Seminar on Probability and Statistics

University of Potsdam, 2019

Non–linear functionals preserving normal distribution and their asymptotic normality

Linda Khachatryan

Institute of Mathematics

National Academy of Sciences of Armenia

Abstract

Limit theorems of probability theory for sums of random variables play important role in many applied problems, in particular, in problems of mathematical statistical physics:
∘ questions of equivalence of ensembles,
∘ asymptotic formulas for calculating the values of thermodynamic functions,
∘ asymptotic behavior of total spins, etc.
This confirms the urgency of the problem of expanding the range of applicability of such theorems.

Limit theorems for sums of independent random variables form a complete theory which presents one of the main parts of the probability theory. Two main directions for further researches:
∘ the validity of the CLT for sums of dependent random variables;
∘ the asymptotic normality of different classes of non–linear functionals (on independent or dependent random variables).

In the talk

• sufficiently wide classes of non-linear functionals preserving Gaussian distribution were introduced

• various conditions under which a sequence of such functionals is asymptotically normal were established

◦ a generalization and sharpening of known results on the CLT for weighted sums (linear functionals) of independent random variables is obtained

Khachatryan L.A., Nahapetian B.S., *Non–linear functionals preserving normal distribution and their asymptotic normality*, Armenian Journal of Mathematics, 2018

Normal distribution

Random variable ξ has a *normal (Gaussian) distribution* $\mathscr{N}(a, \sigma^2)$ if

$$P(\xi < z) = \frac{1}{\sigma\sqrt{2\pi}} \int\limits_{-\infty}^{z} \exp\left\{-\frac{(x-a)^2}{2\sigma^2}\right\} dx, \qquad z \in \mathbb{R}$$

Random variable ξ has the *standard normal (Gaussian) distribution* $\mathscr{N}(0,1)$

$$P(\xi < z) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{z} e^{-x^2/2} dx, \qquad z \in \mathbb{R}$$

Characteristic function

$$Ee^{it\xi} = e^{-t^2/2}$$

Stability of the normal distribution

Let $\xi_1, \xi_2, ..., \xi_n$ be independent random variables with normal distributions

$$\xi_j \sim \mathcal{N}(a_j, \sigma_j^2), \qquad j = 1, 2, ..., n$$

Put

$$S_n = \sum_{j=1}^{n} \xi_j$$

Then

$$\frac{S_n - ES_n}{\sqrt{DS_n}} \sim \mathcal{N}(0, 1)$$

$$P\left(\frac{S_n - ES_n}{\sqrt{DS_n}} < z\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-x^2/2} dx, \qquad z \in \mathbb{Z}^d,\ n \geq 1$$

Asymptotic normality

Theorem (Central Limit Theorem)**.** *Let* $\eta_1, \eta_2, ..., \eta_n$ *be independent identically distributed random variables,* $0 < D\eta_j < \infty$, $j = 1, 2, ..., n$. *Then*

$$\lim_{n\to\infty} P\left(\frac{S_n - ES_n}{\sqrt{DS_n}} < z\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-x^2/2} dx, \qquad z \in \mathbb{Z}^d$$

where $S_n = \sum_{j=1}^{n} \eta_j$.

Theorem (Levy–Lindeberg). *For any* $n \geq 1$ *let* $\eta_{n,1}, \eta_{n,2}, \ldots, \eta_{n,k(n)}$ *be a sequence of independent random variables* ($k(n) \to \infty$ *as* $n \to \infty$) *such that*

$$\sum_{j=1}^{k(n)} E\eta_{n,j} = 0; \qquad \sum_{j=1}^{k(n)} D\eta_{n,j} = 1;$$

and for any $\varepsilon > 0$

$$\sum_{j=1}^{k(n)} E\left(\eta_{n,j}^2 I(|\eta_{n,j}| > \varepsilon)\right) \to 0 \text{ as } n \to \infty.$$

Then for $S_n = \sum_{j=1}^{k(n)} \eta_{n,j}$ *the CLT is valid, i.e.*

$$\lim_{n\to\infty} P\left(\frac{S_n}{\sqrt{DS_n}} < z\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-x^2/2} dx, \qquad z \in \mathbb{Z}^d.$$

1. Functionals preserving the normal distribution

Denote by $f_n = f_n(x_1, x_2, ..., x_n)$ the functional

$$f_n : \mathbb{R}^n \to \mathbb{R}$$

Functional f_n *preserves the normal distribution* if for independent random variables $\xi_1, \xi_2, ..., \xi_n \sim \mathcal{N}(0, 1)$ one has

$$f_n(\xi_1, \xi_2, ..., \xi_n) \sim \mathcal{N}(0, 1),$$

i.e. for any $z \in \mathbb{R}$

$$P(f_n(\xi_1, \xi_2, ..., \xi_n) < z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-x^2/2} dx$$

Examples of functionals preserving the normal distribution

Example 1

$$f_2(x_1, x_2) = \frac{x_1 + x_2}{\sqrt{2}}$$

Example 2

$$f_2(x_1, x_2) = \frac{\sqrt{2} \cdot x_1 x_2}{\sqrt{x_1^2 + x_2^2}}$$

Example 3

$$f_3(x_1, x_2, x_3) = \frac{x_1 x_2 + x_3}{\sqrt{1 + x_1^2}}$$

Superposition of functionals preserving the normal distribution

Any superposition of functionals preserving the normal distribution is again a functional which preserves the normal distribution.

