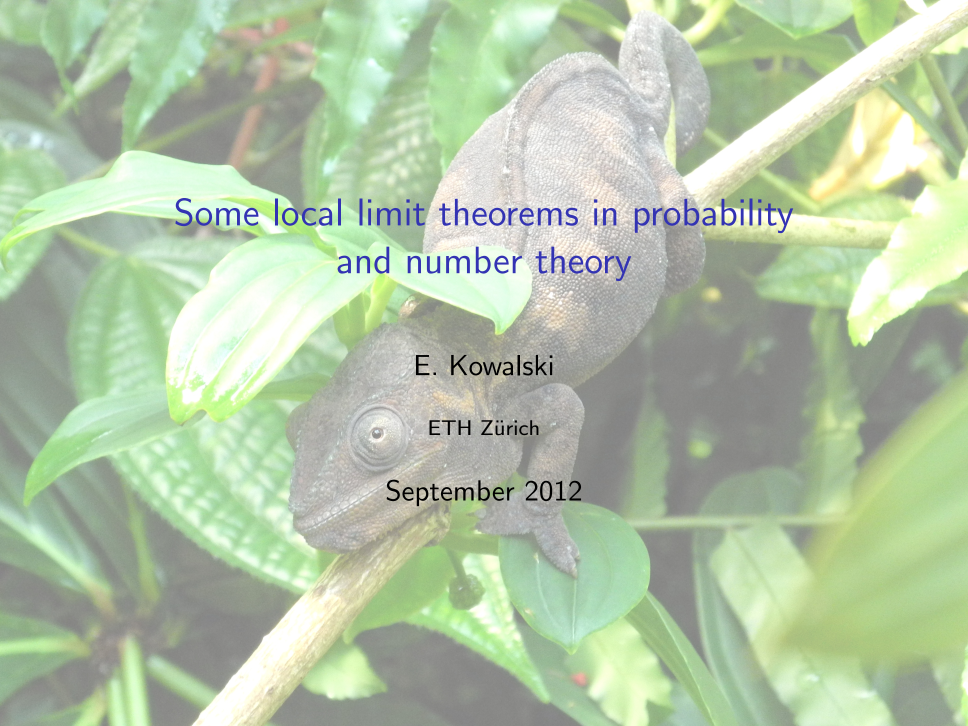


Some local limit theorems in probability and number theory

E. Kowalski

ETH Zürich

September 2012

Introduction

[Joint work with A. Barbour, F. Delbaen, J. Jacod, A. Nikeghbali]

- (J-K-N) Forum Math. 23 (2011), 835–873
- (K-N) IMRN 2010
- (B-K-N) arXiv 0912.1886
- (D-K-N) arXiv 1107.5657
- (K-N) J. London Math. Soc. 86 (2012), 291–319

We consider limit theorems in probability theory which have arithmetic incarnations and applications.
One basic idea is to finds information which lie beyond such universal statements as the Central Limit Theorem.

Example 1

[The Erdös-Kac theorem]

Consider random variables N_n which are uniformly distributed among integers $1 \leq k \leq n$. For an integer $k \geq 1$, let $\omega(k)$ be the number of prime divisors of k. Then

$$\frac{\omega(N_n) - \log\log n}{\sqrt{\log\log n}} \stackrel{law}{\Rightarrow} \mathcal{N}(0,1).$$

Example 2

[Selberg's Normal Limit Theorem]

Consider random variables U_T uniformly distributed on $[0, T]$. Let

$$\zeta(s) = \sum_{n \geq 1} n^{-s} = \frac{s}{s-1} + s \int_1^{+\infty} \{x\} x^{-s-1} dx$$

be the Riemann zeta function, meromorphic on $\mathbf{C}$. Then as $T \to +\infty$, we have

$$\frac{\log |\zeta(1/2 + iU_T)|}{\sqrt{\frac{1}{2} \log \log T}} \overset{law}{\Rightarrow} \mathcal{N}(0, 1).$$

Discussion

These two results have the following drawbacks (for certain purposes):

- The limit distributions are the same, although the quantities $\log|\zeta(1/2+it)|$ and $\omega(k)$ are very different;
- In particular, $\omega(k)$ takes discrete values, whereas $\log|\zeta(1/2+it)|$ is a continuous quantity, and this distinction is lost;
- As a consequence, these two theorems do not give much information on the distribution of *non-typical* values of ω (e.g., of prime powers, such that $\omega(k)=1$) or of $\zeta(1/2+it)$ (e.g., of zeros of ζ on the critical line).

