Nieuw Archief voor Wiskunde (3) I, 129—145 (1953)

ANOTHER FORM OF THE WEAK LAW OF LARGE NUMBERS

BY

PROF. DR. D. VAN DANTZIG

In this paper the socalled (weak) "Law of Large Numbers" (LLN) is proved in a new and very simple "limit-free" form. As a prerequisite to the proof several properties of the functions $\gamma\ (a, b)$, defined below, are needed. As these refer to a single stochastic variable (vector) only, they cannot be said to belong to the proof of the main theorem. A few of them have been listed only because they might perhaps be of some use on other occasions. Most of their proofs are similar to steps occurring in the known proofs of LLN also, e.g. by LÉVY, FELLER and FREUDENTHAL. Neither Feller's "annoying" (according to FREUDENTHAL) supplementary condition, nor the inductive reasoning in Freudenthal's proof of necessity is needed. The somewhat more general assumptions, called the "modern" form of the limit-theorem according to M. LOÈVE, are used. Moreover the theorems are formulated and proved for any (finite) number of dimensions.

In section 1 the definitions and results are formulated, in 2 the auxiliary properties are proved, whereas 3 contains the proof of the main lemma and theorem.

1.

For every positive integer ν n_ν is a positive integer and $\boldsymbol{x}_{\nu 1}, \ldots, \boldsymbol{x}_{\nu n_\nu}$ [1]) are independent stochastic variables; no assumption is made about stochastic variables belonging to different values of the *first* suffix. The suffix ν will always run through the set N of

[1]) Stochastic vectors and variables are printed in bold type. In comparison with e.g. Loève's notations the suffixes ν and n have been interchanged.

positive integers, k through the set N_{n_ν}, consisting of the numbers $1, \ldots, n_\nu$. Further b_ν is a given sequence of positive numbers;

$$s_\nu \stackrel{\text{def}}{=} \Sigma k\ x_{\nu,k} \text{ [2]} \tag{1}$$

The problem is, to find necessary and sufficient conditions, in order that the following statement LLN holds for the *given* sequence b_ν: LLN. For every $\nu \in N$ a real number c_ν exists, such that for all $a > 0$

$$\lim_{\nu \to \infty} \mathcal{P}\{|s_\nu - c_\nu| > a\, b_\nu\} = 0 \text{ [3]}$$

We use the letters N, P, R, R_m for the sets of all natural, positive and real numbers and for m-dimensional Euclidean space respectively, and, if S denotes any set, the symbols '$\forall_x^S \ldots$' and '$\exists_x^S \ldots$.' for 'for all $x \in S \ldots$.' and 'an $x \in S$ exists such that ....' respectively. Then LLN may be written in the shorter form:

$$\text{LLN:} \qquad \forall_\nu^N \exists_{c_\nu}^R \forall_a^P \quad \lim_{\nu \to \infty} \mathcal{P}\{|\, s_\nu - c_\nu \,| > a\, b_\nu\} = 0$$

With any stochastic variable x, having the (cumulative) distribution function $F(x)$, we associate functions $\gamma(a, b)$ and $\gamma(b)$, where $a \in R$, $b \in P$, defined by [4]

$$\gamma(a, b) \stackrel{\text{def}}{=} \int \min(1, y^2)\, dF(a + by) \tag{2}$$

$$\gamma(b) \stackrel{\text{def}}{=} \inf_{a \in R} \gamma(a, b) \tag{3}$$

The integrand $\min(1, y^2)$, occurring in the definition of $\gamma(a, b)$ could be replaced by a more general function $V(y)$ without invalidating some of our results. In particular one could choose for $V(y)$ any absolutely continuous distribution function vanishing for $y \leqslant 0$ and approximating y^2 for small $y > 0$, i.e. impose the conditions

1. $V(y) = 0$ for $y \leqslant 0$
2. $\lim_{y \to \infty} V(y) = 1$
3. $\lim_{y \downarrow 0} y^{-2} V(y) = 1$
4. $V(y) = \int_0^y v(t)\, dt$

[2]) The symbol $\stackrel{\text{def}}{=}$ denotes an equality, defining the *left* hand member.

[3]) The script $\mathcal{P}$ denotes 'the probability of' ....

[4]) In comparison with Feller's and Loève's notations the letters a and b, and n and ν respectively have been interchanged; $\min(u, v)$ denotes the minimum of u and v.

These conditions are e.g. also satisfied by the function $y^2(1 + y^2)^{-1}$ $(y \geq 0)$, used for similar purposes by P. LÉVY, A. KHINTCHINE, B. GNEDENKO and other authors in the theory of infinitely divisible laws. The common truncation method, corresponds with the choice

$$V(y) = \begin{cases} 1 \text{ if } y \geqslant 1 \\ 0 \text{ if } y < 1 \end{cases}$$

The fact that it does not satisfy condition 3 is the main cause of the necessity of separate conditions for the truncated second moments.

The main advantage of our choice $V(y) = \min(1, y^2)$ above that used by LÉVY, etc. is expressed by our auxiliary properties 13, 14, namely by the fact that by minimizing $\gamma(a, b)$ for constant b, the truncated first moment of $b^{-1}(\boldsymbol{x} - a)$ *vanishes* in the minimum $a = \xi$, whereas for other choices of $V(y)$ it only becomes small (conditions 3, 4 then only cause the vanishing of the expectation of $V(y)$).

We denote by $F_{\nu k}(x)$ and $G_\nu(x)$ the distribution functions, by $f_{\nu k}(t)$ and $g_\nu(t)$ the characteristic functions, and by $\gamma_{\nu k}(a, b)$, $\gamma_{\nu k}(b)$ and $\Gamma_\nu(a, b)$, $\Gamma_\nu(b)$ the newly defined functions belonging to the $\boldsymbol{x}_{\nu k}$ and the $\boldsymbol{s}_\nu$ respectively.

We shall prove the

Lemma 1. *LLN is equivalent with*

$$\lim_{\nu \to \infty} \Gamma_\nu(b_\nu) = 0 \tag{4}$$

This lemma admits of a "limit-free form". Introducing as an abbreviation

$$R_\nu(x) \stackrel{\text{def}}{=} \inf_{c_\nu} \mathcal{P}\{| \boldsymbol{s}_\nu - c_\nu | > x\} \tag{5}$$

it is implied by

Lemma 2. *The infimum in* (5) *is a minimum, and*

$$\forall_\nu^{\mathrm{N}} \, \forall_\alpha^{\mathrm{P}} \, \frac{\Gamma_\nu(b_\nu) - \alpha^2}{1 - \alpha^2} \leqslant R_\nu(\alpha\, b_\nu) \leqslant \frac{\Gamma_\nu(b_\nu)}{\min(1, \alpha^2)}, \tag{6}$$

the first inequality holding if $0 < \alpha < 1$.

In fact, the second inequality shows that (4) implies $R_\nu(ab_\nu) \to 0$ for all $a > 0$, i.e. LLN. On the other hand the first inequality shows, that $R_\nu(ab_\nu) \to 0$ for *some* $a > 0$ and < 1 implies $\limsup_{\nu \to \infty} \Gamma_\nu(b_\nu) \leqslant a^2$.

