Bulletin of TICMI
Vol. 25, No. 2, 2021, 85–91

The Consistent Estimators of Charlier's Statistical Structures in Hilbert Space of Measures

Omar Purtukhia[1*], Zurab Zerakidze[2]

[1]*I. Javakhishvili Tbilisi State University, 2 University St., 0186, Tbilisi, Georgia*;
[2]*Gori State University, 53 Chavchavadze Ave., 1400, Gori, Georgia*
(*Received June 25, 2021; Revised September 01, 2021; Accepted September 09, 2021*)

In this paper, we consider the Charlier statistical Structures in a Hilbert space of measures. Sufficient and necessary conditions for the existence of consistent estimators of parameters are given.

Keywords: Consistent estimators, Orthogonal structure, Weakly separable structure, Separable structure, Strongly separable structure.

AMS Subject Classification: 62H05, 62H12.

1. Introduction

Let (E, S) be a measurable space with a given family of probability measures: $\{\mu_i, i \in I\}$.

We recall some definitions from [1] - [6].

Definition 1.1: An object $\{E, S, \mu_i, i \in I\}$ is called a statistical structure.

Definition 1.2: A statistical structure $\{E, S, \mu_i, i \in I\}$ is called orthogonal (singular) if a family of probability measures $\{\mu_i, i \in I\}$ constists of pairwise singular measures (i.e. $\mu_i \perp \mu_j$, $\forall i \neq j$).

Definition 1.3: A statistical structure $\{E, S, \mu_i, i \in I\}$ is called weakly separable if there exists a family of S-measurable sets $\{X_i, i \in I\}$ such that

$$\mu_i(X_j) = \begin{cases} 1, & if\ i = j; \\ 0, & if\ i \neq j \end{cases} \qquad (i, j \in I).$$

Let $\{\mu_i, i \in I\}$ be Charlier probability measures defined on the measurable space (E, S). For each $i \in I$ we denote by $\overline{\mu}_i$ the completion of the measure μ_i, and by $dom(\overline{\mu}_i)$ – the σ-algebra of all μ_i-measurable subsets of E.

We denote

$$S_1 = \cap_{i \in I} dom(\overline{\mu}_i).$$

*Corresponding author. Email: o.purtukhia@gmail.com

ISSN: 1512-0082 print

Definition 1.4: A statistical structure $\{E, S_1, \overline{\mu}_i, i \in I\}$ is called strongly separable if there exists a family of S_1-measurable sets $\{Z_i, i \in I\}$ such that the following relations are fulfilled:
1) $\mu_i(Z_i) = 1 \ \ \forall i \in I$;
2) $Z_{i_1} \cap Z_{i_2} = \emptyset, \ \ \forall i_1 \neq i_2, \ i_1, i_2 \in I$;
3) $\cup_{i \in I} Z_i = E$.

Let I be the set of hypotheses and let $B(I)$ be σ-algebra of subsets of I which contains all finite subsets of I.

Definition 1.5: We will say that the statistical structure $\{E, S, \overline{\mu}_i, i \in I\}$ admits a consistent estimators of parameters $i \in I$ if there exists at least one measurable mapping $\delta \ : \ (E, S) \ \longrightarrow \ (I, B(I))$, such that

$$\overline{\mu}_i(\{x : \delta(x) = i\}) = 1, \ \ \forall \, i \in I.$$

Let M^σ be a real linear space of all alternating finite measures on S.

