Limit theorems in Probability Theory, Random Matrix Theory and Number Theory

Ashkan Nikeghbali
ashkan.nikeghbali@math.uzh.ch

27/10/2008

1. Outline

- Weak convergence appears in probability theory (convergence in law or distribution), random matrix theory and number theory under non-standard forms.

- Review some of the classical results. How can one prove that there is convergence to the standard Gaussian law?

- Is there a common "probabilistic" framework (a "higher order central limit theorem")? Illustrate the interplay between the three areas.

2. Some classical examples

1. Erdős-Kác. The number of (distinct) prime divisors of a positive integer $n \geq 1$, behaves for large n like a Gaussian random variable with mean $\log\log n$ and variance $\log\log n$:

$$\lim_{N\to+\infty} \frac{1}{N}|\{n \leq N \mid a < \frac{\omega(n) - \log\log N}{\sqrt{\log\log N}} < b\}| = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-t^2/2} dt$$

for any real numbers $a < b$.

2. $U(N)$ is the unitary group endowed with the (probability) Haar measure. Take $A \in U(N)$, define

$$Y_N(z) = det(I - zA)$$

and note $Y_N \equiv Y_N(1)$.

Theorem [Keating-Snaith, 2000]:

$$\lim_{N\to\infty} |\left\{ A \in U(N);\ \frac{\log Y_N}{\sqrt{\frac{1}{2}\log N}} \in B \right\}| = \frac{1}{2\pi} \int\int_B \exp\left(\frac{-\left(x^2+y^2\right)}{2} \right) \mathrm{d}x\mathrm{d}y$$

3. For any complex number λ with $\mathfrak{Re}(\lambda) > -1$, we have:

Theorem [Keating-Snaith, 2000]:

$$\lim_{N\to\infty} \frac{1}{N^{\lambda^2}} \mathbb{E}\left[|Y_N|^{2\lambda}\right] = \frac{(G(1+\lambda))^2}{G(1+2\lambda)},$$

where G is the Barnes (double gamma) function.

4. For any complex number λ with $\mathfrak{Re}(\lambda) > -1$, we should have

Conjecture.[Keating-Snaith, 2000]

$$\lim_{T\to\infty} \frac{1}{(\log T)^{\lambda^2}} \frac{1}{T} \int_0^T \left|\zeta\left(\frac{1}{2}+it\right)\right|^{2\lambda} \mathrm{d}t = M(\lambda)\, A(\lambda)$$

where $M(\lambda)$ is the *random matrix factor*,

$$M(\lambda) = \frac{(G(1+\lambda))^2}{G(1+2\lambda)}$$

while $A(\lambda)$ is the *arithmetic factor* defined by the Euler product

$$A(\lambda) = \prod_p \left(1-\frac{1}{p}\right)^{\lambda^2} \left(\sum_{m=0}^{\infty} \left(\frac{\Gamma(\lambda+m)}{m!\Gamma(\lambda)}\right)^2 p^{-m}\right),$$

5. The classical central limit theorem: let $(X_k)_{k\geq 1}$ be a sequence of independent and identically distributed random variables on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$, such that $\mathbb{E}[X_k] = \mu_j$, and $\mathbb{V}(X_k) = \sigma^2 \in (0, \infty)$. Define

$$S_n = X_1 + \ldots + X_N.$$

Then we have:

$$\frac{S_N - N\mu}{\sigma\sqrt{N}} \xrightarrow{\mathsf{d}} \mathcal{N}(0,1),$$

where $\mathcal{N}(0,1)$ is a standard Gaussian random variable.

3. Central limit theorems in Probability Theory

3.1 Convergence in law and the usual tools

Definition. A sequence of real-valued random variables $(X_k)_{k\geq 1}$ is said to converge in law or in distribution to a random variable X if for any continuous and bounded function f we have:

$$\mathbb{E}[f(X_k)] \to \mathbb{E}[f(X)], \quad k \to \infty.$$

Remark. Alternative form: the sequence of probability measures (μ_k) is said to converge weakly to the probability measure μ if

$$\mu_k(f) = \int f(x)\mathrm{d}\mu_k(x) \to \mu(f) = \int f(x)\mathrm{d}\mu(x), \quad k \to \infty$$

for any continuous and bounded function f.