Hence LLN, i.e. $R_\nu(ab_\nu) \to 0$ for *all* $a > 0$ implies $\Gamma_\nu(b_\nu) = 0$ (as identically $\Gamma_\nu(b) \geqslant 0$).

Feller's theorem, applied to the "modern" case, states: Under the "supplementary condition"

$$\exists^{\mathrm{P}}_{\lambda} \forall^{\mathrm{N}}_{\nu} \forall^{\mathrm{N}n_\nu}_{k} \quad \max(\mathcal{P}\{x_{\nu k} > 0\}, \mathcal{P}\{x_{\nu k} < 0\}) \leqslant 1 - \lambda$$

it is necessary and sufficient for LLN that simultaneously

$$\begin{cases} \lim\limits_{\nu\to\infty} \Sigma^k \mathcal{P}\{| x_{\nu k} | > b_\nu\} = 0 \\ \lim\limits_{\nu\to\infty} b_\nu^{-2} \Sigma^k \int\limits_{|x|\leqslant b_\nu} x^2 \, d\, F_{\nu k}(x) = 0 \end{cases} \tag{7}$$

Recently FREUDENTHAL showed how to avoid the supplementary condition. His theorem (not in his limit-free form) can be stated thus: For LLN it is necessary and sufficient that simultaneously

$$\forall^{\mathrm{N}}_{\nu} \forall^{\mathrm{N}n_\nu}_{k} \exists^{\mathrm{R}}_{a_{\nu k}} \begin{cases} \lim\limits_{\nu\to\infty} \Sigma^k \mathcal{P}\{|x_{\nu k} - a_{\nu k} | > b_\nu\} = 0 \\ \lim\limits_{\nu\to\infty} b_\nu^{-2} \Sigma^k \operatorname{var} x'_{\nu k} = 0 \end{cases} \tag{8}$$

where var $x'_{\nu k}$ denotes the variance of the "truncated" variable

$$x'_{\nu k} \stackrel{\text{def}}{=} \begin{cases} x_{\nu k} & \text{if} \quad | x_{\nu k} - a_{\nu k} | \leqslant b_\nu \\ a_{\nu k} & \text{if} \quad | x_{\nu k} - a_{\nu k} | > b_\nu \end{cases} \tag{9}$$

We shall prove

Theorem 1. *Necessary and sufficient for LLN is*

$$\lim_{\nu\to\infty} \Sigma^k \gamma_{\nu k}(b_\nu) = 0 \tag{10}$$

Using the abbreviation

$$\Sigma_\nu(b) \stackrel{\text{def}}{=} \Sigma^k \gamma_{\nu k}(b), \tag{11}$$

theorem 1 together with lemma 1 state that for $\nu \to \infty$ $\Gamma_\nu(b_\nu) \to 0$ if and only if $\Sigma_\nu(b_\nu) \to 0$. As the Γ_ν, $\gamma_{\nu k}$ and Σ_ν are $\geqslant 0$, it is therefore immediately seen to be a trivial consequence of its "limit-free form":

Theorem 2. *If* $\nu \in \mathrm{N}$:

$$(16\,\pi)^{-1}\, \Sigma_\nu(b_\nu) \leqslant \Gamma_\nu(b_\nu) \leqslant \Sigma_\nu(b_\nu) - \Sigma_\nu(b_\nu) \ln \Sigma_\nu(b_\nu) \tag{12}$$

the left hand and the right hand inequality under the additional condition that $6\pi\Gamma_\nu(b_\nu) \leqslant 1$ and $\Sigma_\nu(b_\nu) \leqslant 1$ *respectively.*

Neither the lemma nor the theorems contain existential quantifiers. Once it has been shown (cf. *13* in section 2) that $a_{\nu k} = \xi_{\nu k}$ exist such that $\gamma_{\nu k}(\xi_{\nu k}, b_\nu) = \gamma_{\nu k}(b_\nu)$, (10) is trivially equivalent with Freudenthal's (8). We shall, however, prove it independently.

Although the "modern form" of the theorems makes it possible to put everywhere $b_\nu = 1$ without loss of generality (as the $\boldsymbol{x}_{\nu k}$, $\boldsymbol{s}_\nu$ and c_ν can, for every ν, simultaneously be divided by b_ν), we shall not make use of this simplification.

We shall give the demonstrations such a form that they (unless special notice is given) remain valid if the stochastic variables $\boldsymbol{x}_\nu$ are replaced by stochastic *vectors* in an m-dimensional Euclidean space. In order not to burden readers interested in the one-dimensional case only with vector-notations we shall maintain the simple notations used for this special case for the general case also, with the following new interpretations. The letters x, y, s, r, a, ξ, c, t (with or without suffixes) denote vectors in R_m, whereas the letters b, α, ε, l, T, q, γ, Γ, Σ, p, R, σ'^2 remain, like before, non-negative (or positive) real numbers. A product like tx or ty of two vectors denotes their scalar product; the square $(x - a)^2$ or y^2 of a vector its scalar product with itself. The length of a vector, e.g. $x - a$ is denoted by $|x - a|$, so that $(x - a)^2 = |x - a|^2$. We have, however, instead of the identity $|tx| = |t||x|$ only Cauchy's inequality $|tx| \leqslant |t||x|$.

Even a somewhat more general interpretation is possible, by interpreting $y^2 = |y|^2$ as an arbitrary positive definite quadratic form in y (and, just so, $(x - a)^2$ etc.), and ty as the corresponding bilinear form in t and y.

With these interpretations the lemmata and theorem 1 remain valid without alterations, whereas in theorem 2 only the numerical coefficients must be changed. The lemmata and theorems are then invariant with respect to arbitrary affine transformations in R_m, provided that the values of the distribution functions are kept invariant. In fact we have

Theorem 2'. In R_m, *if* $\nu \in N$

$$({}^{16}/_{3}\pi(m+2))^{-1}\,\Sigma_\nu(b_\nu) \leqslant \Gamma_\nu(b_\nu) \leqslant \Sigma_\nu(b_\nu) - \Sigma_\nu(b_\nu)\ln \Sigma_\nu(b_\nu) \quad (12')$$

the left hand and the right hand inequality under the additional condition that $2\pi(m+2)\Gamma_\nu(b_\nu) \leqslant 1$ *and* $\Sigma_\nu(b_\nu) \leqslant 1$ *respectively.*

Evidently this implies theorem 2.

2.

In this section we consider a single stochastic variable (generally a vector) $\boldsymbol{x}$ with distribution function $F(x)$ and characteristic function $f(t) = \mathcal{E}\, e^{it\boldsymbol{x}}$ ($\mathcal{E}$ denoting the expectation operator), and we shall derive the prerequisite [5]) properties of the functions

1. $$\gamma(a, b) \overset{\text{def}}{=} \int \min(1, y^2)\, dF(a + by)$$

2. $$\gamma(b) \overset{\text{def}}{=} \inf_{a \in \mathrm{R}} \gamma(a, b)$$

where a is an arbitrary real number (generally a vector in R_m) and b an arbitrary positive number.

