Anatolij Dvurečenskij

Laws of large numbers and the central limit theorems on a logic

Mathematica Slovaca, Vol. 29 (1979), No. 4, 397--410

Persistent URL: http://dml.cz/dmlcz/136220

Math. Slovaca 29, 1979, No. 4, 397—410

LAWS OF LARGE NUMBERS AND THE CENTRAL LIMIT THEOREMS ON A LOGIC

ANATOLIJ DVUREČENSKIJ

In this paper the notion of the independence of observables in a state on a logic, as it was introduced by Gudder [2], will be studied. Some generalized forms of the weak and strong law of large numbers and the central limit theorems for observables of a logic will be proved. The used methods are similar to those of the conventional probability theory.

1. Preliminary definitions and results

Let us suppose that L be a poset with the first and the last elements 0 and 1, respectively, and an orthocomplementation $\perp : a \mapsto a^\perp$ which satisfies (i) $(a^\perp)^\perp = a$ for all $a \in L$; (ii) if $a < b$, then $b^\perp < a^\perp$ for $a, b \in L$; (iii) $a \vee a^\perp = 1$ for all $a \in L$. We say that a, b are orthogonal and write $a \perp b$ if $a < b^\perp$. We further assume that if $a < b$, then $b = a \vee (b \wedge a^\perp)$ and if $\{a_i\}$ is a sequence of mutually orthogonal elements of L, then $\bigvee_i a_i \in L$. A poset L satisfying the above axioms will be called a logic ([7]).

A state is a map m from L into $\langle 0, 1 \rangle$ such that $m(1) = 1$ and $m\left(\bigvee_i a_i\right) = \sum_i m(a_i)$ if $a_i \perp a_j$ for $i \neq j$. A system M of states of a logic L is called a quite full system if the statement $m(b) = 1$ whenever $m(a) = 1$, $m \in M$, implies $a < b$. In [3] it is shown that if M is a quite full system and L has at least three elements, then (i) $M \neq \emptyset$; (ii) if $a \neq 0$, then there is $m \in M$ such that $m(a) = 1$; (iii) $a = b$ iff $m(a) = m(b)$ for all $m \in M$; (iv) $a < b$ iff $m(a) \leq m(b)$ for all $m \in M$.

Lemma 1.1. (i) *Let M be a system of states, $a \in L$, and let us define $a^{1(M)} = \{m \in M: m(a) = 1\}$. Then M is a quite full system of states iff the statement $a^{1(M)} \subset b^{1(M)}$ implies $a < b$.*

(ii) *Let* M_1 *be a quite full system of states and* M_2 *be a system of states, then* $M = M_1 \cup M_2$ *is a quite full system of states.*

(iii) *If* M *is a system of states and* $Co(M) = \left\{ \sum_i c_i m_i \colon c_i \geqslant 0, \sum_i c_i = 1, m_i \in M, i \in I \right\}$, *then* M *is a quite full system iff* $Co(M)$ *is a quite full system.*

(iv) *Let* H *be a separable Hilbert space over real or complex scalars, let* $L(H)$ *be a logic of all closed subspaces of* H *and* $M = \{m_x \colon m_x(E) = (Ex, x), E \in L(H), x \in H, \|x\| = 1\}$. *Then* M *is a quite full system of states. If* $\dim H \geqslant 3$, *then* $Co(M)$ *is the system of all states on* $L(H)$; *if* $\dim H = 2$, *then* $Co(M)$ *is not the system of all states.*

Proof. The propositions (i)—(iii) are corollaries of the definition of the quite full system and (iv) is a corollary of the famous Gleason theorem [6].

Q.E.D.

An observable is a map x from the Borel sets $B(R_1)$ of R_1 into L such that (i) $x(R_1) = 1$; (ii) $x(E) \perp x(F)$ if $E \cap F = \emptyset$, $E, F \in B(R_1)$; (iii) $x\left(\bigcup_i E_i\right) = \bigvee_i x(E_i)$ if $E_i \cap E_j = \emptyset$, $i \neq j$, $E_i \in B(R_1)$. If f is a Borel function on R_1 and x an observable, then $f \circ x \colon E \mapsto x(f^{-1}(E))$, $E \in B(R_1)$, is an observable. We say that an observable x is bounded if there is a compact set C such that $x(C) = 1$. We denote by $\sigma(x)$ the smallest closed set E such that $x(E) = 1$ and $\|x\| = \sup\{|t| : t \in \sigma(x)\}$. The mean value of x in the state m is $m(x) = \int_{R_1} t \, \mathrm{d}m_x(t)$ if the integral on the right-hand side exists and is finite, where m_x is a measure on $B(R_1) \colon m_x(E) = m(x(E))$. $E \in B(R_1)$. In [3, Theorem 6.3] it is shown that an observable x is bounded iff $m(x)$ exists and is finite for every m on L.

Let $x_1, \ldots, x_n$ be observables (x_i may be unbounded for some $i = 1, 2, \ldots, n$) of a logic L. If there is a quite full system M of states and a unique observable z such that $m(x_1), \ldots, m(x_n)$ exist and are finite and $m(z) = m(x_1) + \ldots + m(x_n)$ for every $m \in M$, then z is called the sum of $x_1, \ldots, x_n$ and is written $z = x_1 + \ldots + x_n$.