How to prove convergence in law?

1. The method of moments. If there is at most one probability measure μ such that

$$\lim_{k\to\infty} \int x^n \mathrm{d}\mu_k(x) = \int x^n \mathrm{d}\mu(x),$$

then

$$\mu_k \xrightarrow{\mathrm{d}} \mu.$$

Remark. Requires the existence of all moments.

2. The Fourier method. Let X be a random variable. Its characteristic function is defined, for $u \in \mathbb{R}$, by

$$\varphi_X(u) = \mathbb{E}[e^{\mathrm{i}uX}].$$

The characteristic function always exists and uniquely determines the law of X.

If X and Y are independent random variables, then:

$$\varphi_{X+Y}(u) = \varphi_X(u)\varphi_Y(u).$$

If X is a Gaussian random variable with mean μ and variance σ^2, then

$$\varphi_X(u) = \exp(\mathrm{i}u\mu - \sigma^2 u^2/2).$$

Lévy's theorem. Let $(X_k)_{k\geq 1}$ be a sequence of random variables.

a. If $X_n \xrightarrow{\mathsf{d}} X$, then $\varphi_{X_n}(u)$ converges pointwise to $\varphi_X(u)$.

b. If $\varphi_{X_n}(u)$ converges pointwise to a function h, which is continuous at 0, then there exists a random variable X, such that $h(u) = \varphi_X(u)$ and $X_n \xrightarrow{\mathsf{d}} X$.

3.2 Some classical theorems for sums of iid rv's

The Lyapunov theorem. Let $X_1, \ldots, X_n$ be independent random variables such that $\mathbb{E}\left[X_j\right] = 0$. Assume that there exists $\delta > 0$ such that $\mathbb{E}\left[|X_j|^{2+\delta}\right] < \infty$. Put $\sigma_j^2 = \mathbb{E}\left[X_j^2\right]$; $B_n = \sum_{j=1}^N \sigma_j^2$. If

$$B_n^{-1-\delta/2} \sum_{k=1}^{n} \mathbb{E}[|X_k|^{2+\delta}] \to 0,$$

then

$$\frac{1}{\sqrt{B_n}} \sum_{k=1}^{n} X_k \xrightarrow{\mathsf{d}} \mathcal{N}(0,1).$$

Theorem[Berry-Essen.] Let $(X_n)_{n\geq 1}$ be independent random variables such that $\mathbb{E}\left[X_j\right] = 0$, and $\mathbb{E}\left[|X_j|^3\right] < \infty$. Put $\sigma_j^2 = \mathbb{E}\left[X_j^2\right]$; $B_n = \sum_{j=1}^N \sigma_j^2$; $F_n(x) = \mathbb{P}\left[B_n^{-1/2} \sum_{j=1}^n X_j \leq x\right]$ and

$$L_n = \frac{1}{B_n^{3/2}} \sum_{j=1}^n \mathbb{E}\left[|X_j|^3\right].$$

Then there exist two constants A and C not depending on n such that the following uniform and non uniform estimates hold:

$$\sup_x |F_n(x) - \Phi(x)| \leq AL_n$$

and

$$|F_n(x) - \Phi(x)| \leq \frac{CL_n}{(1+|x|)^3}.$$

These theorems can be very useful in the study of the asymptotics of the characteristic polynomials of random matrices. It can be shown that the characteristic polynomial of random unitary matrices can be decomposed in law as product of independent random variables:

$$\det(I - U) \overset{\text{law}}{=} \prod_{k=1}^{n} X_k$$

where the X_k's are some simple and remarkable random variables. The central limit theorem easily follows (plus extra results): this will be explained by Chris Hughes.

But the moments conjecture and the corresponding random matrix result do not fit in a probabilistic framework.