3. $$\forall_a^{\mathrm{R}_m} \forall_b^{\mathrm{P}} \quad 0 \leqslant \gamma(b) \leqslant \gamma(a, b) \leqslant 1$$

Trivial.

4. $$\forall_a^{\mathrm{R}_m} \forall_b^{\mathrm{P}} \quad \begin{aligned} \gamma(a, b) &\geqslant \mathcal{P}\{|\,\boldsymbol{x} - a\,| > b\} \\ 1 - \gamma(a, b) &\leqslant \mathcal{P}\{|\,\boldsymbol{x} - a\,| \leqslant b\} \end{aligned}$$

Trivial.

5. $$\forall_a^{\mathrm{R}_m} \forall_b^{\mathrm{P}} \quad \begin{aligned} 1 - \gamma(a, b) &= \int_{|y| \leqslant 1} (1 - y^2) dF(a + by) = \\ &= \int_0^1 \mathcal{P}\{|\,\boldsymbol{x} - a\,| \leqslant b\alpha\} d(\alpha^2) \\ &= \mathcal{P}\{|\,\boldsymbol{x} - a\,| \leqslant b\boldsymbol{q}\}, \end{aligned}$$

where $\boldsymbol{q}$ is a stochastic variable, independent of $\boldsymbol{x}$, with distribution function

$$V(q) \overset{\text{def}}{=} \mathcal{P}\{\boldsymbol{q} \leqslant q\} \overset{\text{def}}{=} \begin{cases} 0 & \text{if } q \leqslant 0 \\ q^2 & \text{if } 0 \leqslant q \leqslant 1 \\ 1 & \text{if } 1 \leqslant q \end{cases}$$

Proof: obvious, as

$$\begin{aligned} \mathcal{P}\{|\,\boldsymbol{x} - a\,| \leqslant b\boldsymbol{q}\} &= \int dV(q)\, \mathcal{P}\{|\boldsymbol{x} - a\,| \leqslant bq\} = \\ &= \int dF(x)\, \mathcal{P}\{\boldsymbol{q} \geqslant b^{-1}\,|\,x - a\,|\} \end{aligned}$$

Remark. Evidently also

$$\begin{aligned} \gamma(a, b) &= \mathcal{P}\{|\boldsymbol{x} - a\,| > b\boldsymbol{q}\} = \\ &= 1 - \int_{(x-a)^2 \leqslant b^2} (1 - b^{-2}(x - a)^2) dF(x) \end{aligned}$$

6. $$\forall_a^{\mathrm{R}_m} \forall_b^{\mathrm{P}} \forall_\alpha^{\mathrm{P}} \quad \gamma(a, b) \leqslant \alpha^2 + (1 - \alpha^2)\mathcal{P}\{|\boldsymbol{x} - a\,| > \alpha b\}$$

Proof: As for each $q \geqslant \alpha \geqslant 0$ $\mathcal{P}\{|\boldsymbol{x} - a\,| > qb\} \leqslant \mathcal{P}\{|\,\boldsymbol{x} - a\,| \geqslant \alpha b\}$ we have from *5*

---

[5]) Some of the properties, marked by an asterisk, are not needed for the proofs of the theorems, but are listed for completeness only.

$$\gamma(a, b) = \int_{q^2 \leqslant 1} \mathcal{P}\{|\,\boldsymbol{x} - a\,| > qb\} d(q^2) \leqslant$$
$$\leqslant \int_{q^2 \leqslant a^2} d(q^2) + \int_{a^2 \leqslant q^2 \leqslant 1} \mathcal{P}\{|\,\boldsymbol{x} - a\,| > ab\} d(q^2),$$

whence *6*.

7. $\forall_a^{\mathrm{R}_m} \forall_b^{\mathrm{P}} \forall_a^{\mathrm{P}} \quad \gamma(a, b) \geqslant \min(1, a^2)\, \mathcal{P}\{|\boldsymbol{x} - a\,| > ab\}$

Proof: $\gamma(a, b) = \mathcal{P}\{|\,\boldsymbol{x} - a\,| > \boldsymbol{q}b\}$
as $0 \leqslant \boldsymbol{q} \leqslant 1$, hence $\geqslant \mathcal{P}\{|\,\boldsymbol{x} - a\,| > ab\}$ if $a \geqslant 1$, and, if $a \leqslant 1$, $\geqslant \mathcal{P}\{|\boldsymbol{x} - a\,| > ab$ and $a \geqslant \boldsymbol{q}\} = \mathcal{P}\{|\boldsymbol{x} - a\,| > ab\}\, \mathcal{P}\{\boldsymbol{q} \leqslant a\} =$
$= a^2\, \mathcal{P}\{|\,\boldsymbol{x} - a\,| > ab\}$

whence *7*.

8. $$\forall_b^{\mathrm{P}} \lim_{|a| \to \infty} \gamma(a, b) = 1$$

Proof. Choose $\varepsilon > 0$, $l \geqslant 0$ with $\mathcal{P}\{|\,\boldsymbol{x}\,| > l\} \leqslant \varepsilon$ and $|\,a\,| > b + l$. Then $1 - \gamma(a, b) \leqslant \mathcal{P}\{|\,\boldsymbol{x} - a\,| \leqslant b\} \leqslant \mathcal{P}\{|\,\boldsymbol{x}\,| > l\} \leqslant \varepsilon$.

*9**. $$\forall_a^{\mathrm{R}_m} \lim_{b \to \infty} \gamma(a, b) = \lim_{b \to \infty} \gamma(b) = 0$$

Proof. Choose $\varepsilon > 0$ and $\leqslant 1$ and l so that $\mathcal{P}\{|\,\boldsymbol{x} - a\,| > l\} \leqslant \varepsilon$ and take $b \geqslant \varepsilon^{-1/2} l$. Then, for any $a \in \mathrm{R}_m$:

$$0 \leqslant \gamma(b) \leqslant \gamma(a, b) = \int_{b|y| \leqslant l} y^2 dF(a + by) + \int_{b|y| > l} \min(1, y^2) dF(a + by) \leqslant$$
$$\leqslant b^{-2} l^2 + \mathcal{P}\{|\boldsymbol{x} - a\,| > l\} \leqslant 2\varepsilon$$

*10**. $$\forall_a^{\mathrm{R}_m} \lim_{b \to 0} \gamma(a, b) = 1 - \mathcal{P}\{\boldsymbol{x} = a\}$$

Proof. By *5*

$$\lim_{b \to 0} (1 - \gamma(a, b)) = \lim_{b \to 0} \mathcal{P}\{|\,\boldsymbol{x} - a\,| \leqslant b\boldsymbol{q}\} =$$
$$= \mathcal{P}\{|\,\boldsymbol{x} - a\,| = 0\} = \mathcal{P}\{\boldsymbol{x} = a\}$$

as $\boldsymbol{q} \leqslant 1$.