Refining convergence in law

We attempt to refine convergence in law of normalized sequences

$$X_n = \frac{Y_n - m_n}{\sqrt{\sigma_n}}$$

by looking more carefully at the limiting behavior of the characteristic functions $\varphi_n(t) = \mathbf{E}(e^{itY_n})$ *without normalizing*. We find that this behavior often contains significant information in addition to a possible Normal Limit Theorem for X_n.

Example

Let ϖ_n be a random variable counting the number of distinct cycles in a uniformly chosen permutation σ of $\{1, \ldots, n\}$ (e.g., a transposition σ has $\varpi_n(\sigma) = n - 1$). One also knows that $X_n = (\varpi_n - \log n)/\sqrt{\log n}$ converges to $\mathcal{N}(0, 1)$.
But the characteristic function is given exactly by

$$\mathbf{E}(e^{it\varpi_n}) = \prod_{j=1}^{n} (1 - j^{-1} + j^{-1} e^{it})$$

from which we can extract information.

Example

The product diverges as $n \to +\infty$ for $t \notin 2\pi\mathbf{Z}$. But we can write

$$\mathbf{E}(e^{it\varpi_n}) = \prod_{j=1}^{n} (1 + (e^{it} - 1)/j)(1 + 1/j)^{1-e^{it}} \times \exp((e^{it} - 1)H_n)$$

where $H_n = 1 + 1/2 + \cdots + 1/n$. The second term is the characteristic function of a Poisson random variable P_{H_n} with parameter $\lambda = H_n$ (recall that in general

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

for $k \geq 0$.)

Example

The first term converges as $n \to +\infty$ and in fact

$$\prod_{j\geq 1}\Big(1+\frac{z}{j}\Big)\Big(1+\frac{1}{j}\Big)^{-z}=\frac{1}{\Gamma(1+z)}$$

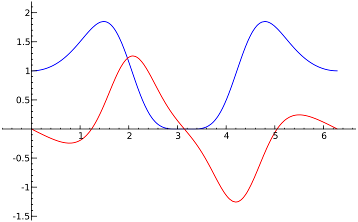
for any $z \in \mathbf{C}$ (Euler) so that for any $t \in \mathbf{R}$, we get

$$\mathbf{E}(e^{it\varpi_n}) \sim \exp((e^{it}-1)H_n)\frac{1}{\Gamma(e^{it})} = \mathbf{E}(e^{itP_{H_n}})\frac{1}{\Gamma(e^{it})}$$

as $n \to +\infty$.

Remarks on this example

- The factorization suggests that there could be a decomposition $\varpi_n = X_n + Y_n$ where $X_n \stackrel{law}{=} P_{H_n}$ and where Y_n is independent of X_n and converges in law to a random variable with characteristic function $1/\Gamma(e^{it})$.
- But $1/\Gamma(e^{it})$ is not a characteristic function!



- We called this type of behavior *mod-Poisson convergence* with *parameters* H_n and *limiting function* $1/\Gamma(e^{it})$.
- This is a type of Poisson approximation that seems widespread but not much studied. (An exception is an early paper of Hwang with different terminology.)