Definition 1.6: A linear subset $M_H \subset M^\sigma$ is called a Hilbert space of measures if:

1) One can introduce on M_H a scalar product (μ, ν) $(\mu, \nu \in H$ so that M_H is a Hilbert space and for every mutually singular measures μ and ν $(\mu, \nu \in H)$ the scalar product $(\mu, \nu) = 0$;

2) If $\nu \in M_H$ and $|f(x)| \leq 1$, then

$$\nu_f(A) = \int_A f(x)\nu(dx) \in M_H,$$

where f is a S_1-measurable real function and $(\nu_f, \nu_f) \leq (\nu, \nu)$;

3) If $\nu_n \in M_H, \nu_n \geq 0, \nu_n(E) < \infty, n = 1, 2, ...$ and $\nu_n \downarrow 0$, then for any $\mu \in M_H$:

$$\lim_{n \to \infty} (\nu_n, \mu) = 0.$$

2. The consistent estimators of Charlier's Statistical Structure in Hilbert space of measures

The normal distribution is symmetrical, that is, the normal distribution density function

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}}$$

is symmetric with respect to the line $x = m$. However, in practice, asymmetric distributions are also often encountered.In the case when the asymmetry in absolute value is not very large, the density can be expressed using the so-called Charlier's law.

The density of Charlier's law is determined by the equality

$$f_{Ch}(x) = f(x) + \frac{1}{\sigma}\Big[\frac{S_k(x)}{6} \cdot z_u \cdot (u^3 - 3u) + \frac{E_k(x)}{24} \cdot z_u \cdot (u^4 - 6u^2 + 3)\Big], \qquad (1)$$

where $f(x)$ is the density of the normal distribution, $u = \frac{x-m}{\sigma}$, $z_u = \frac{1}{\sqrt{2\pi}} e^{-u^2/2}$, $S_k(x) = \mu_3/\sigma^3$ – asymmetry, and $E_k(x) = \mu_4/\sigma^4 - 3$ – kurtosis.

Thus, the second term on the right-hand side of (1) is a correction to the normal distribution. Obviously, for $S_k(x) = 0$ and $E_k(x) = 0$, the Charlier distribution coincides with the normal distribution.

Let μ be the probability measure given on $(R, L(R))$ by the formula

$$\mu(A) = \int_A f_{Ch}(x)dx, \quad A \in L(R).$$

The probability measure determined in this way will be called the Charlier measure.

Definition 2.1: A statistical structure $\{E, S, \mu_i, i \in I\}$ is called a statistical Charlier structure if μ_i, $\forall i \in I$ are Charlier measures.

Let $\{E, S, \overline{\mu}_i, i \in I\}$ be an orthogonal Charlier statistical structure. Next, consider S_1-measurable functions $g_i(x)$, $i \in I$, such that:

$$\sum_{i \in I} \int_E |g_i(x)|^2 \overline{\mu}_i(dx) < +\infty.$$

Consider a measure ν of type

$$\nu_k(B) = \sum_{i \in I_k} \int_B g_i^{(k)}(x) \overline{\mu}_i(dx), \quad B \in S_1, \quad k = 1, 2$$

and define the scalar product on M_H by the formula

$$(\nu_1, \nu_2) = \sum_{i \in I_1 \cap I_2} \int_E g_i^{(1)}(x) g_i^{(2)}(x) \overline{\mu}_i(dx),$$

where I_1 and I_2 are countable subsets of I.

Thus, we can assert that M_H is a Hilbert space of measures and, moreover, M_H is direct sum of Hilbert spaces $H_2(\overline{\mu}_i)$:

$$M_H = \oplus_{i \in I} H_2(\overline{\mu}_i),$$

where $H_2(\overline{\mu}_i)$ are the set of measures of the type

$$\nu(B) = \int_B f(x) \overline{\mu}_i(dx), \quad B \in S_1,$$

with an integrand satisfying the condition

$$\int_E |f(x)|^2 \overline{\mu}_i(dx) < +\infty, \quad i \in I.$$

Let

$$M_H = \oplus_{i \in I} H_2(\overline{\mu}_i)$$

be a Hilbert space of measures, where $cardI \leq c$. Let E be a complete separable metric space and let S_1 be a Borel σ-algebra on E. Let $F = F(M_H)$ be a set of real functions for which $\int_E f(x)\overline{\mu}_i(dx)$ is defined $\forall \overline{\mu}_i \in M_H$. Then we have.