*11**. If $a + b$ and $a - b$ are continuity points of $F(x)$, then

$$\frac{\partial \gamma(a, b)}{\partial b} = -\frac{2}{b} \int_{|y| \leqslant 1} y^2 dF(a + by)$$

Proof. By *5*

$$1 - \gamma(a, b) = 2b^{-2} \int_0^b \mathcal{P}\{|\,\boldsymbol{x} - a\,| \leqslant \beta\} \beta\, d\beta$$

is for $b > 0$ a continuous alnost every where differentiable function of b, as both factors of the last member are. Hence

*11'**. $$\frac{\partial \gamma(a, b)}{\partial b} = \frac{4}{b^3} \int_0^b \mathcal{P}\{|\,\boldsymbol{x} - a\,| \leqslant \beta\} \beta\, d\beta - \frac{2}{b} \mathcal{P}\{|\,\boldsymbol{x} - a\,| \leqslant b\}$$

whence *11** follows by partial integration. In particular *11** shows that $\gamma(a, b)$ for constant a is a monotonous non-increasing function of b.

Remark. *11* is equivalent with

*11''**. $$\frac{\partial \gamma(a, b)}{\partial (b^{-2})} = \int_{|x-a| \leqslant b} | x - a |^2 \, dF(x)$$

which is almost obvious with $\gamma(a, \lambda^{-1/2}) = \int \min(1, \lambda(x-a)^2) dF(x)$ and shows, together with *9** that

*11'''**. $$\gamma(a, b) = \int_b^\infty d(\beta^{-2}) \int_{|x-a| \leqslant \beta} (x - a)^2 dF(x)$$

*12**. If $\mathcal{P}\{| x - a | = b\} = 0$, then

$$\frac{\partial}{\partial a} \gamma(a, b) = -\frac{2}{b} \int_{|y| \leqslant 1} y^* dF(a + by)$$

where $\frac{\partial}{\partial a}$ (for $m > 1$) denotes the gradient of the scalar $\gamma(a, b)$ with respect to a, whereas $y^* = y$ for $m = 1$, and, generally, if y^2 stands for the ordinary scalar product; but, if y^2 is a general positive definite quadratic form, y^2 is the transform of y by the matrix of the quadratic form, i.e.

$$y^* = \frac{\partial}{\partial y} (^1/_2 y^2).$$

Proof. In the second member of the remark to *5* the boundaries of the integral and the integrand are differentiable functions of a, hence $\gamma(a, b)$ is. Differentiation of the boundary gives no contribution, as the integrand vanishes there. Differentiation of the integrand and substitution of $x = a + by$ leads immediately to *12**.

13. $$\forall_b^{\mathrm{P}} \, \exists_\xi^{\mathrm{R}_m} \quad \gamma(b) = \gamma(\xi, b)$$

Proof. As $\gamma(a, b)$ for constant b by *5* (or *12**) is continuous in a, and by *8* tends to 1 for $| a | \to \infty$, its infimum (cf. *2*) is a minimum, which it reaches for some finite $a = \xi$ (in general not uniquely determined), whence *13*.

Henceforth we shall denote by ξ some vector for which *13* holds, and by $M(b)$ the set of all these ξ.

14. $$\forall_b^{\mathrm{P}} \, \forall_\xi^{M(b)} \, \forall_t^{\mathrm{R}_m} \quad \int_{|y| \leqslant 1} ty \, dF(\xi + by) = 0$$

Remark. With the notation used in *12**, this is equivalent with

$$\int_{|y| \leqslant 1} y^* \, d \, F(\xi + by) = 0$$

Proof for $m = 1$. In this special case t is a numerical factor, which may be omitted. If $\partial\gamma/\partial a$ exists in the minimum $a = \xi$, it vanishes there, so that *14* follows from *12**. If, on the other hand for some a, only the left hand and the right hand derivatives exist, partial integration of *12** (using the assumption $m = 1$), leading to

$$\frac{\partial\gamma(a \pm 0, b)}{\partial a} = \frac{2}{b^2} \int_{a-b}^{a+b} F(x)dx - \frac{2}{b}(F(a+b\pm 0) + F(a-b\pm 0))$$

shows that the first term is continuous, whereas $F(a + b)$ can make only *upward* jumps for increasing a. Hence, in a discontinuity $\partial\gamma/\partial a$ can only make *downward* jumps. This, however, can not happen in a minimum. Hence $\partial\gamma/\partial a$ is continuous in a minimum so that the initial conclusion, whence *14*, holds (for $m = 1$).

As this simple proof can not easily be generalized for arbitrary m, we shall give a different proof for general m, not using *12**.

For arbitrary $a \in \mathrm{R}_m$, $a' \in \mathrm{R}_m$,

$$\begin{aligned}\gamma(a, b) - \gamma(a', b) &= \mathcal{P}\{|\boldsymbol{x} - a| > \boldsymbol{q}b\} - \mathcal{P}\{|\boldsymbol{x} - a'| > \boldsymbol{q}b\} = \\ &= \mathcal{P}\{|\boldsymbol{x} - a| > \boldsymbol{q}b \geqslant |\boldsymbol{x} - a'|\} - \mathcal{P}\{|\boldsymbol{x} - a'| > \boldsymbol{q}b \geqslant |\boldsymbol{x} - a|\} = \\ &= \int_{|x-a'|\leqslant \min(b, |x-a|)} \{\min(1, b^{-2}(x-a)^2) - (x-a')^2\}\, dF(x) - \\ &- \int_{|x-a|\leqslant \min(b, |x-a'|)} \{\min(1, b^{-2}(x-a')^2) - (x-a)^2\} dF(x).\end{aligned}$$

Now we put $a' = \xi \in M(b)$, $a = \xi - \varepsilon bt$ with $\varepsilon > 0$, $t \in \mathrm{R}_m$, then, putting $x = \xi + by$, $z = y + \varepsilon t$, we have

$$\begin{aligned}0 \leqslant \gamma(\xi - \varepsilon bt, b) - \gamma(b) &= \int_{|y|\leqslant|z|\leqslant 1} (z^2 - y^2)dF(\xi + by) + \\ + \int_{|y|\leqslant 1<|z|} (1 - y^2)\, dF(\xi + by) &- \int_{|z|\leqslant|y|\leqslant 1} (y^2 - z^2)dF(\xi + by) - \\ &- \int_{|z|\leqslant 1<|y|} (1 - z^2)dF(\xi + by)\end{aligned}$$

as $y^2 - z^2 = -2\varepsilon\, ty - \varepsilon^2 t^2$. Hence the first and the third term are together

$$\leqslant + 2\varepsilon \int_{\max(|y|, |z|)\leqslant 1} ty\, dF(\xi + by) + \varepsilon^2 t^2$$

The last term is $\leqslant 0$, whereas for $|z| \leqslant 1 < |y|$, because of $|y| \leqslant |z| + |y - z|$, $0 \leqslant 1 - |z| \leqslant |y - z| = \varepsilon t$, $1 - z^2 = (1 + |z|)(1 - |z|) \leqslant 2\varepsilon t$. Hence we find after division by $\varepsilon > 0$:

$$\begin{aligned}0 \leqslant \varepsilon^{-1}(\gamma(\xi - \varepsilon bt, b) - \gamma(b)) &\leqslant 2\int_{\max(|y|, |y-\varepsilon t|)\leqslant 1} ty dF(\xi + by) + \\ &+ \varepsilon t^2 + 2|t|\, \mathcal{P}\{1 - \varepsilon|t| < |\boldsymbol{y} - \varepsilon t| \leqslant 1\}\end{aligned}$$

As this holds for every $\varepsilon > 0$, we can pass to the limit $\varepsilon \to 0$. Then the last term in the last member tends to zero, like the second term, and we obtain

$$0 \leqslant \int_{|y|\leqslant 1} ty\, dF(\xi + by)$$

This is true for every $t \in R_m$, hence it remains valid after replacement of t by $-t$, whence *14* follows.

