buczolich and Elekes [1], has been used in [2] in studying the properties of some infinite sample statistics and in explaining why the Null Hypothesis is sometimes rejected for "almost every" infinite sample by some Hypothesis Testing of maximal reliability. To confirm that the conjectures of Jum Nunnally [3] and Jacob Cohen [4] fail for infinite samples, examples of the so called *objective* and *strong objective* infinite sample well-founded estimates of a useful signal in the linear one-dimensional stochastic model were constructed in [2] by using the axiom of choice and a certain partition of the non-locally compact abelian Polish group $\mathbf{R}^{\mathbf{N}}$ constructed in [5]. More late, in [6] has been demonstrated that an estimate constructed in [7] also is a nice counterexample to Nunnally-Cohen conjectures.

In the present paper we focus on the statistical structure defined by the law of the iterated logarithm and study its properties. More precisely, we show that basic statistic is not defined for "almost every" infinite sample in R^N, but admits a strong objective modification.

The paper is organised as follows. In Section 2 some auxiliary notions and facts from the functional analysis and measure theory are considered. In Section 3 the main result is presented.

2. Auxiliary notions and facts from the functional analysis and measure theory

Lemma 2.1. *Let $\mathbf{R}^{\mathbf{N}}$ be a Polish topological vector space of all real valued sequences equipped with Tychonoff metric. For $J \subseteq \mathbf{N}$, we put*

$$A_J = \{(x_i)_{i \in \mathbf{N}} : x_i \geq 0 \text{ for } i \in J \ \&\ x_i < 0 \text{ for } i \in \mathbf{N} \setminus \mathbf{J}\}$$

and $\Phi = \{A_J : J \subseteq \mathbf{N}\}$. Then every element of Φ is Haar ambivalent, $A_{J_1} \cap A_{J_2} = \emptyset$ for all different $J_1, J_2 \subseteq \mathbf{N}$ and Φ is such a partition of $\mathbf{R}^{\mathbf{N}}$ that card$(\Phi) = 2^{\aleph_0}$.

Remark 2.1. The proof of Lemma 2.1 is similar to the proof of Lemma 15.1.3 (see, [8], p. 202). Concerning main notions and facts about shy and Haar ambivalent sets, prevalences and probes in Polish topological vector spaces the reader can consult with [1], [9], [10].

Definition 2.1. The family of probability spaces $(\mathbf{R}^\mathbf{N}, \mathcal{B}(\mathbf{R}^\mathbf{N}), \mu_\theta)_{\theta \in \Theta}$, where $\mathbf{R}^\mathbf{N}$ is the vector space of all real valued sequences, $\mathcal{B}(\mathbf{R}^\mathbf{N})$ is the σ-algebra of all Borel subsets of $\mathbf{R}^\mathbf{N}$ generated by product topology, $(\mu_\theta^\mathbf{N})_{\theta \in \Theta}$ is the family of Borel probability measures in $\mathbf{R}^\mathbf{N}$ and Θ is a non-empty set, is called a statistical structure.

Definition 2.2. Let $\boldsymbol{\Theta}$ be a non-empty set and let $\mathcal{S}(\Theta)$ be a minimal σ-algebra of subsets of $\boldsymbol{\Theta}$ generated by singletons of Θ. A $(B(\mathbf{R}^\mathbf{N}), \mathcal{S}(\boldsymbol{\Theta}))$ - measurable function $\mathbf{T} : \mathbf{R}^\mathbf{N} \to \boldsymbol{\Theta}$ is called an infinite sample well-founded estimate of a parameter θ for the statistical structure $(\mathbf{R}^\mathbf{N}, \mathcal{B}(\mathbf{R}^\mathbf{N}), \mu_\theta)_{\theta \in \boldsymbol{\Theta}}$, if the condition

(i)$\mu_\theta^\mathbf{N}(\{(x_k)_{k \in \mathbf{N}} : \ (x_k)_{k \in \mathbf{N}} \in \mathbf{R}^\mathbf{N} \ \& \ \mathbf{T}((\mathbf{x_k})_{\mathbf{k} \in \mathbf{N}}) = \theta\}) = \mathbf{1}$

holds for each $\theta \in \Theta$.