Example 1'

$$f_n(x_1, x_2, ..., x_n) = f_2\left(x_1, f_{n-1}(x_2, ..., x_n)\right) =$$

$$= \frac{x_1}{\sqrt{2}} + \frac{x_2}{(\sqrt{2})^2} + \frac{x_3}{(\sqrt{2})^3} + ... + \frac{x_{n-2}}{(\sqrt{2})^{n-2}} + \frac{x_{n-1} + x_n}{(\sqrt{2})^{n-1}}, \qquad n > 2$$

Example 2'

$$f_n(x_1, x_2, ..., x_n) = \frac{\sqrt{n} \cdot \prod_{j=1}^{n} x_j}{\sqrt{\sum_{j=1}^{n} \prod_{1 \le i \le n \,:\, i \ne j} x_i^2}}, \qquad n > 2$$

Example 3'

$$f_n(x_1, x_2, ..., x_n) = f_3(x_1, x_2, f_{n-1}(x_1, x_3, x_4, ..., x_n)) =$$

$$= \frac{x_1 x_2}{\sqrt{1+x_1^2}} + \frac{x_1 x_3}{1+x_1^2} + \frac{x_1 x_4}{(1+x_1^2)^{3/2}} + ... +$$

$$+ \frac{x_1 x_{n-1}}{(1+x_1^2)^{(n-2)/2}} + \frac{x_n}{(1+x_1^2)^{(n-2)/2}},$$

$n > 3$

There are other ways to construct sequences of functionals that preserve the normal distribution.

Linear functionals
(weighted sums)

Linear functional with coefficients $\alpha_j^{(n)}$, $j = \overline{1, n}$, has the form

$$f_n(x_1, x_2, ..., x_n) = \sum_{j=1}^{n} \alpha_j^{(n)} x_j.$$

Condition on linear functionals to preserve the normal distribution

Proposition. *The linear functional* f_n *with coefficients* $\alpha_j^{(n)}$*,* $j = \overline{1, n}$*, preserves the normal distribution if and only if*

$$\sum_{j=1}^{n} \left(\alpha_j^{(n)}\right)^2 = 1$$

Proof. Let $\xi_1, \xi_2, ..., \xi_n \sim \mathcal{N}(0,1)$ be independent random variables. Characteristic function φ_n of the random variable $\sum_{j=1}^{n} \alpha_j^{(n)} \xi_j$ has the form

$$\varphi_n(t) = E \exp\left\{ it \sum_{j=1}^{n} \alpha_j^{(n)} \xi_j \right\} = \prod_{j=1}^{n} E e^{it\alpha_j^{(n)} \xi_j}.$$

Since $\alpha_j^{(n)} \xi_j \sim \mathcal{N}\left(0, \left(\alpha_j^{(n)}\right)^2\right)$, we have

$$E \exp\left\{ it\alpha_j^{(n)} \xi_j \right\} = \exp\left\{ -\frac{1}{2} \left(\alpha_j^{(n)} t\right)^2 \right\}.$$

Then

$$\varphi_n(t) = \prod_{j=1}^{n} e^{-(\alpha_j^{(n)} t)^2/2} = \exp\left\{ -\frac{t^2}{2} \sum_{j=1}^{n} \left(\alpha_j^{(n)}\right)^2 \right\}.$$

Hence, $\varphi_n(t) = e^{-t^2/2}$ if and only if $\sum_{j=1}^{n} \left(\alpha_j^{(n)}\right)^2 = 1$. $\square$

Methods for constructing linear functionals preserving the normal distribution

Proposition. *Consider the functional*

$$f_2(x_1, x_2) = \alpha x_1 + \beta x_2, \qquad \alpha, \beta \in \mathbb{R} \backslash \{0, 1\},$$

and the superposition

$$f_n(x_1, x_2, ..., x_n) = f_2\left(x_1, f_{n-1}(x_2, x_3, ..., x_n)\right), \qquad n > 2. \quad (1)$$

1. For any $n \geq 2$, f_n is the linear functional with coefficients $\alpha_j^{(n)}$ of the form

$$\alpha_n^{(n)} = \beta^{n-1}, \qquad \alpha_j^{(n)} = \alpha\beta^{j-1}, \quad 1 \leq j < n-1. \qquad (2)$$

Conversely, any linear functional f_n with coefficients given by (2) *can be represented as the superposition* (1) *where f_2 has coefficients α and β.*

2. The linear functional (1) *preserves the normal distribution if and only if the coefficients α and β of the functional f_2 are such that $\alpha^2 + \beta^2 = 1$.*

Proposition. *Consider the functional*

$$f_2(x_1, x_2) = \alpha x_1 + \beta x_2, \qquad \alpha, \beta \in \mathbb{R} \backslash \{0, 1\},$$

and the superposition

$$f_{2^n}(x_1, ..., x_{2^n}) = f_2\left(f_{2^{n-1}}(x_1, ..., x_{2^{n-1}}), f_{2^{n-1}}(x_{2^{n-1}+1}, ..., x_{2^n})\right),$$

$n > 1$.

