Reports of Enlarged Session of the
Seminar of I. Vekua Institute
of Applied Mathematics
Volume 28, 2014

AN OBJECTIVE INFINITE SAMPLE WELL-FOUNDED ESTIMATE OF A USEFUL SIGNAL IN THE LINEAR ONE-DIMENSIONAL STOCHASTIC MODEL

Pantsulaia G., Kintsurashvili M.

Abstract. It is shown that $\overline{\lim}\widetilde{T_n} := \inf_n \sup_{m \geq n} \widetilde{T_m}$ and $\underline{\lim}\widetilde{T_n} := \sup_n \inf_{m \geq n} \widetilde{T_m}$ are objective infinite sample well-founded estimates of a useful signal θ in the linear one-dimensional stochastic model $\xi_k = \theta + \Delta_k$ $(k \in \mathbb{N})$, where $\#(\cdot)$ denotes a counting measure, Δ_k is a sequence of independent identically distributed random variables on $\mathbb{R}$ with strictly increasing continuous distribution function F, expectation of Δ_1 does not exist and $T_n : \mathbb{R}^{\mathbb{N}} \to \mathbb{R}$ $(n \in \mathbb{N})$ is defined by $T_n((x_k)_{k\in\mathbb{N}}) = -F^{-1}(n^{-1}\#(\{x_1, \cdots, x_n\} \cap (-\infty; 0]))$ for $(x_k)_{k\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$.

Keywords and phrases: An objective infinite-sample well-founded estimate, shy set, prevalence, Haar ambivalent.

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

1. Introduction. In [1], concepts of subjective and objective infinite sample well-founded estimates of a useful signal in the linear one-dimensional stochastic model were introduced by using the notion of a Haar null set introduced by J.P.R. Christensen [2].

In [3], a separation problem for the family of Borel and Baire G-powers of shift-measures on $\mathbb{R}$ for an arbitrary infinite additive group G was studied. It was proved that $T_n : \mathbb{R}^n \to \mathbb{R}$ $(n \in \mathbb{N})$ defined by $T_n(x_1, \cdots, x_n) = -F^{-1}(n^{-1}\#(\{x_1, \cdots, x_n\} \cap (-\infty; 0]))$ for $(x_1, \cdots, x_n) \in \mathbb{R}^n$, was a well-founded estimator of a useful signal θ in the linear one-dimensional stochastic model $\xi_k = \theta + \Delta_k$ $(k \in \mathbb{N})$, where $\#(\cdot)$ denotes a counting measure, Δ_k is a sequence of independent identically distributed random variables on $\mathbb{R}$ with strictly increasing continuous distribution function F and expectation of Δ_1 does not exist.

The purpose of the present manuscript is to show that $\overline{\lim}\widetilde{T_n} := \inf_n \sup_{m\geq n} \widetilde{T_m}$ and $\underline{\lim}\widetilde{T_n} := \sup_n \inf_{m\geq n} \widetilde{T_m}$ are objective infinite sample well-founded estimates of a useful signal θ in the same model.

2. Auxiliary notions and facts from functional analysis and measure theory. Let $\mathbb{V}$ be a complete metric linear space, by which we mean a vector space (real or complex) with a complete metric for which the operations of addition and scalar multiplication are continuous. When we speak of a measure on $\mathbb{V}$ we will always mean a nonnegative measure that is defined on the Borel sets of $\mathbb{V}$ and is not identically zero. We write $S + v$ for the translation of a set $S \subseteq \mathbb{V}$ by a vector $v \in \mathbb{V}$.

Definition 1 ([4], Definition 1, p. 221). A measure μ is said to be transverse to a Borel set $S \subset \mathbb{V}$ if the following two conditions hold:

(i) There exists a compact set $U \subset \mathbb{V}$ for which $0 < \mu(U) < 1$;

(ii) $\mu(S + v) = 0$ for every $v \in \mathbb{V}$.