If there is a quite full system M of states on L such that for any two bounded observables x, y there is a unique observable z such that $m(z) = m(x) + m(y)$ for all $m \in M$, then L is called a sum logic. In [3] it is shown that a sum logic is a lattice. From this moment we shall suppose that L is a sum logic.

Remark 1. Although $z = x + y$ exists on a sum logic, where x, y, z are bounded observables, $m(z) = m(x) + m(y)$ does not hold for every state m on L, in general.

Indeed, let $L = L(R_2)$. Due to (iv) of Lemma 1.1 it may be shown that $L(R_2)$ is a sum logic. Let

$$\mathbf{N}_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \mathbf{N}_2 = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}$$

and

$$\mathbf{N}=\mathbf{N}_1+\mathbf{N}_2=\begin{pmatrix}3/2 & 1/2\\ 1/2 & 1/2\end{pmatrix},$$

where the matrices $\mathbf{N}_1$, $\mathbf{N}_2$ and $\mathbf{N}$, respectively, correspond to the observables x, y and z [6]. If we choose the state m such that $m(N_\varphi)=1$ for every one-dimensional subspace N_φ with a direction angle φ, $0\leq\varphi<\frac{\pi}{2}$, then a straightforward computation shows that $m(\mathbf{N})=(2+\sqrt{2})/2 \neq m(\mathbf{N}_1)+m(\mathbf{N}_2)=2$.

We say that the observables $x_1, \ldots, x_n$ have a joint distribution in the state m if there is an n-dimensional Borel measure m_n such that

$$m_n(E_1\times\ldots\times E_n)=m\left(\bigwedge_{j=1}^{n} x_j(E_j)\right) \text{ for all } E_j\in B(R_1),$$

$$j=1, \ldots, n.$$

In this case we define (i) the joint distribution function $F_{x_1\ldots x_n}$: $(t_1, \ldots, t_n)\mapsto m(x_1(-\infty, t_1) \wedge\ldots\wedge x_n(-\infty, t_n))$ $t_j\in R_1$, $j=1, \ldots, n$; (ii) joint characteristic function $\varphi_{x_1\ldots x_n}$: $(u_1, \ldots, u_n)\mapsto \int_{R_n} \exp\left\{\mathrm{i}\sum_{j=1}^{n} u_jt_j\right\} \mathrm{d}m_n(t_1, \ldots, t_n)$, $u_j\in R_1$, $j=1, \ldots, n$; (iii) the moments $m(x_1^{k_1}, \ldots, x_n^{k_n}) = \int_{R_n} t_1^{k_1}\ldots t_n^{k_n}\, \mathrm{d}m_n(t_1, \ldots, t_n)$ if the integral on the right-hand side exists and is finite.

Due to a one-to-one correspondence between a joint characteristic function and a joint distribution function, respectively, we may transfer the investigation of properties of joint characteristic functions of observables of a logic in a state onto the investigation of joint distribution functions of observables, and conversely. This note is also valid for a characteristic function $\varphi_x(u) = \int_{R_1} \mathrm{e}^{\mathrm{i}ut}\, \mathrm{d}m_x(t)$ and $F_x(t) = m(x(-\infty, t))$ of one observable x.

Let x, x_1, x_2, ... be observables, F_x, F_1, F_2, ..., φ_x, φ_1, φ_2, ... be distribution functions and characteristic functions, respectively, in the state m. We say that a sequence $\{F_n\}$ converges weakly to F_x and write $F_n\xrightarrow{w}F_x$ if $F_n(t)\to F_x(t)$ at each continuity point of F_x. Due to the direct and inverse limit theorems [1] $F_n\xrightarrow{w}F_x$ iff $\varphi_n\to\varphi_x$. This result will be used in the following.

2. Independence

A system of observables $x_1, \ldots, x_n$ is independent in the state m if

$$m(x_1(E_1) \wedge \ldots \wedge x_n(E_n)) = m(x_1(E_1)) \ldots m(x_n(E_n)),$$
$$E_j \in B(R_1), \quad j = 1, \ldots, n.$$

A system of observables $\{x_t \colon t \in T\}$ is independent in m if any finite subsystem is independent. Let $a \in L$. We define the question observable q_a: $q_a(\{1\}) = a$, $q_a(\{0\}) = a^{\perp}$. We say that a system of elements $\{a_t \colon t \in T\}$ of L is independent in m if the corresponding question observables $\{q_{a_t} \colon t \in T\}$ are independent.

Theorem 2.1. *Let $x_1, \ldots, x_n$ be observables and m be a state. Then the following propositions are equivalent.*

(i) $m\left(\bigwedge_{j=1}^{n} x_j(E_j)\right) = \prod_{j=1}^{n} m(x_j(E_j))$, $E_j \in B(R_1)$, $j = 1, \ldots, n$.

(ii) *There is a joint distribution m_n in m such that* $m_n(E_1 \times \ldots \times E_n) = \prod_{j=1}^{n} m(x_j(E_j))$, $E_j \in B(R_1)$, $j = 1, \ldots, n$.