The linear functional f_{2^n} preserves the normal distribution if and only if the coefficients α and β of the functional f_2 are such that $\alpha^2 + \beta^2 = 1$. Moreover,

$$\max_{1 \le j \le n} \left|\alpha_j^{(n)}\right| \to 0 \text{ as } n \to \infty$$

Generalization of Example 3

Example 4. Functionals

$$f_{n+1}(x_0, x_1, x_2, x_3, ..., x_n) = \frac{x_0 x_1 + x_2 + x_3 ... + x_n}{\sqrt{n-1+x_0^2}}, \qquad n \geq 2$$

preserve the normal distribution

One has

$$f_{n+1}(x_0, x_1, x_2, ..., x_n) = \frac{x_0}{\sqrt{n-1+x_0^2}} \cdot x_1 + \sum_{j=2}^{n} \frac{1}{\sqrt{n-1+x_0^2}} \cdot x_j.$$

Note that for any $x_0 \in \mathbb{R}$

$$\left(\frac{x_0}{\sqrt{n-1+x_0^2}}\right)^2 + \sum_{j=2}^{n} \left(\frac{1}{\sqrt{n-1+x_0^2}}\right)^2 = \frac{x_0^2}{n-1+x_0^2} + \frac{n-1}{n-1+x_0^2} = 1.$$

The main class of non–linear functionals preserving the normal distribution

Let $\alpha_j^{(n)}(z)$, $j = \overline{1, n}$, $z \in \mathbb{R}$, be a set of functions. Put

$$f_{n+1}(x_0, x_1, x_2, ..., x_n) = \\ = f_n^{x_0}(x_1, x_2, ..., x_n) = \sum_{j=1}^{n} \alpha_j^{(n)}(x_0) \cdot x_j. \tag{3}$$

Theorem. *The functionals* (3) *preserve the normal distribution if*

$$\sum_{j=1}^{n} \left(\alpha_j^{(n)}(z)\right)^2 = 1 \qquad \text{for any } z \in \mathbb{R}. \tag{4}$$

Proof. Let $\xi_0, \xi_1, \xi_2, ..., \xi_n \sim \mathcal{N}(0,1)$ be independent random variables. Let us show that

$$E\left(\sum_{j=1}^{n} \alpha_j^{(n)}(\xi_0)\xi_j\right)^{2k-1} = 0, \quad E\left(\sum_{j=1}^{n} \alpha_j^{(n)}(\xi_0)\xi_j\right)^{2k} = \frac{(2k)!}{2^k k!},$$

$k = 1, 2, ...$

We have

$$E\left(\sum_{j=1}^{n} \alpha_j^{(n)}(\xi_0)\xi_j\right)^{2k-1} =$$

$$= \sum_{\substack{0 \le m_1, ..., m_n \le 2k-1: \\ m_1 + ... + m_n = 2k-1}} \frac{(2k-1)!}{m_1! m_2! \cdot ... \cdot m_n!} \prod_{i=1}^{n} E\xi_i^{m_i} \cdot E\left(\prod_{i=1}^{n} \left(\alpha_i^{(n)}(\xi_0)\right)^{m_i}\right).$$

Since $m_1 + m_2 + ... + m_n = 2k-1$, for any numbers $0 \le m_1, ..., m_n \le 2k-1$ there exists i, $1 \le i \le n$, such that m_i is odd, and hence $E\xi_i^{m_i} = 0$.

For even moments we have

$$E\left(\sum_{j=1}^{n} \alpha_j^{(n)}(\xi_0)\xi_j\right)^{2k} = \frac{(2k)!}{2^k k!} E\left(\sum_{i=1}^{n}\left(\alpha_i^{(n)}(\xi_0)\right)^2\right)^k.$$

It remains to note that due to the condition (4)

$$E\left(\sum_{i=1}^{n}\left(\alpha_i^{(n)}(\xi_0)\right)^2\right)^k = 1$$

for any k. □

It is not difficult to see that in the proof of the theorem above we do not use any condition on the distribution of ξ_0.

Theorem 1. *Let* $\zeta, \xi_1, \xi_2, ..., \xi_n$ *be independent random variables and let* $\xi_j \sim \mathscr{N}(0,1)$, $j = \overline{1,n}$. *The random variable*

$$f_n^{\zeta}(\xi_1, \xi_2, ..., \xi_n) = \sum_{j=1}^{n} \alpha_j^{(n)}(\zeta)\xi_j$$

has the standard normal distribution if

$$\sum_{j=1}^{n} \left(\alpha_j^{(n)}(z)\right)^2 = 1 \qquad \text{for all } z \in \mathbb{R}.$$

Generalization of Example 4

Example 5. Let $g : \mathbb{R} \to \mathbb{R}$ be a Borel function. Put

$$\alpha_n^{(n)}(z) = \frac{g(z)}{\sqrt{n - 1 + g^2(z)}},$$

$$\alpha_j^{(n)}(z) = \frac{1}{\sqrt{n - 1 + g^2(z)}}, \qquad 1 \le j \le n - 1.$$

Then the functionals

$$f_n^z(x_1, x_2, ..., x_n) = \sum_{j=1}^{n-1} \frac{x_j}{\sqrt{n - 1 + g^2(z)}} + \frac{x_n \cdot g(z)}{\sqrt{n - 1 + g^2(z)}},$$

$n \ge 2$, preserve the normal distribution.