Theorem 2.2: *In order for the Charlier Statistical Structure $\{E, S_1, \overline{\mu}_i, i \in I\}$ to admit a consistent estimators of parameter $i \in I$ in the theory of (ZFC) & (MA) it is necessary and sufficient that the correspondence $f \longleftrightarrow \psi_f$ $(f \in F(M_H))$ defined by the equality*

$$\int_E f(x)\overline{\mu}_i(dx) = (\psi_f, \overline{\mu}_i) \quad \forall \overline{\mu}_i \in M_H$$

was one-to-one.

Proof: Necessity. The existence of a consistent estimators of parameter $\delta : (E, S_1) \longrightarrow (I, B(I))$ implies that $\overline{\mu}_i(\{x : \delta(x) = i\}) = 1 \quad \forall i \in I$. Setting $X_i = \{x : \delta(x) = i\}$ for $i \in I$ we get:

1) $\overline{\mu}_i(X_i) = \overline{\mu}_i(\{x : \delta(x) = i\}) = 1 \ \ \forall i \in I$;
2) $X_{i_1} \cap X_{i_2} = \emptyset$ for all different parameters i_1 and i_2 from I;
3) $\cup_{i \in I} X_i = \{x : \delta(x) \in I\} = E$.

Hence, the statistical structure $\{E, S_1, \overline{\mu}_i, i \in I\}$ is strongly separable. Therefore there exist S_1-measurable sets X_i $(i \in I)$, such that

$$\overline{\mu}_i(X_{i'}) = \begin{cases} 1, & if \ i = i'; \\ 0, & if \ i \neq i'. \end{cases}$$

Let the function $I_{X_i}(x) \in E$ correspond to the measure $\overline{\mu}_i \in H_2(\overline{\mu}_i)$. Then

$$\int\limits_E I_{X_i}(x)\overline{\mu}_i(dx) = \int\limits_E I_{X_i}(x) I_{X_i}(x)\overline{\mu}_i(dx) = (\overline{\mu}_i, \overline{\mu}_i).$$

If now we associate the measure $\overline{\mu}_{i_1} \in H_2(\overline{\mu}_i)$ with the function $f_{i_1}(x) = f_1(x) I_{X_i}(x) \in F(M_H)$ then for all $\overline{\mu}_{i_2} \in M_H(\overline{\mu}_i)$ we can write

$$\int\limits_E f_{i_1}(x) f_{i_2}(x)\overline{\mu}_i(dx) = \int\limits_E f_1(x) f_2(x) I_{X_i}(x) I_{X_i}(x)\overline{\mu}_i(dx)$$

$$= \int\limits_E f_1(x) f_2(x)\overline{\mu}_i(dx) = (\overline{\mu}_{i_1}, \overline{\mu}_{i_2}).$$

Further, we associate the measure

$$\nu(C) = \sum_{i \in I_f \subset I} \int\limits_C g_i(x)\overline{\mu}_i(dx) \in M_H$$

with the function

$$f(x)=\sum_{i\in I_f} g_i(x)I_{X_i}(x)\in F(M_B).$$

Then for the measure

$$\nu_1(C)=\sum_{i\in I_{f_1}\subset I}\int\limits_C g_i^1(x)\overline{\mu}_i(dx)\in M_H$$

we have

$$\int\limits_E f(x)\nu_1(dx)=\int\limits_E\sum_{i\in I_f\cap I_{f_2}} g_i(x)g_i^1(x)\overline{\mu}_i(dx)$$

$$=\sum_{i\in I_f\cap I_{f_2}}\int\limits_E g_i(x)g_i^1(x)\overline{\mu}_i(dx)=(\nu,\nu_1).$$

From the discussion it follows that the above correspondence connects some function $f\in F(M_B)$ into correspondence with some $\nu_f\in M_H$. If in $F(M_B)$ we identify functions that coincide with respect to measures $\{\overline{\mu}_i,\ i\in I\}$, then the correspondence will be bijective.