(iii) *There is a joint distribution function $F_{x_1 \ldots x_n}$ in m such that* $F_{x_1 \ldots x_n}(t_1, \ldots, t_n) = \prod_{j=1}^{n} F_{x_j}(t_j)$, $t_j \in R_1$, $j = 1, \ldots, n$.

(iv) *There is a joint characteristic function $\varphi_{x_1 \ldots x_n}$ in m such that* $\varphi_{x_1 \ldots x_n}(u_1, \ldots, u_n) = \prod_{j=1}^{n} \varphi_{x_j}(u_j)$, $u_j \in R_1$, $j = 1, \ldots, n$.

Proof. Let (i) hold. Let us define a set function μ on $B(R_n)$ by $\mu(E_1 \times \ldots \times E_n) = \prod_{j=1}^{n} m(x_j(E_j))$, $E_j \in B(R_1)$, $j = 1, \ldots, n$. Then it follows, by extending theorems, that there is a unique n-dimensional Borel measure m_n such that $m_n(E_1 \times \ldots \times E_n) = \prod_{j=1}^{n} m(x_j(E_j))$ and hence (ii) holds.

The converse implication is trivial.

The equivalence of (ii) and (iii) may be shown if we put $E_j = (-\infty, t_j)$, $j = 1, \ldots, n$.

Let now (i) hold; then, by (ii), there is a joint distribution m_n and consequently there is a joint characteristic function $\varphi_{x_1 \ldots x_n}$, hence

$$\varphi_{x_1 \ldots x_n}(u_1, \ldots, u_n) = \int_{R_n} \exp\left\{ \mathrm{i} \sum_{j=1}^{n} u_j t_j \right\} \mathrm{d}m_n(t_1, \ldots, t_n) =$$

$$= \prod_{j=1}^{n} \int_{R_1} \exp(\mathrm{i} u_j t_j)\, \mathrm{d}m_{x_j}(t_j) = \prod_{j=1}^{n} \varphi_{x_j}(u_j).$$

To prove (iv) implies (iii), we use the above mentioned one-to-one correspondence between a joint characteristic function and a joint distribution function, respectively, in the state m. Therefore

$$m_n(\langle a_1, b_1) \times \ldots \times \langle a_n, b_n)) =$$

$$= (2\pi)^{-n} \int_{R_n} \prod_{j=1}^{n} \frac{e^{it_j a_j} - e^{it_j b_j}}{it_j} \cdot \varphi_{x_1 \ldots x_n}(t_1, \ldots, t_n)\, dt_1 \ldots dt_n =$$

$$= \prod_{j=1}^{n} (2\pi)^{-1} \int_{R_1} \frac{e^{it_j a_j} - e^{it_j b_j}}{it_j} \varphi_{x_j}(t_j)\, dt_j = \prod_{j=1}^{n} m_{x_j}(\langle a_j, b_j)).$$

Passing to $a_j \to -\infty$, $j = 1, \ldots, n$, we obtain (iii).

Q.E.D.

We say that an observable x is nonnull a.e. $[m]$ if $m(x(\{0\}^c)) > 0$.

Theorem 2.2. *Let $x_1, \ldots, x_n$ be observables which are independent in the state m and nonnull* a.e. $[m]$. *Then the moment $m(x_1, \ldots, x_n)$ exists iff $m(x_j)$, $j = 1, \ldots, n$, exist and then in this case*

$$m(x_1, \ldots, x_n) = m(x_1) \ldots m(x_n).$$

Proof. Let m_n be a joint distribution of $x_1, \ldots, x_n$ in m and π_j, $j = 1, \ldots, n$, be the coordinate functions in R_n. Since π_j are independent on $(R_n, B(R_n), m_n)$, the proof of our theorem follows from the same proposition of the measure theory, therefore

$$m(x_1, \ldots, x_n) = \int_{R_n} \pi_1 \ldots \pi_n \, dm_n = \prod_{j=1}^{n} \int_{R_n} \pi_j \, dm_n = \prod_{j=1}^{n} m(x_j).$$

Q.E.D.

According to Gudder [2] we introduce the notion of a strong independence in a state, which converts into the notion of an idependence in a state in the conventional probability theory. This notion enables us to study properties of a characteristic function of the sum $z = x_1 + \ldots + x_n$ by means of $\varphi_{x_1}, \ldots, \varphi_{x_n}$; indeed, in this case $\varphi_z = \prod_{i=1}^{n} \varphi_{x_i}$. Thus, we say that the observables $x_1, \ldots, x_n$ are strongly independent in the state m if for any n Borel functions $f_1, \ldots, f_n$, for which $f_1 \circ x_1 + \ldots + f_n \circ x_n$ has a sense, we have

$$m_{f_1 \circ x_1 + \ldots + f_n \circ x_n} = m_{f_1 \circ x_1} * \ldots * m_{f_n \circ x_n},$$

where the $*$ denotes the convolution.

As usual, a system $\{x_t : t \in T\}$ of observables is strongly independent in the state m if every finite subsystem is strongly independent in m and a collection $\{a_t : t \in T\}$ is strongly independent in m if $\{q_{a_t} : t \in T\}$ is strongly independent.