Indeed, for any fixed $n \ge 2$

$$\sum_{j=1}^{n} \left(\alpha_j^{(n)}(z)\right)^2 = \frac{n-1}{n - 1 + g^2(z)} + \frac{g^2(z)}{n - 1 + g^2(z)} = 1.$$

∘ Linnik Yu.V., Eidlin V.L., *Remark on analytic transformations of normal vectors*. Theory of Probability and its Applications 13 (4), 1968

A method for constructing non–linear functionals preserving normal distribution (proposed by Shiryaev and Romanovskiy):

Let $P_j(x_m, ..., x_n)$, $j = 1, 2, ..., m-1$, be polynomials in variables $x_m, ..., x_n$, such that

$$\sum_{j=1}^{m-1} P_j^2(x_m, ..., x_n)$$

is a polynomial in the same variables. Then the functional

$$f_n(x_1, x_2, ..., x_n) = \frac{\sum_{j=1}^{m-1} x_j P_j(x_m, ..., x_n)}{\sqrt{\sum_{j=1}^{m-1} P_j^2(x_m, ..., x_n)}}$$

preserves the normal distribution.

Example 6.

Let g, h be functions in one variable. Put

$$P_1(x_3) = f^2(x_3) - g^2(x_3), \qquad P_2(x_3) = 2f(x_3)g(x_3).$$

Then the functional

$$f_3(x_1, x_2, x_3) = \frac{x_1 P_1(x_3) + x_2 P_2(x_3)}{\sqrt{P_1^2(x_3) + P_2^2(x_3)}}$$

$$= \frac{x_1 \cdot (g^2(x_3) - h^2(x_3)) + x_2 \cdot 2g(x_3)h(x_3)}{g^2(x_3) + h^2(x_3)}$$

preserves the normal distribution.

2. Asymptotic normality of sequences of functionals preserving the normal distribution

Let $\eta_1, \eta_2, ..., \eta_n$ be independent random variables (not necessarily Gaussian).

A sequence of functionals $f_n(\eta_1, \eta_2, ..., \eta_n)$, $n \geq 2$, is *asymptotically normal* if their distributions converge to the standard normal one as $n \to \infty$: for any $z \in \mathbb{R}$

$$\lim_{n\to\infty} P(f_n(\eta_1, \eta_2, ..., \eta_n) < z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-x^2/2} dx$$

Theorem 2. *Let random variable* ζ *and functions* $\alpha_j^{(n)}$, $j=\overline{1,n}$, *be such that*

$$\sum_{j=1}^{n} E\left|\alpha_j^{(n)}(\zeta)-\frac{1}{\sqrt{n}}\right| \to 0 \qquad \textit{as } n\to\infty. \tag{5}$$

Let $\eta_1,\eta_2,...,\eta_n$ *be independent random variables with finite second moments which are independent of* ζ *too and for which the CLT holds. Then the sequence of functionals*

$$f_n^{\zeta}(\eta_1,\eta_2,...,\eta_n)=\sum_{j=1}^{n}\alpha_j^{(n)}(\zeta)\eta_j, \qquad n\geq 2,$$

is asymptotically normal.

Proof. For any $n \geq 2$ we can write

$$f_n^\zeta(\eta_1, \eta_2, ..., \eta_n) = \sum_{j=1}^{n} \frac{\eta_j}{\sqrt{n}} + \sum_{j=1}^{n} \left(\alpha_j^{(n)}(\zeta) - \frac{1}{\sqrt{n}} \right) \eta_j.$$

The first summand is asymptotically normal, the second one tends to 0 in probability.

Indeed,

$$E|\eta_j| \leq \left(E\eta_j^2 \right)^{1/2} = C < \infty, \qquad j = \overline{1, n}.$$

Due to the Chebychev inequality for any $\varepsilon > 0$ we have

$$P\left(\left| \sum_{j=1}^{n} \left(\alpha_j^{(n)}(\zeta) - \frac{1}{\sqrt{n}} \right) \eta_j \right| > \varepsilon \right) \leq \frac{1}{\varepsilon} E \left| \sum_{j=1}^{n} \left(\alpha_j^{(n)}(\zeta) - \frac{1}{\sqrt{n}} \right) \eta_j \right| \leq$$

$$\leq \frac{1}{\varepsilon} \sum_{j=1}^{n} E \left| \alpha_j^{(n)}(\zeta) - \frac{1}{\sqrt{n}} \right| E|\eta_i| \leq \frac{C}{\varepsilon} \sum_{j=1}^{n} E \left| \alpha_j^{(n)}(\zeta) - \frac{1}{\sqrt{n}} \right| \to 0$$

as $n \to \infty$. □

Example 5. The functionals

$$f_n^z(x_1, x_2, ..., x_n) = \sum_{j=1}^{n-1} \frac{x_j}{\sqrt{n-1+g^2(z)}} + \frac{x_n \cdot g(z)}{\sqrt{n-1+g^2(z)}},$$

$n \geq 2$, are asymptotically normal.