In addition, if the following two conditions

(ii) $(\forall \theta)(\theta \in \boldsymbol{\Theta} \to \mathbf{T}^{-1}(\theta)$ is Haar ambivalent);

(iii) $(\forall \theta_1, \theta_2)(\theta_1, \theta_2 \in \boldsymbol{\Theta} \ \to$ there exists an isometric (with respect to Tychonov metric) transformation $A_{(\theta_1, \theta_2)}$ of $\mathbf{R}^\mathbf{N}$ such that $\mathbf{A}_{(\theta_1, \theta_2)}(\mathbf{T}^{-1}(\theta_1)) \boldsymbol{\Delta} \mathbf{T}^{-1}(\theta_2)$ is shy)

hold true, then $\mathbf{T}$ is called strong objective infinite sample well-founded estimate of a parameter θ for the statistical structure $(\mathbf{R}^\mathbf{N}, \mathcal{B}(\mathbf{R}^\mathbf{N}), \mu_\theta)_{\theta \in \boldsymbol{\Theta}}$.

3. On a statistical structure defined by the Law of the iterated logarithm

Lemma 3.1. ([11], Theorem 1, p. 385) *Let* $\mathbf{P}_\theta$ *be a probability measure in* $\mathbf{R}$ *defined by the random variable with mean zero and average quadratic deviation* $\theta (\theta > 0)$. *Then for each* $\theta > 0$ *we have*

$$\mathbf{P}_\theta^\mathbf{N}(\{(\mathbf{x_k})_{\mathbf{k} \in \mathbf{N}} : (\mathbf{x_k})_{\mathbf{k} \in \mathbf{N}} \in \mathbf{R}^\mathbf{N} \ \& \ \limsup_{\mathbf{n} \to \infty} \frac{\sum_{\mathbf{k}=\mathbf{1}}^\mathbf{n} \mathbf{x_k}}{\sqrt{\mathbf{2n} \log \log \mathbf{n}}} = \theta\} = \mathbf{1}.$$

Remark 3.1. Note that under condition of Lemma 3.1, the family of probability measures $(\mathbf{R}^\mathbf{N}, \mathcal{B}(\mathbf{R}^\mathbf{N}), \mathbf{P}_\theta)_{\theta > \mathbf{0}}$ stands for the statistical structure defined by the law of iterated logarithm. The result of Lemma 3.2 asserts that the infinite sample statistic $\mathbf{T} : \mathbf{R}^\mathbf{N} \to (\mathbf{0}, \infty)$ defined by $\mathbf{T}((\mathbf{x_k})_{\mathbf{k} \in \mathbf{N}}) = \limsup_{\mathbf{n} \to \infty} \sum_{\mathbf{k}=\mathbf{1}}^\mathbf{n} \mathbf{x_k} / \sqrt{\mathbf{2n} \log \log \mathbf{n}}$ for $(x_k)_{k \in N} \in \mathbf{R}^\mathbf{N}$ is an infinite sample consistent estimate of the unknown parameter θ for the statistical structure $(\mathbf{R}^\mathbf{N}, \mathcal{B}(\mathbf{R}^\mathbf{N}), \mathbf{P}_\theta)_{\theta > \mathbf{0}}$.

Lemma 3.2. *A set* $\mathbf{S}$*, defined by*

$$\mathbf{S} = \{(\mathbf{x_k})_{\mathbf{k} \in \mathbf{N}} : (\mathbf{x_k})_{\mathbf{k} \in \mathbf{N}} \in \mathbf{R}^\mathbf{N} \ \& \ \limsup_{\mathbf{n} \to \infty} \frac{|\sum_{\mathbf{k}=\mathbf{1}}^\mathbf{n} \mathbf{x_k}|}{\sqrt{\mathbf{2n} \log \log \mathbf{n}}} \ \textit{exists and is finite}\}$$

is a Borel shy set in $\mathbf{R}^\mathbf{N}$.