The Rényi-Turán formula

Taking again the example of $\omega(N_n)$, Rényi-Turán proved

$$\frac{1}{n}\sum_{k\leq n} e^{it\omega(k)} = \mathbf{E}(e^{itP_{\log\log n}})\Phi(t)(1+o(1))$$

where

$$\Phi(t) = \frac{1}{\Gamma(e^{it})} \times \prod_p \left(1-\frac{1}{p}\right)^{e^{it}}\left(1+\frac{e^{it}}{p-1}\right).$$

Moreover, the infinite product over p is also the limiting function for

$$X_n = \sum_{p\leq n} B_p$$

where the B_p are independent Bernoulli with $\mathbf{P}(B_p = 1) = p^{-1}$, *which is the "heuristic" probability that a "random" integer be divisible by n.*

The Gaussian case

Maybe the first example that was recognized is in the Gaussian case, where one says that X_n converges in the *mod-Gaussian sense* with *variance* σ_n (usually $\sigma_n \to +\infty$) and *limiting function* $\Phi(t)$ if

$$\mathbf{E}(e^{itX_n}) = e^{-\sigma_n t^2/2}\Phi(t)(1 + o(1))$$

(uniformly for t in compact sets).
The "trivial" case is when

$$X_n \overset{law}{=} N_{\sigma_n} + Y_n$$

where N_σ is centered normal of variance σ and Y_n is independent of N_{σ_n}, converging in law to Y with $\mathbf{E}(e^{itY}) = \Phi(t)$.

Random matrices

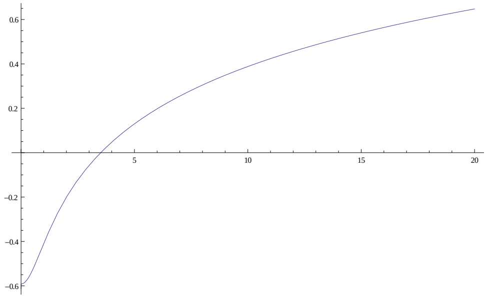
Keating and Snaith proved that if X_n is a random matrix taking values in the unitary group $U(n)$, distributed according to the natural Haar measure, and $P_n(T) = \det(1 - TX_n)$, we have

$$\mathbf{E}(e^{it|P_n(1)|}) = e^{-(\log n)t^2/2} \frac{G(1+it/2)^2}{G(1+it)}(1+o(1))$$

locally uniformly for $t \in \mathbf{R}$. Here $G(z)$ is the *Barnes function*, holomorphic of order 2 such that $G(1) = 1$ and

$$G(z+1) = \Gamma(z)G(z).$$

It is known that $\Phi(t) = G(1+it/2)^2/G(1+it) \approx \exp(t^2 \log t)$ for t large, so this function is far from being a characteristic function!



Plot of $1/t^2 \times \log|\Phi(t)|$

Searching for meaning

We now ask: what is the meaning of such behavior? What are its consequences? One possible answer is: local limit theorems. In fact, one can prove local limit theorems for

$$\mathbf{P}(X_n \in B), \qquad B \subset \mathbf{R}^d \text{ open or Jordan measurable,}$$

for very general sequences of random vectors (X_n) with values in $\mathbf{R}^d$, $d \geq 1$, satisfying some form of “mod”-convergence. These cases go well beyond the Poisson and Gaussian cases, and the conditions are much less stringent.

Mod-ϕ convergence

Fix $d \geq 1$ and a probability measure on $\mathbf{R}^d$ with probability law μ and characteristic function φ. Let (X_n) be random $\mathbf{R}^d$-valued vectors with characteristic functions φ_n.
We say there is mod-φ convergence if there exist $A_n \in \mathrm{GL}_d(\mathbf{R})$, such that

- H1 The characteristic function φ is integrable on $\mathbf{R}^d$;
- H2 Denoting $\Sigma_n = A_n^{-1}$, we have $\Sigma_n \to 0$ and the vectors $Y_n = \Sigma_n(X_n)$ converge in law with limit μ;
- H3 For all $k \geq 0$, we have

$$\sup_{n\geq 1} \int_{\substack{|t|\geq a \\ |\Sigma_n^* t|\leq k}} |\varphi_n(\Sigma_n^* t)| dt \to 0 \text{ as } a \to +\infty.$$

Clarification

H1 This implies that $d\mu = \alpha(t)dt$ for some density α;

H2 Note that $\mathbf{E}(e^{it\cdot Y_n}) = \varphi_n(\Sigma_n^* t)$.