Sufficiency. Let $f\in F(M_H)$ correspond to the measure $\overline{\mu}_i\in M_H$ for which

$$\int\limits_E f(x)\overline{\mu}_i(dx)=(\overline{\mu}_f,\overline{\mu}_i).$$

Then for every $\overline{\mu}_{i'}\in M_H$ we have

$$\int\limits_E f_i(x)\overline{\mu}_{i'}(dx)=(\overline{\mu}_i,\overline{\mu}_{i'})=\int\limits_E f_1(x)f_2(x)\overline{\mu}_i(dx)=\int\limits_E f_i(x)f_2(x)\overline{\mu}_i(dx).$$

So $f_i(x)=f_1(x)$ almost everywhere with respect to the measure $\overline{\mu}_i$. Suppose that $f_i(x)>0$ and

$$\int_E f_i^2(x)\overline{\mu}_i(dx)<+\infty.$$

If now

$$\mu_i^*(C)=\int_C f_i(x)\overline{\mu}_i(dx)$$

then

$$\int\limits_E f_i^*(x)\overline{\mu}_{i'}(dx)=(\mu_i^*,\overline{\mu}_{i'})=0,\ \ i\neq i^{'},$$

where f_i^* is the function which corresponds to the measure μ_i^*.

On the other hand, $\overline{\mu}_i(E \setminus X_i) = 0$, where $X_i = \{x : f_i^*(x) > 0\}$. Hence, we obtain

$$\overline{\mu}_i(X_{i'}) = \begin{cases} 1, & if\ i = i^{'}; \\ 0, & if\ i \neq i^{'}. \end{cases}$$

Therefore the Charlier Statistical Structure $\{E, S_1, \overline{\mu}_i, i \in I\}$ is weakly separable. Further, we represent $\{\overline{\mu}_i, i \in I\}$, $cardI \leq c$, as an inductive sequence $\{\overline{\mu}_i < \omega_1\}$, where ω_1 denotes the first ordinal number of the power of the set I.

Since the Charlier Statistical Structure $\{E, S_1, \overline{\mu}_i, i \in I\}$ is weakly separable, there exists a family of S_1-measurable sets $X_i, i \in I$ such that for all $i \in [0, \omega_1)$ we have

$$\overline{\mu}_i(X_{i'}) = \begin{cases} 1, & if\ i = i^{'}; \\ 0, & if\ i \neq i^{'}. \end{cases}$$

We define ω_1 sequence Z_i of parts of the space E such that the following relations hold:

1) Z_i is a Borel subset of E $\forall i < \omega_1$;
2) $Z_i \subset X_i \ \ \forall i < \omega_1$;
3) $Z_i \cap Z_{i'} = \emptyset$ for all $i < \omega_1,\ i^{'} < \omega_1,\ i \neq i^{'}$;
4) $\overline{\mu}_i(Z_i) = 1 \ \ \forall i < \omega_1$.

Assume that $Z_{i_0} = X_{i_0}$. Suppose further that the partial sequence $\{Z_{i'}\}_{i'<i}$ is already defined for $i < \omega_1$. It is clear that $\overline{\mu}^*(\cup_{i'>i} Z_{i'}) = 0$. Thus there exists a Borel subset Y_i of the space E such that the following relations are valid:

$$\cup_{i'>i} Z_{i'} \subset Y_i \ \ and\ \ \overline{\mu}_i(Y_i) = 0.$$

Assuming that $Z_i = X_i \setminus Y_i$, we construct the ω_1 sequence $\{Z_i\}_{i<\omega_1}$ of disjunctive measurable subsets of the space E. Therefore $\overline{\mu}_h(Z_i) = 1$ for all $i < \omega_1$ and the Charlier statistical structure $\{E, S_1, \overline{\mu}_i, i \in I\}$, $cardI \leq c$, is strongly separable because there exists a family of elements of the σ-algebra $S_1 = \cap_{i\in I} dom(\overline{\mu}_i)$ such that:

1) $\overline{\mu}_i(Z_i) = 1 \ \ \forall i \in I$;
2) $Z_i \cap Z_{i'} = \emptyset$ for all different i and $i^{'}$ from I;
3) $\cup_{i\in I} Z_i = E$.