Indeed, let the function g be such that $Eg^2(\zeta) < \infty$. Then

$$\sum_{j=1}^{n} E\left|\alpha_j^{(n)}(\zeta) - \frac{1}{\sqrt{n}}\right| =$$

$$= E\left|\frac{g(\zeta)}{\sqrt{n-1+g^2(\zeta)}} - \frac{1}{\sqrt{n}}\right| + (n-1)E\left|\frac{1}{\sqrt{n-1+g^2(\zeta)}} - \frac{1}{\sqrt{n}}\right| \leq$$

$$\leq \frac{E|g(\zeta)|}{\sqrt{n-1}} + \frac{1}{\sqrt{n}} + \frac{n-1}{n} \cdot \frac{E|1-g^2(\zeta)|}{\sqrt{n-1}} \to 0$$

as $n \to \infty$.

Theorem 3. *Let functions* $\alpha_j^{(n)}$, $j = \overline{1, n}$, *satisfy the condition*

$$\sum_{j=1}^{n} \left(\alpha_j^{(n)}(z)\right)^2 = 1 \qquad \textit{for all } z \in \mathbb{R},$$

and let random variable ζ *be such that*

$$\max_{1 \leq j \leq n} E \left|\alpha_j^{(n)}(\zeta)\right| \to 0 \qquad \textit{as } n \to 0$$

and for some δ, $0 < \delta < 1/2$,

$$\max_{n \geq 1} \sum_{j=1}^{n} E^{\frac{1}{2(1+\delta)}} \left(\alpha_j^{(n)}(\zeta)\right)^2 \leq C < \infty$$

Let $\eta_1, \eta_2, ..., \eta_n$ *be independent random variables which are independent of* ζ *too and are such that* $E\eta_j = 0$, $D\eta_j = 1$, $j = \overline{1, n}$. *Then the sequence of functionals*

$$f_n^{\zeta}(\eta_1, \eta_2, ..., \eta_n) = \sum_{j=1}^{n} \alpha_j^{(n)}(\zeta)\eta_j, \qquad n \geq 2,$$

is asymptotically normal.

Proof. Let $\xi_1, \xi_2, ..., \xi_n$ be independent standard normally distributed random variables which are also independent of $\zeta, \eta_1, \eta_2, ..., \eta_n$. We have

$$f_n^{\zeta}(\eta_1, \eta_2, ..., \eta_n) = \sum_{j=1}^{n} \alpha_j^{(n)}(\zeta)\xi_j + \sum_{j=1}^{n} \alpha_j^{(n)}(\zeta)(\eta_j - \xi_j).$$

Since the functionals f_n^{ζ} preserve the normal distribution

$$\sum_{j=1}^{n} \alpha_j^{(n)}(\zeta)\xi_j \sim \mathcal{N}(0,1)$$

Applying the Chebyshev inequality, one can show that

$$P\left(\left|\sum_{j=1}^{n} \alpha_j^{(n)}(\zeta)(\eta_j - \xi_j)\right| > \varepsilon\right) \leq \left(\frac{C\sqrt{2}}{\varepsilon}\right)^{1+\delta} \left(\max_{1\leq j\leq n} E\left|\alpha_j^{(n)}(\zeta)\right|\right)^{\delta} \to 0$$

as $n \to \infty$. □

Lemma. *Let functions* $\alpha_j^{(n)}$, $j=\overline{1,n}$, *be such that*

$$\sum_{j=1}^{n}\left(\alpha_j^{(n)}(z)\right)^2=1 \qquad \text{for all } z\in\mathbb{R},$$

for independent random variables $\zeta,\xi_1,...,\xi_n$, *where* $\xi_j\sim\mathcal{N}(0,1)$, *and any* $\varepsilon>0$,

$$\sum_{j=1}^{n}E\left(\left(\alpha_j^{(n)}(\zeta)\xi_j\right)^2 I\left(\left|\alpha_j^{(n)}(\zeta)\xi_j\right|>\varepsilon\right)\right)\to 0 \qquad \text{as } n\to\infty.$$

Then for any independent random variables $\eta_1,...,\eta_n$ *which are independent of* ζ *and such that* $E\eta_j=0$, $D\eta_j=1$, $j=\overline{1,n}$, *and for any* $\varepsilon>0$

$$\sum_{j=1}^{n}E\left(\left(\alpha_j^{(n)}(\zeta)\eta_j\right)^2 I\left(\left|\alpha_j^{(n)}(\zeta)\eta_j\right|>\varepsilon\right)\right)\to 0 \qquad \text{as } n\to\infty,$$

the sequence of functionals

$$f_n^{\zeta}(\eta_1,\eta_2,...,\eta_n)=\sum_{j=1}^{n}\alpha_j^{(n)}(\zeta)\eta_j, \qquad n\geq 2,$$

is asymptotically normal.

Proof. Since the functional f_n^{ζ} preserves the normal distribution, one has

$$E \exp\left\{it \sum_{j=1}^{n} \alpha_j^{(n)}(\zeta)\xi_j\right\} = e^{-t^2/2}, \quad n > 1,$$

where $\xi_1, ..., \xi_n$ are independent random variables with normal distribution which are independent of $\eta_1, ..., \eta_n$ as well.