Definition 2 ([4], Definition 2, p. 222; [5], p. 1579). A Borel set $S \subset \mathbb{V}$ is called shy if there exists a measure transverse to S. More generally, a subset of $\mathbb{V}$ is called shy if it is contained in a shy Borel set. The complement of a shy set is called a prevalent set. We say that a set is a Haar ambivalent set if it is neither shy nor prevalent.

Definition 3 ([4], p. 226). We say "almost every" element of $\mathbb{V}$ satisfies some given property, if the subset of $\mathbb{V}$ on which this property holds is prevalent.

Let $\mathbb{R}^{\mathbb{N}}$ be a topological vector space of all real valued sequences $\mathbb{R}^{\infty}$ equipped with Tychonoff metric ρ defined by $\rho((x_k)_{k\in\mathbb{N}},(y_k)_{k\in\mathbb{N}}) = \sum_{k\in\mathbb{N}} |x_k - y_k|/2^k(1+|x_k-y_k|)$ for $(x_k)_{k\in\mathbb{N}},(y_k)_{k\in\mathbb{N}} \in \mathbb{R}^{\infty}$.

Lemma 1 ([6], Lemma 5, p.78). *Let J be an arbitrary subset of $\mathbb{N}$. We set*

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

Then the family of subsets $\Phi = \{A_J : J \subseteq \mathbb{N}\}$ has the following properties:

(*i*) *every element of Φ is a Haar ambivalent set.*

(*ii*) $A_{J_1} \cap A_{J_2} = \emptyset$ *for all different $J_1, J_2 \subseteq \mathbb{N}$.*

(*iii*) Φ *is a partition of $\mathbb{R}^{\mathbb{N}}$ such that $card(\Phi) = 2^{\aleph_0}$.*

Suppose that Θ is a subset of the infinite-dimensional topological vector space $\mathbb{R}^{\mathbb{N}}$.

In the information transmission theory we consider the linear one-dimensional stochastic system

$$(\xi_k)_{k\in\mathbb{N}} = (\theta_k)_{k\in\mathbb{N}} + (\Delta_k)_{k\in\mathbb{N}}, \tag{1}$$

where $(\theta_k)_{k\in\mathbb{N}} \in \Theta$ is a sequence of useful signals, $(\Delta_k)_{k\in\mathbb{N}}$ is a sequence of independent identically distributed random variables (the so-called generalized "white noise") defined on some probability space $(\Omega, \mathcal{F}, P)$ and $(\xi_k)_{k\in\mathbb{N}}$ is a sequence of transformed signals. Let μ be a Borel probability measure on $\mathbb{R}$ defined by a random variable Δ_1. Then the $\mathbb{N}$-power of the measure μ denoted by $\mu^{\mathbb{N}}$ coincides with the Borel probability measure on $\mathbb{R}^{\mathbb{N}}$ defined by the generalized "white noise", i.e.,

$$(\forall X)(X \in \mathcal{B}(\mathbb{R}^{\mathbb{N}}) \to \mu^{\mathbb{N}}(X) = P(\{\omega : \omega \in \Omega \ \& \ (\Delta_k(\omega))_{k\in\mathbb{N}} \in X\})),$$

where $\mathcal{B}(\mathbb{R}^{\mathbb{N}})$ is the Borel σ-algebra of subsets of $\mathbb{R}^{\mathbb{N}}$.

Following [7], a general decision in the information transmission theory is that the Borel probability measure λ, defined by the sequence of transformed signals $(\xi_k)_{k\in\mathbb{N}}$ coincides with $\left(\mu^{\mathbb{N}}\right)_{\theta_0}$ for some $\theta_0 \in \Theta$ provided that

$$(\exists\theta_0)(\theta_0 \in \Theta \to (\forall X)(X \in \mathcal{B}(\mathbb{R}^{\mathbb{N}}) \to \lambda(X) = \left(\mu^{\mathbb{N}}\right)_{\theta_0}(X))),$$

where $\left(\mu^{\mathbb{N}}\right)_{\theta_0}(X) = \mu^N(X - \theta_0)$ for $X \in \mathcal{B}(\mathbb{R}^{\mathbb{N}})$.