3.3 How about strong limit theorems?

Let $X_1, X_2, \ldots$ be independent random variables such that $\mathbb{E}\left[X_j\right] = 0$, and $\sigma_j^2 = \mathbb{E}\left[X_j^2\right] < \infty$. Set $B_n = \sum_{j=1}^{N} \sigma_j^2$; $F_n(x) = \mathbb{P}\left[B_n^{-1/2} \sum_{j=1}^{n} X_j \leq \right.$ and $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} \mathrm{d}t$. If the conditions

1. $B_n \to \infty$;

2. $\frac{B_{n+1}}{B_n} \to 1$,

3. $\sup_x |F_n(x) - \Phi(x)| = \mathcal{O}\left((\log B_n)^{-1-\delta}\right)$,

are satisfied for some $\delta > 0$, then

$$\limsup \frac{S_n}{\sqrt{2B_n \log\log B_n}} = 1 \quad \text{a.s.}$$

It is a difficult task to give a satisfactory meaning to such strong limit theorems in random matrix theory because the probability spaces $U(N)$ change with the dimension. This raises the problem of constructing a "nice" infinite dimensional space "sitting above."

4. Mod-Gaussian convergence: basic facts.

Recall

$$\lim_{N\to\infty} \frac{1}{N^{\lambda^2}} \mathbb{E}\left[|Y_N|^{2\lambda}\right] = \frac{(G(1+\lambda))^2}{G(1+2\lambda)}.$$

Take $\lambda = \mathrm{i}u$, $u \in \mathbb{R}$, and let $Z_N = \log|Y_N|^2$. Then:

$$\lim_{N\to\infty} e^{u^2 \log N}\, \mathbb{E}[e^{\mathrm{i}uZ_N}] = \lim_{N\to\infty} e^{u^2 \log N}\, \mathbb{E}[e^{\mathrm{i}u \log|Y_N|^2}] = \frac{(G(1+\mathrm{i}u))^2}{G(1+2\mathrm{i}u)}.$$

In probability theory, the characteristic function is a more natural object to consider, because it always exists.

Definition. The sequence (Z_N) is said to *converge in the mod-Gaussian sense* if the convergence

$$e^{-\mathrm{i}u\beta_N + u^2\gamma_N/2}\,\mathbb{E}[e^{\mathrm{i}uZ_N}] \;\to\; \Phi(u)$$

holds for all $u \in \mathbb{R}$, where $\beta_N \in \mathbb{R}$ and $\gamma_N \geq 0$ are two sequences and Φ is a complex-valued function which is continuous at 0 (note that necessarily $\Phi(0) = 1$). We call (β_N, γ_N) the *parameters*, and Φ the associated *limiting function*.

The sequence (Z_N) is said to *strongly converge in the mod-Gaussian sense* if the convergence holds uniformly in u, on every compact subset of $\mathbb{R}$.

Why call it mod-Gaussian?

Proposition. Let (X_N) be a sequence of real random variables converging in law to a limiting variable with characteristic function Φ. If for each N we let

$$Z_N = X_N + G_N,$$

where G_N is a Gaussian random variable independent of X_N, and with mean β_N and variance γ_N, then we have the strong mod-Gaussian convergence of the sequence (Z_N), with limiting function Φ and parameters (β_N, γ_N).

Remark. As we shall see, the above situation does not cover all cases of mod-Gaussian convergences.

Proposition. *(1)* Let (Z_N) be a sequence of real-valued random variables for which the mod-Gaussian convergence holds with parameters (β_N, γ_N) and limiting function Φ. Then the mod-Gaussian convergence holds for some other parameters (β'_N, γ'_N) and limiting function Φ', if and only if the limits

$$\beta = \lim_{N \to +\infty} (\beta_N - \beta'_N), \qquad \gamma = \lim_{N \to +\infty} (\gamma_N - \gamma'_N), \tag{1}$$

exist in $\mathbb{R}$. In this case Φ' is given by

$$\Phi'(u) = e^{i\beta u - u^2 \gamma/2}\, \Phi(u), \tag{2}$$

and if the strong convergence holds with the parameters (β_N, γ_N) it also holds with (β'_N, γ'_N).

(2) Let (Z_N) and (Z'_N) be two sequences of random variable with mod-Gaussian convergence (resp. strong convergence), with respective parameters (β_N, γ_N) and (β'_N, γ'_N), and limiting functions Φ and Φ'. If Z_N and Z'_N are independent for all N, then the sums $(Z_N + Z'_N)$ satisfy mod-Gaussian convergence (resp. strong convergence) with limiting function the product $\Phi\Phi'$ and parameters $(\beta_N + \beta'_N, \gamma_N + \gamma'_N)$.