Hence

$$\left| E \exp\left\{it \sum_{j=1}^{n} \alpha_j^{(n)}(\zeta)\eta_j\right\} - e^{-t^2/2}\right| \to 0$$

$$\Updownarrow$$

$$\left| E \exp\left\{it \sum_{j=1}^{n} \alpha_j^{(n)}(\zeta)\eta_j\right\} - E \exp\left\{it \sum_{j=1}^{n} \alpha_j^{(n)}(\zeta)\xi_j\right\}\right| \to 0$$

as $n \to \infty$

Put

$$\zeta_1^{(n)} = \sum_{j=2}^{n} \alpha_j^{(n)}(\zeta)\xi_j, \qquad \zeta_n^{(n)} = \sum_{j=1}^{n-1} \alpha_j^{(n)}(\zeta)\eta_j,$$

and for any k, $1 < k < n$,

$$\zeta_k^{(n)} = f_n^{\zeta}(\eta_1, ..., \eta_{k-1}, 0, \xi_{k+1}, ..., \xi_n) = \sum_{j=1}^{k-1} \alpha_j^{(n)}(\zeta)\eta_j + \sum_{j=k+1}^{n} \alpha_j^{(n)}(\zeta)\xi_j.$$

Then

$$\left| E \exp\left\{ it \sum_{j=1}^{n} \alpha_j^{(n)}(\zeta)\eta_j \right\} - E \exp\left\{ it \sum_{j=1}^{n} \alpha_j^{(n)}(\zeta)\xi_j \right\} \right| =$$

$$= \left| \sum_{k=1}^{n} \left(E \exp\left\{ it(\zeta_k^{(n)} + \alpha_k^{(n)}(\zeta)\eta_k) \right\} - E \exp\left\{ it(\zeta_k^{(n)} + \alpha_k^{(n)}(\zeta)\xi_k) \right\} \right) \right| \leq$$

$$\leq \sum_{k=1}^{n} \left| E e^{it\zeta_k^{(n)}} \cdot e^{it\alpha_k^{(n)}(\zeta)\eta_k} - E e^{it\zeta_k^{(n)}} \cdot e^{it\alpha_k^{(n)}(\zeta)\xi_k} \right|.$$

Since for any k

$$Ee^{it\zeta_k^{(n)}}\alpha_k^{(n)}(\zeta)\eta_k = Ee^{it\zeta_k^{(n)}}\alpha_k^{(n)}(\zeta)E\eta_k = 0 = Ee^{it\zeta_k^{(n)}}\alpha_k^{(n)}(\zeta)\xi_k,$$

$$Ee^{it\zeta_k^{(n)}}\left(\alpha_k^{(n)}(\zeta)\eta_k\right)^2 = Ee^{it\zeta_k^{(n)}}\left(\alpha_k^{(n)}(\zeta)\right)^2 E\eta_k^2 = Ee^{it\zeta_k^{(n)}}\left(\alpha_k^{(n)}(\zeta)\xi_k\right)^2,$$

we can write

$$\sum_{k=1}^{n}\left|Ee^{it\zeta_k^{(n)}}\cdot e^{it\alpha_k^{(n)}(\zeta)\eta_k} - Ee^{it\zeta_k^{(n)}}\cdot e^{it\alpha_k^{(n)}(\zeta)\xi_k}\right| \leq$$

$$\leq \sum_{k=1}^{n} E\left|e^{it\alpha_k^{(n)}(\zeta)\eta_k} - 1 - it\alpha_k^{(n)}(\zeta)\eta_k - \frac{(it)^2}{2}(\alpha_k^{(n)}(\zeta)\eta_k)^2\right| +$$

$$+ \sum_{k=1}^{n} E\left|e^{it\alpha_k^{(n)}(\zeta)\xi_k} - 1 - it\alpha_k^{(n)}(\zeta)\xi_k - \frac{(it)^2}{2}(\alpha_k^{(n)}(\zeta)\xi_k)^2\right|.$$

Using the well-known inequality

$$\left| e^{it} - \sum_{k=0}^{N} \frac{(it)^k}{k!} \right| \le \min \left\{ \frac{2\,|t|^N}{N!}, \frac{|t|^{N+1}}{(N+1)!} \right\}, \qquad N = 0, 1, 2, ...$$

for any $\varepsilon > 0$ we obtain

$$\sum_{k=1}^{n} E \left| e^{it\alpha_k^{(n)}(\zeta)\eta_k} - 1 - it\alpha_k^{(n)}(\zeta)\eta_k - \frac{(it)^2}{2}(\alpha_k^{(n)}(\zeta)\eta_k)^2 \right| \le$$

$$\le \frac{|t|^3}{6}\varepsilon + t^2 \sum_{k=1}^{n} E \left(\left(\alpha_k^{(n)}(\zeta)\eta_k \right)^2 I \left(\left| \alpha_k^{(n)}(\zeta)\eta_k \right| > \varepsilon \right) \right).$$

Similarly,

$$\sum_{k=1}^{n} E \left| e^{it\alpha_k^{(n)}(\zeta)\xi_k} - 1 - it\alpha_k^{(n)}(\zeta)\xi_k - \frac{(it)^2}{2}(\alpha_k^{(n)}(\zeta)\xi_k)^2 \right| \le$$

$$\le \frac{|t|^3}{6}\varepsilon + t^2 \sum_{k=1}^{n} E \left(\left(\alpha_k^{(n)}(\zeta)\xi_k \right)^2 I \left(\left| \alpha_k^{(n)}(\zeta)\xi_k \right| > \varepsilon \right) \right).$$