Let a be real, we define an observable I_a by $I_a(\{a\}) = 1$ and it may be shown that $aI_1 = I_a$, where on the left-hand side we have a constant function $a(t) = a$, $t \in R_1$, and $aI_1 \equiv a \circ I_1$.

If $m(x^2)$ is finite, then $m(x)$ is finite, too, and we define the variance $V_m(x)$ of an observable x in the state m by

$$V_m(x) = m((x - m(x)I_1))^2,$$

and in this case the Chebyshev inequality $m((x - m(x)I_1)\ ((-\varepsilon, \varepsilon)^c)) \leq V_m(x)\varepsilon^{-2}$ holds.

Observables $x_1, \dots, x_n$ are uncorrelated in the state m if

$$V_m(x_1 + \dots + x_n) = V_m(x_1) + \dots + V_m(x_n),$$

if the sum $x_1 + \dots + x_n$ exists, and a system $\{x_t : t \in T\}$ of observables is uncorrelated in m; if every finite subsystem is uncorrelated in m.

In [2, Theorem 4.5, Lemma 4.7] it is shown that if a system $\{x_t : t \in T\}$ of observables is strongly independent in the state m, then it is independent and uncorrelated in m. The converse question, that is, whether the independence in m implies the strong independence, seems to be open for some classes of logics or systems of observables and states. A partial answer is given in [2] for one-dimensional subspaces in a pure state on $L(H)$.

3. Convergence of observables

For purposes of the last three sections of this paper we need to introduce some types of convergences of observables with respect to a state. These forms of convergences are equivalent to the corresponding convergences in the conventional probability theory. If $x, x_1, x_2, \dots$ are observables and m is a state, then we say that (i) x_n converges in the state m to x, and write $x_n \xrightarrow{m} x$ if $\lim_{n\to\infty} m((x - x_n)\ ((-\varepsilon, \varepsilon)^c)) = 0$ for every $\varepsilon > 0$; (ii) x_n converges almost everywhere $[m]$ to x, and write $x_n \to x$ a.e. $[m]$ if $m(\lim_n \sup\ (x_n - x)\ ((-\varepsilon, \varepsilon)^c)) = 0$ for every $\varepsilon > 0$; (iii) x_n converges in the square mean (m) to x if $\lim_{n\to\infty} m((x - x_n)^2) = 0$.

Lemma 3.1. *There holds $x_n \xrightarrow{m} I_a$ iff $F_n \xrightarrow{w} F_a$, where F_n, F_a are distribution functions of x_n and I_a, respectively, in the state m.*

Proof. It is easy to see that $F_a(t) = 0$ for $t \leq a$ and $F_a(t) = 1$ for $t > a$. Let $x_n \xrightarrow{m} I_a$ and $t < a$; then for every $\varepsilon > 0$ such that $t + \varepsilon < a$ we have

$$F_n(t) = m(x_n(-\infty, t)) \leq m(x_n(-\infty, a - \varepsilon)) \leq m((x_n - I_a)((-\varepsilon, \varepsilon)^c)) \to 0,$$

by the assumption. Similarly, for $t > a$ and $\varepsilon > 0$ such that $t \geq a + \varepsilon$ we have

$$1-F_n(t)=m(x_n(\langle t,\infty)))\leq m(x_n(\langle a+\varepsilon,\infty)))\leq m((x_n-I_a)((-\varepsilon,\varepsilon)^c))\to 0,$$

hence $F_n(t)\to 1$ and the necessity is proved.

To prove the sufficiency, let $F_n\xrightarrow{w}F_a$, that is $F_n(t)\to F_a(t)$ for every $t\neq a$. Then for an arbitrary $\varepsilon>0$

$$F_n(a+\varepsilon)-F_n(a-\varepsilon)=m(x_n(\langle a-\varepsilon,a+\varepsilon)))=m((x_n-I_a)(\langle -\varepsilon,\varepsilon)))\to 1,$$

therefore $x_n\xrightarrow{m}I_a$.

Q.E.D.

Lemma 3.2. *Let $x_n\xrightarrow{m}I_a$ and let f be a Borel function which is continuous at a. Then $f\circ x_n\xrightarrow{m}I_{f(a)}$.*

Proof. From the continuity of f at a it follows that for each $\varepsilon>0$ there is a $\delta>0$ such that $|f(t)-f(a)|<\varepsilon$ for $|t-a|<\delta$. The convergence $x_n\xrightarrow{m}I_a$ implies that for this δ and for sufficiently large $n m((x_n-I_a)\ ((-\delta,\delta)^c))$ is enough small. Moreover, $m((f\circ x_n-I_{f(a)})\ ((-\varepsilon,\varepsilon)^c))\ =\ m(f\circ(x_n-I_a)\ ((-\varepsilon,\varepsilon)^c))\ \leq\ m((x_n-I_a)\ ((-\delta,\delta)^c))\to 0$.

Q.E.D.

4. Weak laws of large numbers

From this moment we suppose in theses three last sections that $\{x_i\}$ is such a sequence of observables (they may be unbounded, too) of a sum logic L that for all $n=1,2,\ldots$ the sum $x_1+\ldots+x_n$ exists, and m is such state that if $m(x_1),\ldots,m(x_n)$ exist and they are finite, $n=1,2,\ldots$, then $m(x_1+\ldots+x_n)=m(x_1)+\ldots+m(x_n)$ (see Remark 1).