H3 This is a uniform-integrability condition. It holds, for instance, if there exist fixed integrable functions h_k such that

$$|\varphi_n(\Sigma_n^* t)| \leq h_k(t)$$

for all n and all t such that $|\Sigma_n^* t| \leq k$, since then

$$\int_{\substack{|t|\geq a \\ |\Sigma_n^* t|\leq k}} |\varphi_n(\Sigma_n^* t)| dt \leq \int_{a\leq |t|} |h_k(t)| dt \to 0.$$

H2, H3 If **H1** holds, and

$$\varphi_n(t) = \Phi(t)\varphi(A_n^* t)(1 + o(1)) \text{ as } n \to +\infty$$

for some continuous Φ, *and the convergence holds uniformly* on sets of the form $|t| \leq A_n^* k$ for $k > 0$, then we have mod-ϕ convergence.

Local limit theorem for mod-ϕ convergence

Theorem (Delbaen-K-Nikeghbali)

Assume mod-ϕ convergence for X_n. Then for f continous and compactly supported we have

$$\mathbf{E}(f(X_n)) = \alpha(0)|\det(A_n)|^{-1}\Big(\int_{\mathbf{R}^d} f(x)dx\Big)(1 + o(1))$$

as $n \to +\infty$.

Remark

This applies also to the case $\alpha(0) = 0$, but in that case it is more interesting to apply it, e.g., to $X_n + A_n c$, where $c \neq 0$ is a constant vector.

Proof

We take $d = 1$, so $A_n(x) = a_n x$ with $a_n \neq 0$, and $\Sigma_n^* t = a_n^{-1} t$. By an approximation argument, it is enough to prove the result when the Fourier transform $\hat{f}$ has compact support. Let μ_n be the law of X_n. We have

$$\begin{aligned}
\mathbf{E}(f(X_n)) &= \int_{\mathbf{R}} f(x) d\mu_n(x) \\
&= \frac{1}{2\pi} \int_{\mathbf{R}} \int_{\mathbf{R}} \hat{f}(t) e^{itx} dt d\mu_n(x) \\
&= \frac{1}{2\pi} \int_{\mathbf{R}} \hat{f}(t) \varphi_n(t) dt \\
&= \frac{1}{2\pi a_n} \int_{\mathbf{R}} \hat{f}(a_n^{-1} s) \varphi_n(a_n^{-1} s) ds.
\end{aligned}$$

Proof (II)

Let $k \geq 1$ be such that $\mathrm{Supp}(\hat{f}) \subset [-k, k]$, so that

$$\mathbf{E}(f(X_n)) = \frac{1}{2\pi a_n} \int_{|s| \leq a_n k} \hat{f}(a_n^{-1} s) \varphi_n(a_n^{-1} s) ds.$$

By **H2** and the Lévy criterion, the integrand converges pointwise to $\varphi(s)\hat{f}(0)$. Uniform integrability then implies convergence in L^1: for any $\varepsilon > 0$, and $a > 0$ large enough we have

$$\int_{a < |s| \leq k a_n} |\varphi_n(a_n^{-1} s) \hat{f}(a_n^{-1} s)| ds \leq \|\hat{f}\|_\infty \int_{a < |s| \leq k a_n} |\varphi_n(a_n^{-1} s)| ds < \varepsilon$$

for all n by **H3**.