Remark. Given a family $\mathcal{F} = (\mu_\lambda)_{\lambda \in \Lambda}$ of probability distributions parametrized by some set Λ, such that the Fourier transforms

$$\widehat{\mu}(\lambda, u) = \int_{\mathbb{R}} e^{itx} d\mu_\lambda(t)$$

are non-zero for all $u \in \mathbb{R}$, one would say that a sequence of random variables (Z_N) converges in the mod-$\mathcal{F}$ sense if, for some sequence $\lambda_N \in \Lambda$, we have

$$\lim_{N \to +\infty} \widehat{\mu}(\lambda_N, u)^{-1} \, \mathbb{E}(e^{iuZ_N}) = \Phi(u)$$

for all $u \in \mathbb{R}$, the limiting function Φ being continuous at 0.

5. Mod-Gaussian convergence in probability theory

5.1 The CLT for mod-Gaussian convergence

Let (X_i^n), for $n \geq 1$ and $1 \leq i \leq n$, be random variables, where the variables

$$X_1^n, \ldots, X_n^n$$

in each row are i.i.d. with law denoted by μ_n. Let

$$S_n = X_1^n + \ldots + X_n^n.$$

Consider the *logarithmic mean* of the S_n:

$$Z_N = \sum_{n=1}^{N} \frac{S_n}{n} = \sum_{n=1}^{N} \frac{1}{n}(X_1^n + \ldots + X_n^n)$$

with variance:

$$\mathbb{V}(Z_N) = H_N = \sum_{n=1}^{N} \frac{1}{n}.$$

For a numerical sequence (u_n) that converges to a limit α, the analogue *logarithmic means*

$$v_N = \frac{1}{\log N} \sum_{n=1}^{N} \frac{u_n}{n}$$

also converge to this limit. This shows that, intuitively, the Z_N can "amplify" the sums S_N by a logarithmic factor.

Theorem. Let $\left(X_i^n\right)_{i.n\geq 1}$ be a triangular array of random variables, *all independent*, and such that the variables in the nth row have the same law μ_n, and let us assume that μ_n has mean zero, variance 1 and third absolute moment satisfying

$$\sum_{n=1}^{\infty} \frac{m_n}{n^2} < \infty, \qquad \text{where} \qquad m_n = \int_{\mathbb{R}} |x|^3 \mu_n(\mathrm{d}x).$$

Then the logarithmic means Z_N strongly converge in the mod-Gaussian sense, with parameters $(0, H_N)$, or with parameters $(0, \log N)$.

Corollary. $Z_N/\sqrt{\log N}$ converges in law to the standard Gaussian law $\mathcal{N}(0,1)$.

5.2 A Gnedenko-Kolmogorov type theorem

Definition. Denote by μ^{n*} the n-fold convolution of a probability measure μ with itself. The probability measure μ on $\mathbb{R}$ is said to be infinitely divisible if for any positive integer n, there is a probability measure μ_n on $\mathbb{R}$ such that $\mu = \mu_n^{n*}$.

Theorem. The following properties hold:

(1) If μ is an infinitely divisible distribution on $\mathbb{R}$, then for $u \in \mathbb{R}$, we have

$$\widehat{\mu}(u) \equiv \int_{-\infty}^{\infty} e^{\mathrm{i}ux}\mu(\mathrm{d}x) = \exp\left[-\frac{1}{2}\sigma u^2 + \mathrm{i}\beta u + \int_{\mathbb{R}}\left(e^{\mathrm{i}ux} - 1 - \mathrm{i}ux\mathbb{1}_{|x|\leq 1}\right)\nu(\mathrm{d}x)\right] \quad (3)$$

where $\sigma \geq 0$, $\beta \in \mathbb{R}$ and ν is a measure on $\mathbb{R}$, called the Lévy measure, satisfying

$$\nu(\{0\}) = 0 \text{ and } \int_{\mathbb{R}}\left(x^2 \wedge 1\right)\nu(\mathrm{d}x) < \infty. \quad (4)$$

(2) The representation of $\widehat{\mu}(u)$ in *(3)* by σ, β and ν is unique.

(3) Conversely, if $\sigma \geq 0$, $\beta \in \mathbb{R}$ and ν is a measure satisfying *(4)*, then there exists an infinitely divisible distribution μ whose characteristic function is given by *(3)*.