For $x \in E$, we put $\delta(x) = i$, where i is the unique parameter from the set I for which $x \in Z_i$. The existence of such a unique parameter from I can be proved using conditions 2), 3).

Now let $Y \in B(I)$. Then $\{x : \delta(x) \in Y\} = \cup_{i\in Y} Z_i$. We must show that $\{x : \delta(x) \in Y\} \in dom(\overline{\mu}_{I_0})$ for $i_0 \in I$.

If $i_0 \in Y$, then

$$\{x : \delta(x) \in Y\} = \cup_{i\in Y} Z_i = Z_{i_0} \cup (\cup_{i\in Y\setminus\{i_0\}} Z_i).$$

On the one hand, from the validity of the condition 1), 2), 3) it follows that

$$Z_{i_0} \in S_1 = \cap_{i\in I} dom(\overline{\mu}_i) \subseteq dom(\overline{\mu}_{i_0}).$$

On the other hand, the validity of the condition

$$\cup_{i \in Y \setminus \{i_0\}} Z_i \subseteq (E \setminus Z_{i_0})$$

implies that

$$\overline{\mu}_{i_0}(\cup_{i \in Y \setminus \{i_0\}} Z_i) = 0.$$

The last equality yields that

$$\cup_{i \in Y \setminus \{i_0\}} Z_i \in dom(\overline{\mu}_{i_0}).$$

Since $dom(\overline{\mu}_{i_0})$ is a σ-algebra, we deduce that

$$\{x : \delta(x) \in Y\} = Z_{i_0} \cup (\cup_{i \in Y \setminus \{i_0\}} Z_i) \in dom(\overline{\mu}_{i_0}).$$

If $i_0 \notin Y$, then

$$\{x : \delta(x) \in Y\} = \cup_{i \in Y} Z_i \subseteq (E \setminus Z_{i_0})$$

and we conclude that $\overline{\mu}_{i_0}\{x : \delta(x) \in Y\} = 0$. Hence, we obtain that

$$\{x : \delta(x) \in Y\} \in dom(\overline{\mu}_{i_0}).$$

Thus we have shown the validity of the relation

$$\{x : \delta(x) \in Y\} \in dom(\overline{\mu}_{i_0})$$

for an arbitrary $i_0 \in I$. Hence,

$$\{x : \delta(x) \in Y\} \in \cap_{i \in I} dom(\overline{\mu}_i) = S_1.$$

Because $B(I)$ contains all singletons of I, we conclude that

$$\overline{\mu}_h(\{x : \delta(x) = i\}) = \overline{\mu}_i(Z_i) = 1, \ \ \forall i \in I.$$

□

References

[1] I. Ibramkhalilov, A. Skorokhod, *Consistent Estimates of Parameters of Random Processes* (Russian), "Naukova Dumka", Kiev, 1980

[2] T. Jech, *Set Theory*, Springer-Verlag, Berlin, 2003

[3] A. Kharazishvili, *Topological Aspects of Measure Theory* (Russian), "Naukova Dumka", Kiev, 1984

[4] Z. Zerakidze, *Hilbert space of measures*, Ukrainian. Math. J., **38**, 2 (1986), 147-153

[5] Z. Zerakidze, *Generalization criteria of Neyman-Pearson*, In: International scientific conference "Information Technologies", Georgian Technical University, Tbilisi (2008), 22-24

[6] O. Purtukhia, Z. Zerakidze, *On consistent criteria of hypotheses testing for non-separable complete metric space*, Bulletin of TICMI, **23**, 2 (2019), 97-108

[7] P.E. Danko, A.G. Popov, T.Ya. Kozhevnikova, S.P. Danko, *Higher Mathematics in Exercises and Problems, II* (Russian), "Mir i Obrazovanie", Moscow, 2009