In the same way *12** could have been proved, by using Fréchet's definition of the differential of a functional, in casu

$$t \frac{\partial}{\partial a} \gamma(a, b) = \lim_{\varepsilon \to 0} \frac{\gamma(a + \varepsilon t) - \gamma(a)}{\varepsilon}.$$

This one is the most important among the properties of $\gamma(a, b)$ which allows us to avoid Feller's supplementary condition, as well as the complete induction in Freudenthal's proof of necessity. The main difference between their methods and ours lies in the fact that they have to consider all a, which (substituted for ξ) make the left member of *14 small*, whereas we, by minimizing $\gamma(a, b)$, could make it *zero*.

In the following theorems t always denotes a real vector.

15. $\quad \forall_b^{\mathrm{P}} \forall_\xi^{M(b)} \; |\int \sin ty \, dF(\xi + by)| \leqslant (\pi - 2)\gamma(b)$ if $|t| \leqslant 2$

Proof. For $|y| > 1 \quad |\sin ty| \leqslant 1$. For $|y| \leqslant 1 \quad \sin ty = ty - (ty - \sin ty)$. The integral of ty vanishes because of *14*. Moreover, for $0 \leqslant z \leqslant \pi$, $z^{-2}(z - \sin z)$ is a monotonous increasing function[6]) of z, so that for $|z| \leqslant T \leqslant \pi \;\; |z - \sin z| \leqslant (T - \sin T) z^2 T^{-2}$, and for $|y| \leqslant 1$, $|t| \leqslant T$: $|ty - \sin ty| \leqslant (T - \sin T) y^2$, since $T^{-2}(ty)^2 \leqslant T^{-2} t^2 y^2 \leqslant y^2$. Hence, taking $T = 2$ with $T - \sin T = 1{,}091 < \pi - 2$,

$$\begin{aligned} |\textstyle\int \sin ty \, dF(\xi + by)| &\leqslant |\textstyle\int_{|y| \leqslant 1}| + |\textstyle\int_{|y|>1}| \leqslant \\ &\leqslant (\pi - 2) \textstyle\int_{|y| \leqslant 1} y^2 dF(\xi + by) + \textstyle\int_{|y|>1} dF(\xi + by) \leqslant \\ &\leqslant (\pi - 2)\gamma(\xi, b) = (\pi - 2)\gamma(b) \end{aligned}$$

This theorem replaces Feller's [1] (16) and Freudenthal's [3] (12.4).

16. $\quad \forall_b^{\mathrm{P}} \forall_\xi^{M(b)} \; |\int (1 - e^{ity}) dF(\xi + by)| \leqslant \pi \gamma(b)$ if $|t| \leqslant 2$

Proof.

$$\begin{aligned} 0 &\leqslant \textstyle\int (1 - \cos yt) dF(\xi + by) = 2 \textstyle\int \sin^2 {}^1/_2 ty \, dF(\xi + by) \leqslant \\ &\leqslant 2 \textstyle\int \min ({}^1/_4 (ty)^2, 1) dF(\xi + by) \leqslant 2\gamma(\xi, b) = 2\gamma(b) \text{ for } |t| \leqslant 2 \end{aligned}$$

Hence

$$\begin{aligned} &|\textstyle\int (1 - e^{ity}) dF(\xi + by)| \leqslant \\ &\leqslant \textstyle\int (1 - \cos ty) \, dF(\xi + by) + |\textstyle\int \sin ty \, dF(\xi + by)| \leqslant \\ &\leqslant 2\gamma(b) + (\pi - 2)\gamma(b) = \pi\gamma(b) \end{aligned}$$

[6]) In fact, the derivative of $z^{-2}(z - \sin z)$ is

$$\frac{2 \sin z - z(1 + \cos z)}{z^3} = \frac{4 \cos^2 \frac{1}{2} z}{z^3} (\operatorname{tg} \tfrac{1}{2} z - \tfrac{1}{2} z),$$ which is positive for $0 < z < \pi$.

The argument is similar to Freudenthal's [3] (12.2)—(12.5).

17. $$\forall_b^{P}\,\forall_\xi^{M(b)}\ \left(\int_{|t|\leqslant T} dt\right)^{-1}\int_{|t|\leqslant T} dt \int \sin^2 ty\, dF(\xi + by) \geqslant$$
$$\geqslant \frac{3}{2(m+2)}\left(1 - \frac{1}{2T}\right)\gamma(b) \text{ if } 0 < T \leqslant {}^1/_2\pi.$$

Proof. By means of an affine transformation (cf. the remark before theorem 2′) we may reduce the quadratic form y^2 to its diagonal form $\Sigma\,(y^\lambda)^2$, where $\lambda = 1, \ldots, m$, and y^λ are the coordinates of y. The denominator then is the volume $T^m I_m$ of a sphere with radius T, where

$$I_m = \frac{\{(-\tfrac{1}{2})!\}^m}{(\tfrac{1}{2}m)!}$$

As the whole expression then is invariant under rotations, we may, after having interchanged the order of the integrations, for any fixed y, choose the orthogonal coordinates $t_1, \ldots, t_m$ of t so that the t_1-axis has the direction of y, i.e. $ty = t_1\,|\,y\,|$, and put $t_2^2 + \ldots + t_m^2 = T^2 r^2$. Then assuming first $m \geqslant 2$, the integration over $r =$ const. can be performed, and gives $O_{m-2}(Tr)^{m-2}$, where

$$O_{m-2} = \frac{2.\{(-\tfrac{1}{2})!\}^{m-1}}{\dfrac{m-3}{2}\,!}$$

is the area of a unit-sphere in $(m-1)$-dimensional Euclidean space.