Proof (III)

For $|s| \leq a$, we have dominated convergence

$$|\hat{f}(a_n^{-1}s)\varphi_n(a_n^{-1}s)| \leq \mathbf{1}_{|s|\leq a}\|\hat{f}\|_\infty$$

so

$$\frac{1}{2\pi}\int_{|s|\leq a}\hat{f}(a_n^{-1}s)\varphi_n(a_n^{-1}s)ds \longrightarrow \hat{f}(0)\int_{|s|\leq a}\varphi(s)ds.$$

For a large enough, this differs from $\hat{f}(0)\int\varphi(s)ds$ by at most ε, and then

$$\frac{1}{2\pi}\int_{|s|\leq a_n k}\hat{f}(a_n^{-1}s)\varphi_n(a_n^{-1}s)ds \longrightarrow \frac{1}{2\pi}\hat{f}(0)\int_{\mathbf{R}}\varphi(s)ds$$

$$= \alpha(0)\int_{\mathbf{R}} f(s)ds.$$

□

Examples

Example 1 (sums of i.i.d variables) Let (S_n) be a sequence of i.i.d variables, symmetric, not supported on a lattice, and let

$$X_n = \frac{S_1 + \cdots + S_n}{b_n}$$

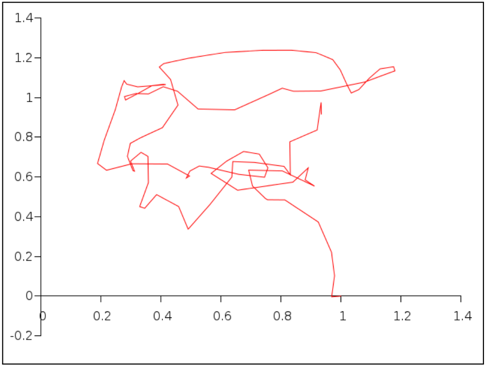
for suitable b_n so that X_n converges in law to some μ. Classical results (Shepp, Borovkov–Mogulskii, Stone, Bretagnolle–Dacunha-Castelle) imply local limit theorems for

$$b_n \mathbf{P}(S_1 + \cdots + S_n \in B)$$

which the general theorem recovers. *Usually* we do *not* have the strong convergence

$$\mathbf{E}(e^{it\cdot(S_1 + \cdots S_n)}) = \varphi(b_n t)\Phi(t)(1 + o(1)).$$

Example 2 (Winding number). Let W_u, $u \geq 0$, be a complex Brownian motion starting at 1. Let θ_u denote the argument of W_u, starting with $\theta_0 = 0$ and defined by continuity.



Spitzer proved that

$$\frac{2\theta_u}{\log u} \overset{law}{\Rightarrow} \frac{1}{\pi} \frac{dx}{1+x^2}$$

as $u \to +\infty$ (Cauchy law).

Theorem

If $u_n \to +\infty$, we have mod-ϕ convergence for θ_{u_n} with $\varphi(t) = \exp(-|t|)$, $a_n = \frac{1}{2}(\log u_n)$. In particular

$$\frac{\log u}{2} \mathbf{P}(W_u \in [a, b]) \to \frac{1}{\pi}(b - a) \text{ as } u \to +\infty$$

for any real $a < b$.

This follows easily from the fact that Spitzer computed exactly the characteristic function of θ_u in terms of Bessel functions. One even gets the stronger "mod-Cauchy convergence" with limiting function

$$\Phi(t) = 8^{-|t|/2} \frac{\Gamma(1/2)}{\Gamma((|t| + 1)/2)}.$$

Example 3 (Random matrices).
Let again X_n be a random matrix taking values in the unitary group $U(n)$, distributed according to the natural Haar measure. (One can deal similarly with unitary symplectic groups and orthogonal groups.) Now put $P_n = \log\det(1 - X_n) \in \mathbf{C} = \mathbf{R}^2$. The characteristic function is known (Keating–Snaith) to be

$$\mathbf{E}(e^{it\cdot P_n}) = \prod_{1\leq j\leq n} \frac{\Gamma(j)\Gamma(j+it_1)}{\Gamma(j+\frac{1}{2}(it_1+t_2))\Gamma(j+\frac{1}{2}(it_1-t_2))}$$

for $t = (t_1, t_2) \in \mathbf{R}^2$.