Let $\{x_i\}$ be a sequence of observables and m be a state; we shall investigate convergences of $s_n=\frac{1}{n}\sum_{i=1}^{n}(x_i-m(x_i)I_1)$ if $m(x_i)$, $i=1,2,\ldots$, is finite, with respect to the state m.

We say that for a sequence $\{x_i\}$

(i) the weak law of large numbers (w.l.l.n.) in the state m holds if $s_n\xrightarrow{m}0$;

(ii) the strong law of large numbers (s.l.l.n.) in the state m holds if $s_n\to 0$ a.e. $[m]$;

(iii) the square mean law of large numbers (s.m.l.l.n.) in the state m holds if $m(s_n^2)\to 0$;

(iv) the central limit theorem (c.l.t.) in the state m holds if distribution functions of ns_n/B_n convergence to $F(t)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{t}\exp\left(-\frac{u^2}{2}\right)du$ at each $t\in R_1$, where $B_n^2=V_m(x_1+\ldots+x_n)$.

Theorem 4.1. [2] (Chebyshev) *If $\{x_i\}$ is a sequence of uncorrelated observables in the state m and there is a constant K such that $m(x_i^2)\leq K$, $i=1, 2, \ldots$, then for $\{x_i\}$ the w.l.l.n. and the s.m.l.l.n. hold.*

Proof. Due to the finiteness and boundedness of $m(x_i^2)$, $i=1, 2, \ldots$, and by the Chebyshev inequality, we have

$$m(s_n((-\varepsilon,\varepsilon)^c))\leq\frac{1}{\varepsilon^2}V_m\left(\frac{1}{n}(x_1+\ldots+x_n)\right)=$$

$$=\frac{1}{\varepsilon^2 n^2}\sum_{i=1}^{n}V_m(x_i)\leq\frac{1}{\varepsilon 2_n}K\to 0.$$

Similarly, $m(s_n^2)=V_m\left(\frac{1}{n}(x_1+\ldots+x_n)\right)\to 0.$

Q.E.D.

Corollary 4.1.1. *If $\{x_i\}$ is a sequence of observables which are uncorrelated in the state m and uniformly bounded, that is, there is a constant K such that $\|x_i\|\leq K$, $i=1, 2, \ldots$, then for $\{x_i\}$ the w.l.l.n. and the s.m.l.n. hold.*

Proof. Since $\|x_i\|\leq K$, then $m(x_i^2)\leq K^2$ and the rest of the proof follows from Theorem 4.1.

Q.E.D.

Corollary 4.1.2. *Let $\{x_i\}$ be a sequence of uncorrelated observables which have the same distribution function in the state m. Let $m(x_1^2)$ be finite and $m(x_1)=a$; then $\frac{1}{n}(x_1+\ldots+x_n)$ converges to I_a* (i) *in the state m*; (ii) *in the square mean (m).*

Proof. It is easy to verify that the proposition fulfils the assumptions of the Chebyshev theorem.

Q.E.D.

Theorem 4.2. (Kchinchin) *Let $\{x_i\}$ be a sequence of strongly independent observables, which have the same distribution function in the state m and let $m(x_1)=a$; then*

$$\frac{1}{n}(x_1+\ldots+x_n)\xrightarrow{m}I_a.$$

Proof. By substituting x_i-I_a for x_i, we may assume that $a=0$. According to Lemma 3.1 it suffices to show that a sequence of distribution functions of $y_n=\frac{1}{n}(x_1+\ldots+x_n)$ converges weakly to the distribution function of I_0, but due to the known theorems of the probability theory it is sufficient to examine a convergence of the corresponding functions φ_n of y_n in the state m to $\varphi_0(u) \equiv 1$, $u \in R_1$.

If φ is a characteristic function of x_1, then the strong independence of $\{x_i\}$ in m implies $\varphi_n(u) = \varphi^n\left(\frac{u}{n}\right)$ for all $u \in R_1$, $n = 1, 2, \ldots$ Since φ' exists, $\varphi'(0) = 0$, and φ' is continuous, then, by Taylor's theorem, we have $\varphi\left(\frac{u}{n}\right) = 1 + O\left(\frac{u}{n}\right)$, where $\lim_{n\to\infty} O\left(\frac{u}{n}\right) = 0$, $u \in R_1$. Therefore $\varphi^n\left(\frac{u}{n}\right) = \left(1 + O\left(\frac{u}{n}\right)\right)^n \to 1$ for every $u \in R_1$, and the proof is finished.

Q.E.D.

Theorem 4.3. (Markov) *If a sequence of observables satisfies in the state m the Markov condition*

$$\frac{1}{n^2} V_m(x_1 + \ldots + x_n) \to 0 \tag{4.1}$$

then for $\{x_i\}$ *the w.l.l.n. and the s.m.l.l.n. hold in the state* m.

Proof. Using the Chebyshev inequality we have

$$m(s_n((-\varepsilon, \varepsilon)^c)) \leq \frac{1}{\varepsilon^2 n^2} V_m(x_1 + \ldots + x_n) \to 0.$$

Q.E.D.