We find that the left hand member of *17* is

$$\geqslant \frac{O_{m-2}}{I_m T}\int_0^1 r^{m-2}\,dr \int_{|t_1|\leqslant T\sqrt{1-r^2}} dt_1 \int \sin^2 t_1\,|\,y\,|\,dF(\xi + by) =$$
$$= \frac{O_{m-2}}{I_m}\int_0^1 r^{m-2} dr\sqrt{1-r^2}\int\left(1 - \frac{\sin 2T|y|\sqrt{1-r^2}}{2T\,|\,y\,|\,\sqrt{1-r^2}}\right) dF(\xi + by).$$

Abbreviating for a moment $|\,y\,|\,\sqrt{1-r^2}$ by u, we have

$$\left|\frac{\sin 2Tu}{2Tu}\right| \leqslant \frac{1}{2Tu},$$

hence, for $|\,u\,| > 1$:

$$1 - \frac{\sin 2Tu}{2Tu} > 1 - \frac{1}{2T} \geqslant \left(1 - \frac{1}{2T}\right)\min\{1, u^2\}.$$

Moreover, for $0 < z \leqslant \pi$, $z^{-3}(z - \sin z)$ is a monotonous decreasing function. In fact, it equals $\dfrac{1}{6} - \dfrac{z^2}{120} + \dfrac{z^4}{5040} - \ldots$; its derivative

$-\frac{z}{60}+\frac{z^3}{1260}-\dots$ also is an alternating series, hence it is $=-\vartheta\frac{z}{60}$ with $0<\vartheta<1$ if $z^2\leqslant\pi^2<21=\frac{1260}{60}$, hence it is negative. For $|u|\leqslant 1$, $0<2T\leqslant\pi$ we have then

$$1-\frac{\sin 2Tu}{2Tu}\geqslant\left(1-\frac{\sin 2T}{2T}\right)u^2\geqslant\left(1-\frac{1}{2T}\right)u^2\geqslant\left(1-\frac{1}{2T}\right)\min(1,u^2)$$

Therefore, for all y and $0<2T\leqslant\pi$

$$1-\frac{\sin 2T|y|\sqrt{1-r^2}}{2T|y|\sqrt{1-r^2}}\geqslant\left(1-\frac{1}{2T}\right)\min(1,\ y^2(1-r^2)\geqslant$$

$$\geqslant\left(1-\frac{1}{2T}\right)(1-r^2)\min(1,y^2)$$

Now, as

$$\int_0^1 r^{m-2}(1-r^2)^{3/2}dr=\tfrac{1}{2}\mathrm{B}\left(\frac{5}{2}.\frac{m-1}{2}\right)=\frac{m-3}{2}!\frac{3}{2}!\left(2.\left(\frac{m+2}{2}\right)!\right)^{-1}$$

$$=\frac{3}{4(m+2)}\frac{m-3}{2}!\,(-\tfrac{1}{2})!\,\{(\tfrac{1}{2}m)!\}^{-1}=\frac{3I_m}{2(m+2)O_{m-2}},$$

we see that the left member of *17* is

$$\geqslant\frac{3}{2(m+2)}\left(1-\frac{1}{2T}\right)\int\min(1,y^2)dF(\xi+by),$$

so that *17* has been proved for $m\geqslant 2$. For $m=1$ the integration over r does not occur; except for some simplifications the proof remains the same.

3.

It is now easy to prove lemma 2. The fact that the infimum in (5) is a minimum follows exactly in the same way as *13*. By applying *6* and *7* to $\boldsymbol{s}_\nu$ instead of $\boldsymbol{x}$ we have at once for any $a>0$, $b>0$ and any vector c_ν

$$\frac{\Gamma_\nu(c_\nu,b_\nu)-a^2}{1-a^2}\leqslant\mathcal{P}\{|\,\boldsymbol{s}_\nu-c_\nu\,|>ab_\nu\}\leqslant\frac{\Gamma_\nu(c_\nu,b_\nu)}{\min(1,a^2)}$$

(the left hand inequality for $a<1$ only).
For any c_ν the left hand member is $\geqslant(1-a^2)^{-1}(\Gamma_\nu^*(b_\nu)-a^2)$. Hence, choosing for c_ν a minimum η_ν of Γ_ν, so that $\Gamma_\nu(\eta_\nu,b_\nu)=\Gamma_\nu(b_\nu)$:

$$\frac{\Gamma_\nu(b_\nu)-a^2}{1-a^2}\leqslant R_\nu\leqslant\mathcal{P}\{|\,\boldsymbol{s}_\nu-\eta_\nu\,|>ab_\nu\}\leqslant\frac{\Gamma_\nu(b_\nu)}{\min(1,a^2)}.\qquad(13)$$

Now we can pass to the proof of theorem 2. As only one value of the suffix ν occurs in (12), we can consider any fixed ν and drop the suffixes ν altogether. Then we have to prove for any fixed sufficiently large positive b:

$$(16\pi)^{-1}\,\Sigma(b) \leqslant \Gamma(b) \leqslant \Sigma(b) - \Sigma(b)\ln\Sigma(b) \tag{12'}$$

where

$$\Sigma(b) \overset{\text{def}}{=} \Sigma^k \gamma_k(b) \tag{11'}$$

$\gamma_k(b)$ and $\Gamma(b)$ being the functions defined by (2), (3), belonging to the stochastic vectors $\boldsymbol{x}_k$ and $\boldsymbol{s} = \Sigma\,\boldsymbol{x}_k$ respectively, the distribution functions of which are $F_k(x)$ and $G(x)$ respectively.

For every $k \in \mathrm{N}_n = \{1, \ldots, n\}$ we choose according to *13* a ξ_k such that

$$\gamma_k(\xi_k, b) = \gamma_k(b) \tag{14}$$

Moreover we introduce for any $k \in \mathrm{N}_n$ the following abbreviations

$$p_k \overset{\text{def}}{=} \mathcal{P}\{|\,\boldsymbol{x}_k - \xi_k\,| > b\} \tag{15}$$

$$\boldsymbol{y}_k \overset{\text{def}}{=} b^{-1}(\boldsymbol{x}_k - \xi_k) \tag{16}$$

$$\boldsymbol{x}'_k \overset{\text{def}}{=} \begin{cases} \boldsymbol{x}_k & \text{if } |\,\boldsymbol{x}_k - \xi_k\,| \leqslant b \\ \xi_k & \text{if } |\,\boldsymbol{x}_k - \xi_k\,| > b \end{cases} \tag{17}$$

$$\boldsymbol{s}' \overset{\text{def}}{=} \Sigma^k\,\boldsymbol{x}'_k \tag{18}$$

$$c \overset{\text{def}}{=} \Sigma^k\,\xi_k \tag{19}$$

$$\sigma_k'^2 \overset{\text{def}}{=} \operatorname{var} \boldsymbol{x}'_k \tag{20}$$

$$\beta'^2 \overset{\text{def}}{=} \operatorname{var} \boldsymbol{s}' \tag{21}$$

Then, as

$$\mathcal{E}\,\boldsymbol{x}'_k = \xi_k \tag{22}$$

by *14*,

$$\sigma_k'^2 = \mathcal{E}(\boldsymbol{x}'_k - \xi_k)^2 = b^2 \int_{|y|\leqslant 1} y^2 dF_k(\xi_k + by) \tag{23}$$