One can deduce that

$$\mathbf{E}(e^{it\cdot P_n}) = \Phi(t)e^{-(\log n)|t|^2/4}(1+o(1))$$

with

$$\Phi(t_1,t_2) = \frac{G(1+(it_1-t_2)/2)G(1+(it_1+t_2)/2)}{G(1+it_1)}$$

uniformly for $|t| \leq Cn^{1/6}$. This is (more than) mod-ϕ convergence for the Gaussian distibution and $A_n(t_1,t_2) = \frac{1}{2}(\log n)(t_1,t_2)$. Hence

$$\mathbf{P}(P_n \in B) \sim \frac{1}{2\pi}\frac{\sqrt{2}}{\sqrt{\log n}}\int_B dx$$

as $n \to +\infty$ for any Jordan-measurable subset B of $\mathbf{R}^2$.

Example 4 (the Riemann zeta function).
Selberg's Theorem can be generalized to

$$\frac{\log\zeta(1/2+iU_T)}{\sqrt{\frac{1}{2}\log\log T}} \overset{law}{\Rightarrow} \text{standard complex Gaussian}$$

where U_T is uniform on $[0, T]$. Thus we have **H1** and **H2** for $X_T = \log\zeta(1/2+iU_T)$ with $A_T(t_1, t_2) = (\frac{1}{2}\log\log T)(t_1, t_2)$. *We conjecture that* **H3** *holds*. In particular:

Conjecture

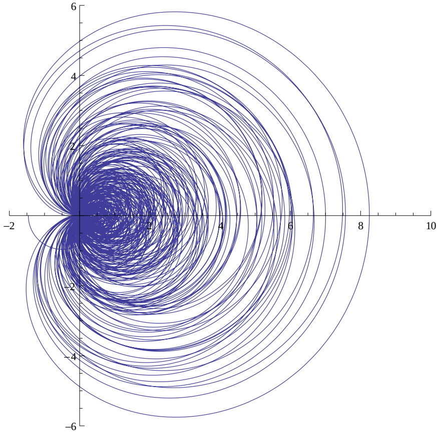
For any Jordan measurable set $B \subset \mathbf{C}$, we have

$$\frac{1}{T}\lambda(\{t \in [0, T] \mid \log\zeta(1/2+it) \in B\}) \sim \frac{1}{\sqrt{\frac{1}{2}\log\log T}}\lambda(B).$$

This would imply the following (answering a question of Ramachandra):

Corollary (Conditional)

The set of values $\zeta(1/2 + it)$, for $t \in \mathbf{R}$, is dense in $\mathbf{C}$.



The analogue of this has been known for a long time when $1/2$ is replaced with any fixed $\sigma \in]1/2, 1[$. In that case, Bohr and Jessen showed that $\log \zeta(\sigma + iU_T)$ converges in law to some measure μ_σ as $T \to +\infty$, and that $\mathrm{Supp}(\mu_\sigma) = \mathbf{C}$.
In fact, one has

$$\mathbf{E}(e^{it\cdot\mu_\sigma}) = \prod_p \mathbf{E}(e^{it\cdot Z_p})$$

where Z_p is distributed like

$$\log\Big(\frac{1}{1 - p^{-\sigma}\Theta_p}\Big)$$

with Θ_p uniform on the unit circle. This is the characteristic function of the series

$$\sum_p \log\Big(\frac{1}{1 - p^{-\sigma}\Theta_p}\Big).$$

This should be compared with the formula

$$\zeta(\sigma + i\tau) = \prod_p \frac{1}{1 - p^{-\sigma} p^{-i\tau}}$$

(which holds for $\sigma > 1$). This shows that statistically, the zeta function on a fixed vertical line with real part in $]1/2, 1[$ behaves as if the factors in the Euler product were completely independent.