Remark 2. If $\{x_i\}$ is a sequence of uncorrelated observables in m, then the Markov condition (4.1) can be rewritten in the form $n^{-2} \sum_{i=1}^{n} V_m(x_i) \to 0$, therefore the Chebyshev theorem follows from the Markov theorem.

If x is an observable, then $y = \frac{x^2}{1+x^2}$ is an observable if we put $y = f \circ x$, where $f(t) = t^2/(1+t^2)$.

Theorem 4.4. *In order that the w.l.l.n. may hold in the state m for a sequence* $\{x_i\}$ *it is necessary and sufficient that*

$$\lim_{n\to\infty} m\left(\frac{\left(\sum_{i=1}^{n}(x_i - m(x_i)I_1)\right)^2}{n^2 + \left(\sum_{i=1}^{n}(x_i - m(x_i)I_1)\right)^2}\right) = 0. \tag{4.2}$$

Proof. The necessity follows from this calculus

$$m(s_n((-\varepsilon, \varepsilon)^c)) = \int_{|t| \geq \varepsilon} dm_{s_n}(t) \geq \int_{|t| \geq \varepsilon} t^2/(1+t^2)\, dm_{s_n}(t) =$$

$$= \int t^2/(1+t^2)\, \mathrm{d}m_{s_n}(t) - \int_{|t|<\varepsilon} t^2/(1+t^2)\, \mathrm{d}m_{s_n}(t) \geq$$

$$\geq \int t^2/(1+t^2)\, \mathrm{d}m_{s_n}(t) - \varepsilon^2 = m\left(\frac{s_n^2}{1+s_n^2}\right) - \varepsilon^2.$$

Therefore

$$0 \leq m\left(\frac{s_n^2}{1+s_n^2}\right) \leq \varepsilon^2 + m(s_n((-\varepsilon, \varepsilon)^c))$$

and the (4.2) holds.

The sufficiency. Let (4.2) hold, then for each $\varepsilon > 0$ we have

$$m(s_n((-\varepsilon, \varepsilon)^c)) = \int_{|t|\geq\varepsilon} \mathrm{d}m_{s_n}(t) \leq (1+\varepsilon^2)\varepsilon^{-2} \int_{|t|\geq\varepsilon} t^2/(1+t^2)\, \mathrm{d}m_{s_n}(t) \leq$$

$$\leq (1+\varepsilon^2)\varepsilon^{-2} \int t^2/(1+t^2)\, \mathrm{d}m_{s_n}(t) = (1+\varepsilon^2)\varepsilon^{-2} m\left(\frac{s_n^2}{1+s_n^2}\right) \to 0.$$

Q.E.D.

Remark 3. We can show that Theorems 4.1—4.3 follow from Theorem 4.4. For example, if the Markov condition (4.1) is satisfied, then

$$m\left(\frac{s_n^2}{1+s_n^2}\right) = \int t^2/(1+t^2)\, \mathrm{d}m_{s_n}(t) \leq \int t^2\, \mathrm{d}m_{s_n}(t) = n^{-2} V_m(x_1 + \ldots + x_n) \to 0.$$

Theorem 4.5. *If a sequence of observables $\{x_i\}$ satisfies the following condition in the state m*

$$\lim_{n\to\infty} n^{-(1+\delta)} m\left(\left|\sum_{i=1}^{n}(x_i - m(x_i)I_1)\right|^{1+\delta}\right) = 0 \tag{4.3}$$

for $0 < \delta \leq 1$, then the w.l.l.n. holds.

Proof. According to Theorem 4.4, it suffices to verify (4.2).

Indeed,

$$m\left(\frac{s_n^2}{1+s_n^2}\right) = \int t^2/(1+t^2)\, \mathrm{d}m_{s_n}(t) \leq \int |t|^{1+\delta}\, \mathrm{d}m_{s_n}(t) =$$

$$= n^{-(1+\delta)} m\left(\left|\sum_{i=1}^{n}(x_i - m(x_i)I_1)\right|^{1+\delta}\right) \to 0.$$

Q.E.D.

Remark 4. If in (4.3) we put $\delta = 1$, then the Markov condition is satisfied.

5. Strong law of large numbers

Because of a somewhat more complicated investigation of the properties of the convergence almost everywhere $[m]$ we introduce the following definition.

If every finite subsystem $x_{i_1}, \ldots, x_{i_n}$ of a sequence $\{x_i\}$ has a joint distribution m_n in the state m, then there is a unique measure μ (μ is the Kolmogorov measure on $(R_\infty, B(R_\infty))$ which is determined by all m_n. We shall say that a sequence $\{x_i\}$ of observables which has the Kolmogorov measure μ in the state m satisfies the Cezaro joint distribution condition in the state m if

$$m\left(\bigvee_{k=n}^{m} \frac{1}{k} \sum_{i=1}^{k}(x_i - m(x_i)I_1)((-\varepsilon, \varepsilon)^c)\right) = \mu\left(\bigcup_{k=n}^{m}\left\{t \in R_\infty : \left|\frac{1}{k}\sum_{i=1}^{k}(\pi_i(t) - m(x_i))\right| \geq \varepsilon\right\}\right) \tag{5.1}$$

holds for every $\varepsilon > 0$ and $n \leq m$, where $\pi_i(t) = t_i$, $i = 1, 2, \ldots, t = (t_1, t_2, \ldots) \in R_\infty$.