Finally,

$$\left| E \exp\left\{ it \sum_{j=1}^{n} \alpha_j^{(n)}(\zeta)\eta_j \right\} - e^{-t^2/2} \right| \leq \frac{|t|^3}{3}\varepsilon +$$

$$+ t^2 \sum_{k=1}^{n} E\left(\left(\alpha_k^{(n)}(\zeta)\eta_k \right)^2 I\left(\left| \alpha_k^{(n)}(\zeta)\eta_k \right| > \varepsilon \right) \right) +$$

$$+ t^2 \sum_{k=1}^{n} E\left(\left(\alpha_k^{(n)}(\zeta)\xi_k \right)^2 I\left(\left| \alpha_k^{(n)}(\zeta)\xi_k \right| > \varepsilon \right) \right)$$

$\square$

Theorem 4. *Let functions* $\alpha_j^{(n)}$, $j = \overline{1, n}$, *be such that*

$$\sum_{j=1}^{n} \left(\alpha_j^{(n)}(z)\right)^2 = 1 \qquad \textit{for all } z \in \mathbb{R},$$

and

$$\max_{1 \le j \le n} \sup_{z \in \mathbb{R}} \left|\alpha_j^{(n)}(z)\right| \to 0 \qquad \textit{as } n \to \infty.$$

Then for any independent random variables $\zeta, \eta_1, ..., \eta_n$ *such that* $E\eta_j = 0$, $D\eta_j = 1$, $j = \overline{1, n}$, *and for some* $\delta > 0$

$$\sup_{1 \le j \le n} E\left|\eta_j\right|^{2+\delta} = C_\delta < \infty,$$

the sequence of functionals

$$f_n^{\zeta}(\eta_1, \eta_2, ..., \eta_n) = \sum_{j=1}^{n} \alpha_j^{(n)}(\zeta)\eta_j, \qquad n \ge 2,$$

is asymptotically normal.

Proof. We need to check that conditions of the Lemma are fulfilled. Since for any $\varepsilon > 0$

$$E\left(\left(\alpha_j^{(n)}(\zeta)\eta_j\right)^2 I\left(\left|\alpha_j^{(n)}(\zeta)\eta_j\right| > \varepsilon\right)\right) \le \frac{C_\delta}{\varepsilon^\delta} E\left|\alpha_j^{(n)}(\zeta)\right|^{2+\delta},$$

$1 \le j \le n$, we have

$$\sum_{j=1}^{n} E\left(\left(\alpha_j^{(n)}(\zeta)\eta_j\right)^2 I\left(\left|\alpha_j^{(n)}(\zeta)\eta_j\right| > \varepsilon\right)\right) \le$$

$$\frac{C_\delta}{\varepsilon^\delta}\left(\max_{1\le j\le n}\sup_{z\in\mathbb{R}}\left|\alpha_j^{(n)}(z)\right|\right)^\delta \to 0$$

as $n \to \infty$.

Similarly, for independent standard normally distributed random variables $\xi_1, \xi_2, ..., \xi_n$, which are independent of ζ, we have

$$E\left(\left(\alpha_j^{(n)}(\zeta)\xi_j\right)^2 I\left(\left|\alpha_j^{(n)}(\zeta)\xi_j\right| > \varepsilon\right)\right) \leq \frac{3^{(2+\delta)/4}}{\varepsilon^\delta} E\left|\alpha_j^{(n)}(\zeta)\right|^{2+\delta},$$

and hence

$$\sum_{j=1}^{n} E\left(\left(\alpha_j^{(n)}(\zeta)\xi_j\right)^2 I\left(\left|\alpha_j^{(n)}(\zeta)\xi_j\right| > \varepsilon\right)\right) \leq$$

$$\leq \frac{3^{(2+\delta)/4}}{\varepsilon^\delta} \left(\max_{1\leq j\leq n} \sup_{z\in\mathbb{R}} \left|\alpha_j^{(n)}(z)\right|\right)^\delta \to 0$$

as $n \to \infty$. $\square$

Asymptotic normality of sequences of linear functionals

Conditions under which a sequence of linear functionals (weighted sums)

$$f_n(x_1, x_2, ..., x_n) = \sum_{j=1}^{n} \alpha_j^{(n)} x_j$$

is asymptotically normal were obtained in several works:

∘ Weber M., *A weighted central limit theorem*. Statistics and Probability Letters 76, 2006

• conditions on the 4-th power of the coefficients $\alpha_j^{(n)}$ as well as existence of $E\eta_j^p$ for $p > 4$;

◦ Fisher E., *A Skorohod representation and an invariance principle for sums of weighted i.i.d. random variables*. Rocky Mount. J. Math. 22, 1992

◦ Kevei P., *A note on asymptotics on linear combinations of i.i.d. random variables*. Periodica Mathematica Hungarica 60 (1), 2010

• conditions on the rate of convergence to 0 for the coefficients.