Proof. It is obvious that $\mathbf{S}$ is a vector subspace of $\mathbf{R}^\mathbf{N}$. We have to show that $\mathbf{S}$ is a Borel subset of $\mathbf{R}^\mathbf{N}$. For $i \in \mathbf{N}$, we denote by Pr_i i-th projection on $\mathbf{R}^\mathbf{N}$ defined by$Pr_i((x_k)_{k \in \mathbf{N}}) = x_i$ for $(x_k)_{k \in \mathbf{N}} \in \mathbf{R}^\mathbf{N}$. We put $\mathbf{T_n} = |\sum_{\mathbf{i}=\mathbf{1}}^\mathbf{n} \mathbf{Pr_i}| / \sqrt{\mathbf{2n} \log \log \mathbf{n}}$ for $n \in \mathbf{N}$. We get

$$\{(x_k)_{k \in N} : (x_k)_{k \in N} \in \mathbf{R}^\mathbf{N} \ \& \ \limsup_{\mathbf{n} \to \infty} \frac{|\sum_{\mathbf{k}=\mathbf{1}}^\mathbf{n} \mathbf{x_k}|}{\sqrt{\mathbf{2n} \log \log \mathbf{n}}} \ \text{exists and is finite}\}$$

$$= \{(x_k)_{k\in N} : (x_k)_{k\in N} \in \mathbf{R^N} \ \& \ \limsup_{\mathbf{n}\to\infty} \frac{|\sum_{\mathbf{k=1}}^{\mathbf{n}} \mathbf{x_k}|}{\sqrt{\mathbf{2n}\log\log\mathbf{n}}} < \infty\}$$

$$= \cup_{s=1}^{\infty}\{(x_k)_{k\in N} : (x_k)_{k\in N} \in \mathbf{R^N} \ \& \ \limsup_{\mathbf{n}\to\infty} \frac{|\sum_{\mathbf{k=1}}^{\mathbf{n}} \mathbf{x_k}|}{\sqrt{\mathbf{2n}\log\log\mathbf{n}}} < \mathbf{s}\}.$$

Since $\overline{\lim}\mathbf{T_n} = \inf_{\mathbf{n}} \sup_{\mathbf{m}\geq\mathbf{n}} \mathbf{T_n}$ is a Borel measurable function in $\mathbf{R^N}$ we claim that $\mathbf{S}$ is a Borel subset in $\mathbf{R^N}$.

We put $v = (v_n)_{n\in N}$ where $v_n = 0$ for $1 \leq n \leq 10$, $v_{11} = 11\sqrt{22\log\log 11)}$ and $v_n = (n+1)\sqrt{2(n+1)\log\log(n+1)} - n\sqrt{2n\log\log n)}$ for $n > 11$. Let us show that v spans a line L such that every translate of L meets S in at most one point; in particular, L is a probe for the complement of S. Indeed, assume the contrary. Then there will be an element $(z_k)_{k\in\mathbf{N}} \in \mathbf{R^N}$ and two different parameters $t_1, t_2 \in \mathbf{R}$ such that $(z_k)_{k\in\mathbf{N}} + t_1 v \in S$ and $(z_k)_{k\in\mathbf{N}} + t_2 v \in S$. Since S is a vector space we deduce that $(t_2 - t_1)v \in S$. Using the same argument we claim that $v \in S$ because $t_2 - t_1 \neq 0$, but the latter relation is false. This ends the proof of Lemma 3.2.