Theorem 5.1. *Let a sequence $\{x_i\}$ of independent observables in the state m satisfy the Cezaro joint distribution condition in m and let $\sum_{i=1}^{\infty} i^{-2} V_m(x_i) < \infty$; then the s.l.l.n. holds.*

Proof. The independence of $\{x_i\}$ in m implies that the Kolmogorov measure μ exists, and it is a product measure on $(R_\infty, B(R_\infty))$, that is, $\mu = m_{x_1} \times m_{x_2} \times \ldots$ Therefore a sequence of coordinate functions $\pi_i : t \to t_i$, $i = 1, 2, \ldots, t = (t_1, t_2, \ldots)$, is independent and $\int \pi_i \, d\mu = m(x_i)$, $\sigma^2(\pi_i) = V_m(x_i)$. Because of the validity of the strong law of large numbers for $\{\pi_i\}$ [4] we have

$$\frac{1}{n}\sum_{i=1}^{n}(\pi_i - \int \pi_i \, d\mu) \to 0 \quad \text{a.e.} \quad [\mu].$$

Let now $\varepsilon > 0$; then

$$m(\lim_n \sup s_n((-\varepsilon, \varepsilon)^c)) = m\left(\bigwedge_{n=1}^{\infty} \bigvee_{k=n}^{\infty} s_k((-\varepsilon, \varepsilon)^c)\right) =$$

$$= \lim_{n\to\infty} \lim_{m\to\infty} m\left(\bigvee_{k=n}^{m} s_k((-\varepsilon, \varepsilon)^c)\right) =$$

$$= \lim_{n\to\infty} \lim_{m\to\infty} \mu\left(\bigcup_{k=n}^{m}\left\{t \in R_\infty : \left|\frac{1}{k}\sum_{i=1}^{k}(\pi_i(t) - m(x_i))\right| \geq \varepsilon\right\}\right) =$$

$$= \mu\left(\lim_n \sup \left\{t \in R_\infty : \left|\frac{1}{n}\sum_{i=1}^{n}(\pi_i(t) - m(x_i))\right| \geq \varepsilon\right\}\right) = 0,$$

as it follows by [4, Theorem 4], and the proof is complete. Q.E.D.

Theorem 5.2. *(Kolmogorov's strengthened law of large numbers.) Let $\{x_i\}$ be a sequence of independent observables which have the same distribution function in the state m and let $\{x_i\}$ satisfy the Cezaro joint distribution condition (5.1) in m and let $m(x_1) = a$; then $\frac{1}{n}(x_1 + \ldots + x_n) \to I_a$* a.e. $[m]$.

Proof. A sequence of integrable coordinate functions $\{\pi_i\}$ is independent on $(R_\infty, B(R_\infty), \mu)$ and $\{\pi_i\}$ have the same distribution function with respect to $\mu = m_{x_1} \times m_{x_2} \times \ldots$ Because of the strengthened law of large numbers $\frac{1}{n} \sum_{i=1}^{n} (\pi_i - m(x_i)) \to 0$ a.e. $[\mu]$ [4]. The rest of the proof is analogical to the proof of the above theorem.

Q.E.D.

6. Central limit theorems

The role of the Gaussian distributions in the conventional probability theory and its applications are, doubtless, very important. Moreover, in the quantum field theory, which motivates the theory of logic, they give a satisfactory description of the radiation field of coherent sources, as it was remarked in [5, p. 372]. We now show the role of the Gaussian distribution from a purely probabilistic viewpoint for the logic theory. If a distribution function F_x of an observable x in the state m has the form $F_x(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} \exp\left(-\frac{u^2}{2}\right) du$, then F_x is called the Gaussian distribution.

Theorem 6.1. (Lindeberg—Levy) *Let $\{x_i\}$ be a sequence of strongly independent observables which have the same distribution function in the state m and let $m(x_1^2)$ be finite; then for $\{x_i\}$ the c.l.t. holds in m.*

Proof. The proof is analogical to that of the conventional probability theory. Let φ be a characteristic function of $(x_1 - m(x_1)I_1)/V_m(x_1)$; then $\varphi(0) = 1$, $\varphi'(0) = 0$, $\varphi''(0) = -1$ and φ'' is continuous. By Taylor's theorem we have $\varphi(u) = 1 - \frac{u^2}{2} + O(u^2)$, where $\lim_{u \to 0} \frac{O(u^2)}{u^2} = 0$. If φ_n is a characteristic function of $y_n = B_n^{-1} \sum_{i=1}^{n} (x_i - m(x_i)I_1)$ in m, then, by the strong independence,

$$\varphi_n(u) = \left(\varphi\left(\frac{u}{\sqrt{n}}\right)\right)^n = \left(1 - \frac{u^2}{2n} + O\left(\frac{u^2}{n}\right)\right)^n \to \exp\left(-\frac{u^2}{2}\right), \quad u \in R_1,$$

and this property implies a weak convergence of the corresponding distribution functions.