Now, $\boldsymbol{s} = \boldsymbol{s}'$ unless $\boldsymbol{x}_k \neq \boldsymbol{x}'_k$ for at least one k. The probability of this latter case is $1 - \mathcal{P}\{\boldsymbol{x}_1 = \boldsymbol{x}'_1, \ldots, \boldsymbol{x}_k = \boldsymbol{x}'_k\} = 1 - \Pi\mathcal{P}\{\boldsymbol{x}_k = \boldsymbol{x}'_k\}$ (as the $\boldsymbol{x}_k$ are independent) $= 1 - \Pi\mathcal{P}\{|\boldsymbol{x}_k - \xi_k| \leqslant b\} = 1 - \Pi(1 - p_k) \leqslant$ $\leqslant \Sigma\,p_k$. Hence, for any $\alpha > 0$:

$$R(\alpha\,b) \leqslant \mathcal{P}\{|\,\boldsymbol{s} - c\,| > ab\} \leqslant$$
$$\leqslant \mathcal{P}\{|\,\boldsymbol{s}' - c\,| > ab\} + \Sigma p_k \leqslant (ab)^{-2}\beta'^2 + \Sigma p_k$$

by Bienaymé's inequality, and because of $\mathcal{E}\,\boldsymbol{s}' = c$ by (18), (19), (22),

$$= (ab)^{-2}\,\Sigma\,\sigma_k'^2 + \Sigma\,p_k$$

as $\beta'^2 = \Sigma^k \sigma_k'^2$ because the $\boldsymbol{x}_k'$ like the $\boldsymbol{x}_k$ are independent

$$= \Sigma^k(\int_{|y|>1} dF_k(\xi_k + by) + a^{-2}\int_{|y|\leqslant 1} y^2 dF_k(\xi_k + by))$$

by (15) and (23)

$$\leqslant \max(1, a^2)\, \Sigma^k \gamma_k(\xi_k, b),$$

i.e.

$$\mathcal{P}\{|\, \boldsymbol{s} - c \,| > ab\} \leqslant \max(1, a^{-2})\, \Sigma(b) \tag{24}$$

This proves the sufficiency of the condition $\Sigma_k = \Sigma^k \gamma_k(b) \to 0$ for the LLN in its *original* form (p. 1). (The argument is practically the same as in Freudenthal's proof). Moreover:

$$\begin{aligned} \Gamma(b) &\leqslant \Gamma(c, b) = \int \min(1, z^2)\, dG(c + bz) = \\ &= \int_0^1 \mathcal{P}\{|\, \boldsymbol{s} - c \,| > a\, b\} d(a^2) \leqslant && \text{by } 5 \\ &\leqslant \int_0^1 d(a^2) \min(1, \Sigma(b) a^{-2}) = && \text{by (24)} \\ &= \int_{a^2 \leqslant \Sigma(b)} d(a^2) + \int_{\Sigma(b) \leqslant a^2 \leqslant 1} \Sigma(b)\, a^{-2} da^2 = && \text{if } \Sigma(b) \leqslant 1 \\ &= \Sigma(b) - \Sigma(b) \ln \Sigma(b). \end{aligned}$$

This proves the second inequality in (12'), hence the *sufficiency* of the condition $\Sigma(b) \to 0$ for $\Gamma(b) \to 0$.

In order to prove its necessity, following partly Feller's, partly Freudenthal's lines, we define

$$\boldsymbol{r}_k \overset{\text{def}}{=} \Sigma_{j \neq k}\, \boldsymbol{x}_j = \boldsymbol{s} - \boldsymbol{x}_k \tag{25}$$

and call $H_k(x)$ the distribution function of $\boldsymbol{r}_k$. Then we have for each $k \in \mathrm{N}_n$, if η is a value with $\Gamma(\eta, b) = \Gamma(b)$, then

$$\begin{aligned} \Gamma(b) &= \int_{a^2 \leqslant 1} \mathcal{P}\{|\, \boldsymbol{x}_k + \boldsymbol{r}_k - \eta \,| > ab\} d(a^2) = \\ &= \int_{a^2 \leqslant 1} d(a^2) \int d\, H_k(r) \mathcal{P}\{|\, \boldsymbol{x}_k + r - \eta \,| > ab\} = && (26) \\ &= \int dH_k(r) \gamma_k(\eta - r, b) \geqslant \gamma_k(b) && \text{by } 5 \text{ and (3).} \end{aligned}$$

Hence, by *16*,

$$\pi\Gamma(b) \geqslant \pi\, \gamma_k(b) \geqslant |\int (1 - e^{ity}) dF_k(\xi_k + by)| \tag{27}$$

for any vector with $|\, t \,| \leqslant 2$. If $\Gamma(b) \leqslant (6\pi)^{-1}$, the last member of (27) is $\leqslant \frac{1}{6}$ for each $k \in \mathrm{N}_n$. Applying *16* again, but now to $\boldsymbol{s}$ instead of $\boldsymbol{x}$, and still assuming $|\, t \,| \leqslant 2$, we have

$$\begin{aligned} \pi\Gamma(b) &\geqslant |\int (1 - e^{itz}) dG(\eta + bz)| = |1 - \mathcal{E}\, e^{itb^{-1}(\Sigma|\boldsymbol{x}_k| - \eta)}| = \\ &= |\, 1 - e^{itb^{-1}(c - \eta)} \prod \mathcal{E} e^{itb^{-1}(\boldsymbol{x}_k - \xi_k)} \,| \geqslant 1 - \prod |\mathcal{E} e^{it\boldsymbol{y}_k}| \geqslant && (28) \\ &\geqslant \tfrac{1}{2} \Sigma^k (1 - |\, \mathcal{E}\, e^{it\boldsymbol{y}_k}|) \geqslant \\ &\geqslant \tfrac{1}{4} \Sigma^k \operatorname{var}\{\sin t\, \boldsymbol{y}_k\}, \end{aligned}$$

as

$$|\mathcal{E}\,\mathrm{e}^{ity_k}|^2 = (\int \cos ty\, dF_k(\xi_k + by))^2 + (\int \sin ty\, dF_k(\xi_k + by))^2 \leqslant$$
$$\leqslant \int \cos^2 ty\, dF_k(\xi_k + by) + (\mathcal{E} \sin ty_k)^2 =$$
$$= 1 - (\mathcal{E} \sin^2 ty_k - (\mathcal{E} \sin ty_k)^2) =$$
$$= 1 - \mathrm{var}\,\{\sin ty_k\},$$

whence

$$1 - |\mathcal{E}\,\mathrm{e}^{ity_k}| \geqslant (1 + |\mathcal{E}\,\mathrm{e}^{ity_k}|)^{-1} \mathrm{var}\{\sin ty_k\} \geqslant \tfrac{1}{2}\, \mathrm{var}\{\sin ty_k\}$$