Corollary 3.1. For $(x_k)_{k\in N} \in \mathbf{R^N}$ we put $\mathbf{T}((\mathbf{x_k})_{\mathbf{k}\in\mathbf{N}}) = \limsup_{\mathbf{n}\to\infty} \frac{|\sum_{\mathbf{k=1}}^{\mathbf{n}} \mathbf{x_k}|}{\sqrt{\mathbf{2n}\log\log\mathbf{n}}}$ if $\limsup_{n\to\infty} \frac{|\sum_{k=1}^{n} x_k|}{\sqrt{2n\log\log n}}$ exists and is finite. Then for "almost every" infinite sample the function $\mathbf{T}$ is not defined.

Proof. Let us denote by A the set of all points $(x_k)_{k\in N} \in R^N$ for which the function $\mathbf{T}$ is not defined. It is obvious that $A = \mathbf{R^N} \setminus \mathbf{S}$, where $\mathbf{S}$ comes from Lemma 3.2. By Lemma 3.2 we know that $\mathbf{S}$ is a Borel shy set. Hence $A = \mathbf{R^N} \setminus \mathbf{S}$ is prevalence and Corollary 3.1 is proved.

Theorem 3.1. *Let μ_θ be a Borel probability measure in $\mathbf{R}$ defined by the distribution function of the random variable Y with means zero and θ^2 variance. For $(x_k)_{k\in N} \in \mathbf{R^N}$ we put $\mathbf{T_1}((x_k)_{k\in N}) = \limsup_{n\to\infty} \frac{|\sum_{k=1}^{n} x_k|}{\sqrt{2n\log\log n}}$ if $\limsup_{n\to\infty} \frac{|\sum_{k=1}^{n} x_k|}{\sqrt{2n\log\log n}}$ exists and is finite, and $\mathbf{T_1}((x_k)_{k\in N}) = 1$, otherwise. Then $\mathbf{T_1}$ is a subjective infinite sample consistent estimate of the parameter $\theta \in (0, \infty)$.*

Proof. According to Lemma 3.1, we have

$$\mu_\theta^{\mathbf{N}}(\{(x_k)_{k\in\mathbf{N}} : (x_k)_{k\in\mathbf{N}} \in \mathbf{R^N} \ \& \ \mathbf{T_1}((x_k)_{k\in\mathbf{N}}) = \theta\}) = 1$$

for each $\theta > 0$.

Note that $\mathbf{T_1}$ is subjective because the set $\mathbf{T_1^{-1}(1)}$ is a complement of the shy set $\mathbf{S} \setminus \mathbf{S_1}$, where $\mathbf{S}$ comes from Lemma 3.2 and

$$\mathbf{S_1} = \{(x_k)_{k\in\mathbf{N}} : (x_k)_{k\in\mathbf{N}} \in \mathbf{R^N} \ \& \ \limsup_{\mathbf{n}\to\infty} \frac{|\sum_{\mathbf{k=1}}^{\mathbf{n}} \mathbf{x_k}|}{\sqrt{\mathbf{2n}\log\log\mathbf{n}}} = \mathbf{1}\}.$$

This ends the proof of the theorem.

Example 3.1. Let us denote by $\mathcal{P}(\mathbf{N})$ powerset of the set of all natural numbers, and by ϕ a one-to-one mapping of $\mathbf{R}^+$ onto $\mathcal{P}(\mathbf{N})$, where $\mathbf{R}^+ = (\mathbf{0}, +\infty)$. We put

$$S_\theta = \{(x_k)_{k\in N} : (x_k)_{k\in N} \in R^N \ \& \ \limsup_{n\to\infty} \frac{|\sum_{k=1}^{n} x_k|}{\sqrt{2n\log\log n}} = \theta\}$$

for each $\theta \in \mathbf{R}^+$.

Since $S = \cup_{\theta \in \mathbf{R}^+} S_\theta$ and $S_{\theta_1} \cap S_{\theta_2} = \emptyset$ for different $\theta_1, \theta_2 \in \mathbf{R}^+$, by Lemma 3.2 we deduce that S_θ is shy for each $\theta \in \mathbf{R}^+$.