The parameters (σ, β, ν) are called the generating triplet of μ.

Remark. Let h be a *truncation function*, that is a real function on $\mathbb{R}$, bounded, and such that $h(x) = x$ on a neighborhood of 0. Then for every $u \in \mathbb{R}$, $x \mapsto (e^{\mathrm{i}ux} - 1 - uh(x))$ is integrable with respect to ν, and (3) may be rewritten as:

$$\widehat{\mu}(u) = \exp\left[-\frac{1}{2}\sigma u^2 + \mathrm{i}\beta_h u + \int_{\mathbb{R}} \left(e^{\mathrm{i}ux} - 1 - \mathrm{i}uh(x)\right) \nu(\mathrm{d}x)\right]$$

where

$$\beta_h = \beta + \int_{-\infty}^{\infty} (h(x) - x\mathbb{1}_{|x|\leq 1})\nu(\mathrm{d}x).$$

The triplet (σ, β_h, ν) is called the generating triplet of μ with respect to the truncation function h.

Theorem. In the setting of the previous Theorem, assume further that the probability measures μ_n are infinitely divisible. Then the sequence (Z_N) strongly converges in the mod-Gaussian sense, with the parameters $(0, H_N)$ and limiting function $\Phi = e^{\Psi}$, where

$$\Psi(u) = \int_{-\infty}^{\infty} \left(e^{\mathrm{i}ux} - 1 - \mathrm{i}ux + \frac{u^2 x^2}{2} \right) \nu(\mathrm{d}x)$$

and

$$\nu = \sum_{n=1}^{\infty} n\nu'_n, \tag{5}$$

and the measures ν'_n are defined by

$$\nu'_n(A) = \int_{-\infty}^{\infty} \mathbb{1}_A(x/n)\nu_n(\mathrm{d}x),$$

for any Borel set A. Consequently, ν is a positive measure which satisfies $\nu(\{0\}) = 0$ and $\int_{-\infty}^{\infty} |x|^3 \nu(\mathrm{d}x) < \infty$.

Theorem. Let (Z_N) be a sequence of real-valued random variables whose respective laws μ_N are infinitely divisible, with generating triplets (σ_N, b_N, ν_N) relative to a fixed truncation function h. Then we have mod-Gaussian strong convergence *if and only if* the following two conditions hold:

(1) The sequence $\kappa_N = \int_{-\infty}^{\infty} h(x)^3 \nu_N(\mathrm{d}x)$ converges to a finite limit κ;

(2) There exists a nonnegative measure ν satisfying

$$\nu(\{0\}) = 0, \quad \int_{\mathbb{R}} (x^4 \wedge 1)\nu(\mathrm{d}x) < +\infty$$

and such that $\nu_N(f) \to \nu(f)$ for any continuous function f with $|f(x)| \leq C(x^4 \wedge 1)$ for $x \in \mathbb{R}$ and some constant $C \geq 0$.

Under these conditions, one may take the parameters

$$\beta_N = b_N, \qquad \gamma_N = \sigma_N + \nu_N(h^2),$$

and the limiting function is then $\Phi = \exp(\Psi)$, where

$$\Psi(u) = -\mathrm{i}\frac{u^3}{6}\kappa + \int_{-\infty}^{\infty}\left(e^{\mathrm{i}ux} - 1 - \mathrm{i}uh(x) + \frac{u^2h(x)^2}{2} + \mathrm{i}\frac{u^3h(x)^3}{6}\right)\nu(\mathrm{d}x).$$

6. Mod-Gaussian convergence in Number Theory

6.1 The arithmetic factor in the moment conjecture

Recall the moment conjecture for the Riemann zeta function:

$$\lim_{T\to+\infty} \frac{1}{T(\log T)^{\lambda^2}} \int_0^T |\zeta(1/2+it)|^{2\lambda} \mathrm{d}t = A(\lambda)M(\lambda)$$

for any complex number λ such that $\mathfrak{Re}(\lambda) > -1$, where

$$M(\lambda) = \frac{G(1+\lambda)^2}{G(1+2\lambda)}, \qquad G(z) \text{ the Barnes double-gamma function,} \tag{6}$$

$$A(\lambda) = \prod_p \left(1-\frac{1}{p}\right)^{\lambda^2} \left\{ \sum_{m\geq 0} \left(\frac{\Gamma(m+\lambda)}{m!\Gamma(\lambda)}\right)^2 p^{-m} \right\}. \tag{7}$$

Proposition. There exists a sequence (Z_N) of positive real-valued random variables and positive real numbers $\gamma_N > 0$ such that

$$e^{u^2\gamma_N/2}\,\mathbb{E}(e^{iuZ_N}) \to A(iu)$$

locally uniformly for $u \in \mathbb{R}$.