The strong Keating–Snaith conjecture

Keating–Snaith conjecture that $\log\zeta(1/2+iU_T)$ exhibits mod-gaussian convergence

$$\mathbf{E}(e^{it\cdot\log\zeta(1/2+iU_T)}) = \Phi_1(t)\Phi_2(t)e^{-t^2(\log\log T)/4}(1+o(1))$$

where

$$\Phi_1(t) = \frac{G(1+(it_1-t_2)/2)G(1+(it_1+t_2)/2)}{G(1+it_1)}$$

and

$$\Phi_2(t) = \prod_p \Big(1-\frac{1}{p}\Big)^{-|t|^2/4}\mathbf{E}(e^{it\cdot Z_p})$$

where Z_p is distributed like

$$\log\frac{1}{1-p^{-1/2}\Theta_p}.$$

This conjecture suggests the possibility of a probabilistic model of the values of the Riemann zeta function on the critical line which combines, with some subtle dependency structure, two ingredients:

- Random unitary matrices (of size $\sim \log T$ if $t \in [T/2, T]$);
- A product over small primes with independent random variables.

This is also very similar to the lessons of the Rényi-Turán formula: for the number of prime divisors of k, it involves

- Random permutations (of size $\log k$ if $n \in [k/2, k]$);
- A sum of independent Bernoulli variables.

It is a challence to construct or understand such models. For the number of irreducible factors of polynomials over finite fields, we have however a very convincing explanation (K-N), which is encouraging.

Going beyond the local limit theorem

The local limit theorem does not indicate a rate of convergence. In

$$\mathbf{P}(X_n \in B) \sim \frac{1}{(2\pi)^d} \frac{1}{|\det(A_n)|} \int_B dx,$$

the *location* of B does not appear, only its size. If we ask

How large must n be before $\mathbf{P}(X_n \in B)$ is of the right size?

the answer must also depend on the location of B.
At least in the Gaussian case, one can prove a quantitative local limit theorem if one assumes sufficiently uniform version of mod-gaussian convergence.

For random variables (X_n), assume that $\sigma_n \to +\infty$ are such that

- We have $c > 0$ and $a > 0$ (small) such that

$$\mathbf{E}(e^{itX_n}) = \Phi(t)e^{-\sigma_n t^2/2}\Big(1 + O\Big(\frac{1}{\exp(\sigma_n^c)}\Big)\Big)$$

  uniformly for $|t| \leq \sigma_n^a$;
- The function Φ is C^1 for $|t| \leq 2$;
- The function Φ satisfies $\Phi(t) = O(e^{|t|^A})$ for some $A > 0$ (large).

Let Y_n be a centered Gaussian variable with variance σ_n.

Theorem (K-N)

Under these conditions, there exists $\delta > 0$ such that for any open interval $I =]x_0 - \varepsilon, x_0 + \varepsilon[\subset \mathbf{R}$, we have

$$\mathbf{P}(X_n \in I) = \mathbf{P}(Y_n \in I) + O\Big(\frac{1}{\sigma_n^{1/2+\delta}} + \frac{1}{\varepsilon \sigma_n}\Big)$$

uniformly in terms of x_0 and ε.

The location enters from the main term:

$$\mathbf{P}(|Y_n - x_0| < \varepsilon) \gg \frac{\varepsilon}{\sqrt{\sigma_n}} e^{-\frac{1}{2}x_0^2/\sigma_n}$$

so we need roughly $\sigma_n > \sqrt{x_0}$ to have a chance that the second term is smaller than the first.

Applications

This result applies to

- Values of characteristic polynomials of unitary (symplectic, orthogonal) matrices;
- The "model" sums

$$\sum_{p \leq X} \log\Big(\frac{1}{1 - p^{-1/2}\Theta_p}\Big);$$

- And we conjecture it does for $\log \zeta(1/2 + it)$ (for $d = 2$). This would give a quantitative answer to Ramachandra's question: how large should T be before we can be sure to find $t \leq T$ with $\zeta(1/2 + it)$ in a given open ball in $\mathbf{C}$.