Q.E.D.

We say that a sequence $\{x_i\}$ of observables of a logic satisfies the Lindeberg condition in the state m if for every $\tau > 0$

$$\lim_{n \to \infty} B_n^{-2} \sum_{i=1}^{n} \int_{|t - m(x_i)| > \tau B_n} (t - m(x_i))^2 \, dm_{x_i}(t) = 0. \tag{6.1}$$

Theorem 6.2. (Lindeberg) *If a sequence $\{x_i\}$ of strongly indenpendent observables satisfies the Lindeberg condition (6.1) in the state m, then the c.l.t. holds for $\{x_i\}$ in m.*

Proof. Since a sequence of coordinate functions $\{\pi_i\}$ is independent on $(R_\infty, B(R_\infty), \mu)$ and they satisfy the Lindeberg condition, then, by the known result of the probability theory [1], it follows that the distribution functions F_n of $\sum_{i=1}^{n}(\pi_i - E(\pi_i))/\sigma\left(\sum_{i=1}^{n}\pi_i\right)$ converge weakly to the distribution function of $N(0, 1)$. Because of the equality of F_n and the distribution function of y_n, respectively, the remaining part of the proof is shown.

Remark 5. Theorem 6.1 follows from Theorem 6.2. Indeed, let $\tau > 0$; then $B_n = \sigma\sqrt{n}$, where $\sigma^2 = V_m(x_1)$. If $m(x_1) = a$, then

$$\sum_{i=1}^{n} B_n^{-2}\int_{|t-m(x_i)|>\tau B_n}(t-m(x_i))^2\,\mathrm{d}m_{x_i}(t) =$$

$$= \sigma^{-2}\int_{|t-a|>\tau\sigma_n}(t-a)^2\,\mathrm{d}m_{x_1}(t) \to 0.$$

Theorem 6.3. (Ljapunov) *If for a sequence $\{x_i\}$ of strongly independent observables in the state m we may choose an $\delta > 0$ such that*

$$\lim_{n\to\infty} B_n^{-(2+\delta)}\sum_{i=1}^{n}\int|t-m(x_i)|^{2+\delta}\,\mathrm{d}m_{x_i}(t) = 0, \tag{6.2}$$

then for $\{x_i\}$ the c.l.t. holds in m.

Proof. It suffices to verify that the condition (6.2) implies the validity of (6.1). Indeed,

$$B_n^{-2}\sum_{i=1}^{n}\int_{|t-m(x_i)|>\tau B_n}(t-m(x_i))^2\,\mathrm{d}m_{x_i}(t) \leqslant$$

$$\leqslant B_n^{-2}(\tau B_n)^{-\delta}\sum_{i=1}^{n}\int_{|t-m(x_i)|>\tau B_n}|t-m(x_i)|^{2+\delta}\,\mathrm{d}m_{x_i}(t) \leqslant$$

$$\leqslant \tau^{-\delta}B_n^{-(2+\delta)}\sum_{i=1}^{n}\int|t-m(x_i)|^{2+\delta}\,\mathrm{d}m_{x_i}(t) \to 0.$$

Q.E.D.

Corollary 6.3.1. *If a sequence of uniformly bounded observables $\{x_i\}$ is strongly independent in the state m and if $B_n \to \infty$, then the c.l.t. holds in m.*

Proof. If $\|x_i\| \leqslant K$, $i = 1, 2, \ldots$, for some K, then

$$\int|t-m(x_i)|^{2+\delta}\,\mathrm{d}m_{x_i}(t) \leqslant (2K)^{2+\delta}/B_n^{2+\delta} \to 0$$

for $n \to \infty$.

Q.E.D.

REFERENCES

[1] ГНЕДЕНКО, Б. В.: Курс теории вероятностей. Москва 1965.
[2] GUDDER, S. P.: Hilbert space, independence, and generalized probability. J. Math. Analysis and Appl., 20, 1, 1967, 48—61.
[3] GUDDER, S. P.: Uniqueness and existence properties of bounded observables. Pac. J. Math., 19, 1, 1966, 81—83.
[4] ХАЛМОШ, П. Р.: Теория меры. Москва 1953.
[5] HOLEVO, A. S.: Statistical decision theory for quantum systems. J. Multivar. Anal., 3, 4, 1973, 337—394.
[6] VARADARAJAN, V. S.: Geometry of quantum theory. New York 1968.
[7] VARADARAJAN, V. S.: Probability in physics and a theorem on simultaneous observability. Comm. Pure appl. Math., 15, 1962, 189—217.

Received November 11, 1977

Ústav merania a meracej techniky SAV
Dúbravská cesta
885 27 Bratislava

ЗАКОНЫ БОЛЬШИХ ЧИСЕЛ И ЦЕНТРАЛЬНЫЕ ПРЕДЕЛЬНЫЕ ТЕОРЕМЫ НА ЛОГИКЕ

Анатолий Двуреченский

Резюме

В работе исследуется понятие независимости на логике, как его завел Gudder [2]. Некоторые обобщенные формы слабого и сильного закона больших чисел и центральные теоремы для наблюдаемых на логике доказаны. Использованые методы подобные методам класической теории вероятностей.