Remark. let (X_p) be a sequence of independent random variables identically and uniformly distributed on the unit circle. Then the sequence of random variables defined by

$$\sum_{p\leq N} \log\left|1-\frac{X_p}{\sqrt{p}}\right|^{-2} = \log \prod_{p\leq N}\left|1-\frac{X_p}{\sqrt{p}}\right|^{-2}$$

converges as $N\to+\infty$, in the mod-Gaussian sense, with limiting function given by the arithmetic factor evaluated at iu, and parameters $(0, 2\log(e^{\gamma}\log N))$.

Remark. Compare it with the formal identity:

$$|\zeta(1/2+it)|^2 \quad \text{“}=\text{”} \quad \prod_p \left|1-\frac{1}{p^{1/2+it}}\right|^{-2},$$

6.2 Revisiting Erdős-Kác

$$\lim_{N\to+\infty} \frac{1}{N}|\{n \le N \mid a < \frac{\omega(n) - \log\log N}{\sqrt{\log\log N}} < b\}| = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-t^2/2} dt$$

for any real numbers $a < b$.

Let

$$\omega'(n) = \omega(n) - 1$$

for $n \ge 2$.

A Poisson random variable P_λ with parameter $\lambda > 0$ is one taking integer values $k \ge 0$ with

$$\mathbb{P}(P_\lambda = k) = \frac{\lambda^k}{k!} e^{-\lambda}.$$

The characteristic function is then given by

$$\mathbb{E}(e^{iuP_\lambda}) = \exp(\lambda(e^{iu} - 1)),$$

and strong mod-Poisson convergence of a sequence Z_N of random variables with parameters λ_N means that the limit

$$\lim_{N \to +\infty} \exp(\lambda_N(1 - e^{iu}))\, \mathbb{E}(e^{iuZ_N}) = \Phi(u)$$

exists for every $u \in \mathbb{R}$, and the convergence is locally uniform. The *limiting function* Φ is then continuous and $\Phi(0) = 1$.

Proposition. For $u \in \mathbb{R}$, let

$$\Phi(u) = \frac{1}{\Gamma(e^{iu}+1)} \prod_p \left(1-\frac{1}{p}\right)^{e^{iu}} \left(1+\frac{e^{iu}}{p-1}\right). \tag{8}$$

This Euler product is absolutely and locally uniformly convergent. Moreover, for any $u \in \mathbb{R}$, we have

$$\lim_{N\to+\infty} \frac{(\log N)^{(1-e^{iu})}}{N} \sum_{2\le n\le N} e^{iu\omega'(n)} = \frac{1}{\Gamma(e^{iu}+1)} \prod_p \left(1-\frac{1}{p}\right)^{e^{iu}} \left(1+\frac{e^{iu}}{p-1}\right)$$

and the convergence is locally uniform.

Corollary. Consider random variables M_N, for $N \geq 2$, such that

$$\mathbb{P}(M_N = n) = \frac{1}{N-1}, \qquad 2 \leq n \leq N,$$

and let $Z_N = \omega'(M_N)$. Then the sequence (Z_N) converges strongly in the mod-Poisson sense with limiting function Φ given by (8) and parameters $\lambda_N = \log\log N$.

Corollary. Erdős-Kác

Proof. We have to show that:

$$Y_N = \frac{\omega(M_N) - \log\log N}{\sqrt{\log\log N}}$$

converges in law to a standard Gaussian variable. Elementary asymptotic manipulations show that:

$$\mathbb{E}(e^{iuY_N}) \to \exp\left(-\frac{u^2}{2}\right).$$

The result follows from Lévy's theorem.