Here we consider a particular case of the above model (1) when a set of useful signals Θ has the form

$$\Theta = \{(\theta, \theta, \cdots) : \theta \in \mathbb{R}\}.$$

For $\theta \in \mathbb{R}$, a measure $\mu_\theta^{\mathbb{N}}$ defined by

$$\mu_\theta^N = \mu_\theta \times \mu_\theta \times \cdots,$$

where μ_θ is a θ-shift of μ (i.e., $\mu_\theta(X) = \mu(X - \theta)$ for $X \in \mathcal{B}(\mathbb{R})$), is called the $\mathbb{N}$-power of the θ-shift of μ on $\mathbb{R}$. It is obvious that $\mu_\theta^{\mathbb{N}} = \left(\mu^{\mathbb{N}}\right)_{(\theta,\theta,\cdots)}$.

Following concepts of the theory of statistical decisions, a triplet $(\mathbb{R}^{\mathbb{N}}, \mathcal{B}(\mathbb{R}^{\mathbb{N}}), \mu_\theta^{\mathbb{N}})_{\theta\in\Theta}$ with $\Theta \subseteq \mathbb{R}$ is called a statistical structure describing the linear one-dimensional stochastic model (1).

Definition 4. We set $\overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$. Let $\mathcal{S}(\overline{\mathbb{R}})$ denote a minimal σ-algebra of subsets of $\overline{\mathbb{R}}$ generated by singletons of $\overline{\mathbb{R}}$. A $(\mathcal{B}(\mathbb{R}^{\mathbb{N}}), \mathcal{S}(\overline{\mathbb{R}}))$-measurable function $T : \mathbb{R}^{\mathbb{N}} \to \overline{\mathbb{R}}$ is called an infinite sample well-founded estimate of a parameter θ for the family $(\mu_\theta^{\mathbb{N}})_{\theta\in\Theta}$ if the condition

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

holds for each $\theta \in \Theta := \mathbb{R}$.

Lemma 2 ([3], Theorem 4.2, p. 483). *Let F be a strictly increasing continuous distribution function on R and let μ be a Borel probability measure on R defined by F. For $\theta \in R$, we set $F_\theta(x) = F(x-\theta)(x \in R)$ and denote by μ_θ a Borel probability measure on $\mathbb{R}$ defined by F_θ. Then estimators $\overline{\lim}\widetilde{T_n} := \inf_n \sup_{m\geq n} \widetilde{T_m}$ and $\underline{\lim}\widetilde{T_n} := \sup_n \inf_{m\geq n} \widetilde{T_m}$ are infinite sample consistent estimators of a parameter θ for the family $(\mu_\theta^{\mathbb{N}})_{\theta\in\mathbb{R}}$, where $\widetilde{T_n} : \mathbb{R}^{\mathbb{N}} \to \mathbb{R}$ is defined by*

$$(\forall (x_k)_{k\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}} \to \widetilde{T_n}((x_k)_{k\in\mathbb{N}}) = -F^{-1}(n^{-1}\#(\{x_1, \cdots, x_n\} \cap (-\infty; 0]))).$$

Below we give a certain modification of Definition 5.1 introduced in [1] (see, p. 66).

Definition 5. An infinite sample well-founded estimate $T : \mathbb{R}^{\mathbb{N}} \to \overline{\mathbb{R}}$ of a parameter $\theta \in \mathbb{R}$ is called subjective if there is a null hypothesis $H_0 : \theta = \theta_0$ $(\theta_0 \in \mathbb{R})$ which is accepted or rejected for "almost every" infinite sample by T or a region of the rejection of the Main Assumption *that the Borel probability measure λ defined by the sequence of transformed signals $(\xi_k)_{k\in\mathbb{N}}$ coincides with $\left(\mu^{\mathbb{N}}\right)_{\theta_0}$ for some $\theta_0 \in \mathbb{R}$* is shy or prevalent. Otherwise, the estimate T is called objective.