In the mentioned papers only *identically distributed* random variables were considered

Theorem 5. *Let coefficients* $\alpha_j^{(n)}$, $j = \overline{1,n}$, *of the linear functionals* f_n *be such that*

$$\sum_{j=1}^{n} \left(\alpha_j^{(n)}\right)^2 = 1$$

and

$$\max_{1 \le j \le n} \left|\alpha_j^{(n)}\right| \to 0 \quad \text{as } n \to \infty.$$

Let $\eta_1, \eta_2, ..., \eta_n$ *be independent random variables such that* $E\eta_j = 0$, $E\eta_j^2 = 1$, $j = \overline{1,n}$, *and for some* $\delta > 0$

$$\sup_{1 \le j \le n} E\left|\eta_j\right|^{2+\delta} = C_\delta < \infty.$$

Then the sequence of linear functionals $f_n(\eta_1, \eta_2, ..., \eta_n)$, $n \ge 2$, *is asymptotically normal.*

Future research directions

- Construction of other classes of non–linear functionals preserving the normal distribution, and the establishment of conditions for their asymptotic normality;

- Extension of results to functionals in dependent random variables (mixing processes, martingales).

On Lindeberg condition

For any $n \geq 1$ let a sequence of independent random variables $\eta_{n,1}, \eta_{n,2}, ..., \eta_{n,k(n)}$ be given, $k(n) \to \infty$ as $n \to \infty$. The *Lindeberg condition* for this sequence is fulfilled if for any $\varepsilon > 0$

$$\sum_{j=1}^{k(n)} E\left(\eta_{n,j}^2 I(|\eta_{n,j}| > \varepsilon)\right) \to 0 \text{ as } n \to \infty.$$

The Lindeberg condition is a sufficient condition for the validity of the CLT for sums $S_n = \sum_{j=1}^{k(n)} \eta_{n,j}$.

The usual explanation of the meaning of the Lindeberg condition is that this condition guarantees the uniform asymptotic negligibility of random summands. Indeed,

$$P\left(\max_{1\le j\le k(n)} |\eta_{n,j}| > \varepsilon\right) \le \sum_{j=1}^{k(n)} P\left(|\eta_{n,j}| > \varepsilon\right) = \sum_{j=1}^{k(n)} EI\left(|\eta_{n,j}| > \varepsilon\right) \le$$

$$\le \sum_{j=1}^{k(n)} E\left(\frac{\eta_{n,j}^2}{\varepsilon^2} I(|\eta_{n,j}| > \varepsilon)\right) = \frac{1}{\varepsilon^2}\sum_{j=1}^{k(n)} E\left(\eta_{n,j}^2 I(|\eta_{n,j}| > \varepsilon)\right) \to 0$$

as $n \to \infty$.

However, the real meaning of the Lindeberg condition is revealed in the following little-known Khinchin theorem.

Khinchin theorem

Theorem. *For any* $n \geq 1$ *let* $\eta_{n,1}, \eta_{n,2}, ..., \eta_{n,k(n)}$ *be a sequence of independent random variables* ($k(n) \to \infty$ *as* $n \to \infty$) *such that*

$$\sum_{j=1}^{k(n)} P\left(|\eta_{n,j}| > \varepsilon\right) \to 0 \text{ as } n \to \infty.$$

Then

$$\lim_{n\to\infty} P(S_n < z) = G(z), \qquad z \in \mathbb{Z}^d,$$

where $G(z)$ *is a normal distribution function with parameters 0 and* σ^2, $\sigma > 0$.

In this theorem, there are no restrictions on the moments; nevertheless, the limiting distribution is Gaussian. However, this theorem is not a CLT, since the parameter σ is not necessarily equal to one: its value may depend on the considered array of random variables.

From the conditions of the Levy–Lindeberg theorem it follows that

$$DS_n = D\left(\sum_{j=1}^{k(n)} \eta_{n,j}\right) = \sum_{j=1}^{k(n)} D\eta_{n,j} = 1.$$

Hence

$$\lim_{n\to\infty} DS_n = \lim_{n\to\infty} \int\limits_{-\infty}^{\infty} x^2 dP(S_n < x) = 1.$$

If it is possible to *take the limit under the integral sign*, then

$$\lim_{n\to\infty} \int\limits_{-\infty}^{\infty} x^2 dP(S_n < x) = \int\limits_{-\infty}^{\infty} x^2 \lim_{n\to\infty} dP(S_n < x) = \int\limits_{-\infty}^{\infty} x^2 dG(x) = \sigma^2,$$

and hence $\sigma^2 = 1$. Thus, the CLT is valid.

One can take the limit under the integral sign if and only if the corresponding sequence of squares of random variables is uniformly integrable:

$$\int\limits_{S_n^2 > C} S_n^2 P(d\omega) \to 0$$

as $C \to \infty$ uniformly on n. Due to the Billingsley inequality

$$\int\limits_{S_n^2 > C} S_n^2 P(d\omega) \leq K \left(\frac{1}{C} + \sum_{|\eta_{n,j}| \geq \frac{1}{4} C} \eta_{n,j}^2 P(d\omega) \right),$$

where K is a constant.

the Lindeberg condition

$$\Downarrow$$

uniform square integrability of the sequence S_n, $n = 1, 2, ...$

$$\Downarrow$$

the possibility to take the limit under the integral sign

$$\Downarrow$$

$\sigma^2 = 1$ in the Khinchin theorem

Thank you for your attention!

Vielen Dank für Ihre Aufmerksamkeit!