As (28) holds for all vectors t with $|t| \leqslant 2$, we have by *15* and $\pi - 2 < \frac{6}{5}$;

$$4\pi\Gamma(b) \geqslant \max_{|t| \leqslant 2} \Sigma^k\, \mathrm{var}\,\{\sin ty_k\} \geqslant$$
$$\geqslant \max_{|t| \leqslant 2} \Sigma^k\, \{\int \sin^2 ty\, dF_k(\xi_k + by) - \tfrac{36}{25} (\gamma_k(b))^2\} \geqslant$$
$$\geqslant (\int_{|t| \leqslant T} dt)^{-1} \int_{|t| \leqslant T} dt\, \Sigma^k \{\int \sin^2 ty\, dF_k(\xi_k + by) - \tfrac{36}{25} (\gamma_k(b))^2\}$$

if $0 < T \leqslant 2$,

$$= \Sigma^k\, (\int_{|t| \leqslant T} dt)^{-1} \int_{|t| \leqslant T} dt \int \sin^2 ty\, dF_k(\xi_k + by) - \tfrac{36}{25}\, \Sigma^k\, (\gamma_k(b))^2 \geqslant$$
$$\geqslant \frac{3}{2(m+2)}\, \Sigma^k \left(1 - \frac{1}{2T}\right) \gamma_k(b) - \tfrac{36}{25}\, \Gamma(b)\, \Sigma^k\, \gamma_k(b)$$

by (26) and *17*, if we restrict T to $T \leqslant \dfrac{\pi}{2}$, and $\Gamma(b)$ to $2\pi(m+2)\Gamma(b) \leqslant 1$,

$$= \frac{3}{(m+2)} \left(\tfrac{1}{2} - \frac{1}{4T} - \tfrac{36}{25}\frac{1}{6\pi}\right) \Sigma\,(b) > \frac{3}{4(m+2)}\, \Sigma\,(b)$$

if we take $2T = 3 < \pi$. Hence $16\pi(m+2)\, \Gamma(b) > 3\, \Sigma\,(b)$, which proves the first inequality in (12') and completes the proof of theorem 2'. [7])

ADDITIONAL NOTE

The theorem proved in the text can be brought into another form by introducing an appropriate metric for stochastic variables and their sums.

We remind that $\varrho(x, y)$ defines a *distance* between two elements x, y of a given set E, if $\varrho(x, y) \geqslant 0$, $\varrho(x, y) = 0 \rightleftarrows x = y$, $\varrho(x, y) + \varrho(y, z) \geqslant \varrho(z, x)$, whence $\varrho(x, y) = \varrho(y, x)$. Moreover, it is well known that, if $\varrho(x, y)$ is a distance on E, then also the "truncated" value

$$\varrho_b(x, y) \overset{\mathrm{def}}{=} \min(\varrho(x, y), b)$$

[7]) I wish to thank Dr W. Peremans and Dr H. J. A. Duparc for reading the MS and suggesting some improvements of the text.

is, where b is any positive number. It determines the same topology as $\varrho(x, y)$.

We apply this to the case where x and y are real numbers, or, more generally, vectors in a Euclidean space R_m. Instead of the distances it is then sufficient to consider the norms, i.e. distances from zero. We define the "truncated norm"

$$| x|_b \stackrel{\text{def}}{=} \min (| x |, b)$$

where $| x |$ in the first case denotes the absolute value, in the second one the length of the vector x. The corresponding truncated distance $\varrho_b(x, y)$ is then the truncated norm of their difference:

$$\varrho_b(x, y) = | x - y |_b.$$

Moreover, if $\boldsymbol{x}$ is a stochastic quantity (vector) we can, following the ideas developed by M. FRÉCHET, define its truncated norm $|| x ||_b$ e.g. by the generalized Pythagorean addition (integration) by

$$|| \boldsymbol{x} ||_b \stackrel{\text{def}}{=} (\mathcal{E} | \boldsymbol{x} |_b^2)^{1/2}$$

The corresponding truncated distance will be denoted by the symbol σ_b:

$$\sigma_b(\boldsymbol{x}, \boldsymbol{y}) \stackrel{\text{def}}{=} || \boldsymbol{x} - \boldsymbol{y} ||_b = (\mathcal{E}\, \sigma_b(\boldsymbol{x}, \boldsymbol{y})^2)^{1/2}.$$

(Of course instead of $\sigma(x, y) = | x - y |$ any other distance could be taken, and instead of the root mean square any other function preserving the triangle inequality).

With these definitions the quantity $\gamma(a, b)$ defined by (2) is the square of the truncated distance between the stochastic vector $\boldsymbol{x}$ and the constant one a, and generally

$$\gamma_{\nu k}(a, b_\nu) = \sigma_{b_\nu}(\boldsymbol{x}_{\nu k}, a)^2.$$

Moreover, defining as customary, the distance between an element x of a set E and one of its subsets S as the infimum of the distances between x and all elements of S, we have, if R_m is considered as the set of all "univalued" (= "constant") stochastic vectors:

$$\gamma_{\nu k}(b_\nu) = \sigma_{b_\nu}(\boldsymbol{x}_{\nu k}, R_m)^2.$$

The quantities $\sigma_b(\boldsymbol{x}, R_m)$ can be considered as norms mod R_m and will be denoted shortly by

$$||| \boldsymbol{x} |||_b \stackrel{\text{def}}{=} \sigma_b(\boldsymbol{x}, R_m).$$

Then

$$||| \boldsymbol{x}_{\nu k} |||_{b_\nu} = \gamma_{\nu k}(b_\nu)^{1/2}.$$

In the same way we have

$$||| s_\nu |||_{b\nu} = \Gamma_\nu(b_\nu)^{1/2}$$

where $s_\nu = x_{\nu,1} + \ldots + x_{\nu,n_\nu}$ like before.

Then Lemma 1 states that LLN holds if and only if

$$\lim_{\nu \to \infty} ||| s_\nu |||_{b\nu} = 0$$

i.e. if the truncated norms modulo R_m of the sum s_ν tend to zero.

Finally theorem 1 states that this is the case if and only if

$$\lim_{\nu \to \infty} (\Sigma^k ||| x_{\nu k} |||_{b\nu})^{1/2} = 0.$$

Hence theorem 1 states that the topology determined by the truncated norms modulo R_m is equivalent with the one, obtained by taking as the norm of a sum of *independent* stochastic vectors the Pythagorean sum of the truncated norms mod R_m of the summands.

REFERENCES

[1] W. Feller. Über das Gesetz der grossen Zahlen, Acta Szeged *8*, 191—201, 1937.
[2] P. Lévy. Propriétés asymptotiques des sommes de variables aléatoires indépendantes ou enchaînées, Jn. de Math. p. et appl. *40*, 347—402, 1935.
[3] H. Freudenthal. A limit free formulation of the weak Law of Large Numbers, Proc. Kon. Ned. Ak. *55*, 427—432, 1952, or Indagationes Math. *14*, 4, 1952.
[4] H. Cramér and H. Wold. Some theorems on distribution functions, Jn. London M. Soc. *11*, 290—294, 1936.
[5] M. Loève. Fundamental limit theorems of probability theory, Ann. M. Stat. *21*, 321—338, 1950.

(Received 25.1.'53; Additional note: 11.4.'53)

8252