3. Main result. The following statement takes place.

Theorem. *Under conditions of Lemma 2, $\overline{\lim}\widetilde{T_n} := \inf_n \sup_{m\geq n} \widetilde{T_m}$ and $\underline{\lim}\widetilde{T_n} := \sup_n \inf_{m\geq n} \widetilde{T_m}$ are objective infinite sample well-founded estimates.*

Proof. We consider the proof of the theorem only for $\overline{\lim}\widetilde{T_n}$. The proof of the theorem for $\underline{\lim}\widetilde{T_n}$ can be obtained similarly.

By Lemma 2 we know that $\overline{\lim}\widetilde{T_n} := \inf_n \sup_{m\geq n} \widetilde{T_m}$ is an infinite sample well-founded estimate of a parameter $\theta \in \mathbb{R}$. We have to show that $(\overline{\lim}\widetilde{T_n})^{-1}(\theta)$ is a Haar ambivalent set for each $\theta \in \overline{\mathbb{R}}$. Let $\theta \in \overline{\mathbb{R}}$ and $(x_k)_{k\in\mathbb{N}} \in (\overline{\lim}\widetilde{T_n})^{-1}(\theta)$. This means that

$$\inf_n \sup_{m\geq n} \big(-F^{-1}(m^{-1}\#(\{x_1, \cdots, x_m\} \cap (-\infty; 0]))\big) = \theta.$$

Setting $J = \{k : x_k \leq 0\}$, let consider a set

$$A_J = \{(y_k)_{k\in\mathbb{N}}\ :\ y_k \leq 0 \text{ for } k \in J \ \&\ y_k > 0 \text{ for } k \in \mathbb{N} \setminus \mathbb{J}\}.$$

Then

$$A_J \subseteq (\overline{\lim}\widetilde{T_n})^{-1}(\theta).$$

By Lemma 1 we know that A_J is a Haar ambivalent set which implies that $(\overline{\lim}\widetilde{T_n})^{-1}(\theta)$ is not shy. Since for $\theta_1 \neq \theta$ we have $(\overline{\lim}\widetilde{T_n})^{-1}(\theta) \cap (\overline{\lim}\widetilde{T_n})^{-1}(\theta_1) = \emptyset$ we deduce that $(\overline{\lim}\widetilde{T_n})^{-1}(\theta)$ is a Haar ambivalent set.

Notice that the region of the rejection of all null hypothesis $H_0 : \theta = \theta_0$ $(\theta_0 \in \mathbb{R})$ coincides with a set $(\overline{\lim}\widetilde{T_n})^{-1}(-\infty) \cup (\overline{\lim}\widetilde{T_n})^{-1}(+\infty)$ which is also a Haar ambivalent set.

This ends the proof of the theorem.

R E F E R E N C E S

1. 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.
2. Christensen J.R. Topology and Borel Structure. *Amsterdam: North-Holland Publishing Company*, 1974.
3. Zerakidze Z., Pantsulaia G., Saatashvili G. On the separation problem for a family of Borel and Baire G-powers of shift-measures on $\mathbb{R}$. *Ukrainian Math. J.,* **65**, 4 (2013), 470-485.
4. 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.
5. 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.
6. Pantsulaia G.R. Invariant and Quasiinvariant Measures in Infinite-Dimensional Topological Vector Spaces. *Nova Science Publishers, Inc., New York*, 2007.
7. Ibramkhallilov I.Sh., Skorokhod A.V. On Well-off Estimates of Parameters of Stochastic Processes. (Russian), *Kiev*, 1980.

Received 22.05.2014; revised 10.10.2014; accepted 25.12.2014.

Authors' addresses:

G. Pantsulaia
I. Vekua Institute of Applied Mathematics of
Iv. Javakhishvili Tbilisi State University
2, University St., Tbilisi 0186
Georgia
E-mail: g.pantsulaia@gtu.ge

M. Kintsurashvili
Department of Mathematics
Georgian Technical University
77, Kostava St., Tbilisi 0175
Georgia
E-mail: m.kintsurashvili@gtu.ge