We set $D_\theta = (A_{\phi(\theta)} \setminus S) \cup S_\theta$ for $\theta \in \mathbf{R}^+$, where $A_{\phi(\theta)}$ comes from Lemma 2.1. Then $(D_\theta)_{\theta \in \mathbf{R}}$ will be Borel partition of $\mathbf{R}^{\mathbf{N}}$ such that D_θ is a Haar ambivalent for $\theta \in \mathbf{R}^+$ and for all $\theta_1, \theta_2 \in \mathbf{R}^+$ there exists an isometric (w.r.t. Tychonov metric) transformation $A_{(\theta_1,\theta_2)}$ of the $\mathbf{R}^\infty$ such that $A_{(\theta_1,\theta_2)}(D_{\theta_1}) \Delta D_{\theta_1}$ is shy. We can define $A_{(\theta_1,\theta_2)}$ as follows: for $(x_k)_{k \in \mathbf{N}} \in \mathbf{R}^{\mathbf{N}}$ we put $A_{(\theta_1,\theta_2)}((x_k)_{k \in \mathbf{N}}) = (y_k)_{k \in \mathbf{N}}$, where $y_k = x_k$ if $k \in \phi(\theta_1) \cap \phi(\theta_2)$ and $y_k = -x_k$ otherwise.

We put $T^\circ((x_k)_{k \in \mathbf{N}}) = \theta$ if $(x_k)_{k \in \mathbf{N}} \in D_\theta$.

Now it is not hard to show that under conditions of Lemma 3.1, T° is a strong objective infinite sample well-founded estimate of a parameter θ for the family $(\mu_\theta^{\mathbf{N}})_{\theta \in \mathbf{R}^+}$.

R E F E R E N C E S

1. Balka R., Buczolich Z., Elekes M. Topological Hausdorff dimension and level sets of generic continuous functions on fractals. *Chaos Solitons Fractals*, **45**, 12 (2012), 1579-1589.

2. Pantsulaia G., Kintsurashvili M. Why is Null Hypothesis rejected for "almost every" infinite sample by some Hypothesis Testing of maximal reliability?. *Journal of Statistics: Advances in Theory and Applications*, **11**, 1 (2014), 45-70.

3. Nunnally J. The place of statistics in psychology, *Educational and Psychological Measurement*, **20**, 4 (1960), 641-650.

4. Cohen J. The Earth is round ($p < .05$). *American Psychologist*, **49**, 12 (1994), 997-1003.

5. Pantsulaia G.R. On a certain partition of the non-locally compact abelian Polish group $\mathbb{R}^{\mathbb{N}}$. *Proc. A. Razmadze Math. Inst.*, **149** (2009), 75-86.

6. Pantsulaia G., Kintsurashvili M. An objective infinite-sample well-founded estimate of a useful signal in one-dimensional linear stochastic model. *Rep. Enlarged Sess. Semin. I. Vekua Appl. Math.*, **28** (2014), 90-93.

7. Zerakidze Z., Pantsulaia G., Saatashvili G. On the separation problem for a family of Borel and Baire G-powers of shift-measures on $\mathbf{R}$. *Ukrainian Math. J.*, **65**, 4 (2013), 470-485.

8. Pantsulaia G.R. Invariant and Quasiinvariant Measures in Infinite-Dimensional Topological Vector Spaces. *Nova Science Publishers, Inc., New York*, 2007.

9. Christensen J.R. Topology and Borel Structure. *Amsterdam: North-Holland Publishing Company*, 1974.

10. Hunt B.R., Sauer T., Yorke J.A. Prevalence: A translation-invariant "almost every" on infinite-dimensional spaces.*Bulletin (New Series) of the American Mathematical Society*, **27(2)**, 10 (1992), 217-238.

11. Shiryaev A.N. Probability (Russian). *Izd. Nauka, Moscow*, 1980.

Received 15.05.2015; revised 15.09.2015; accepted 12.11.2015.

Author's address:

T. Kiria
Department of Mathematics
Georgian Technical University
77, Kostava St., Tbilisi 0175
Georgia
E-mail: kiria8@gmail.com