Central Limit Theorems and Lee-Yang Zeros

Joel L. Lebowitz
Rutgers University

We consider a sequence of increasing sets Λ_N, $N \to \infty$, and define the random variable $X_{\Lambda_N} \in \{0, 1, \ldots, N\}$ as the number of points (objects) in Λ_N. Λ_N could be a set in $\mathbb{R}^d, \mathbb{Z}^d$ or the edges of a graph G. In all cases, due to constraints the number of points is at most N. Let

$$\text{Prob}(X = m) = p_m z^m / P(z), \quad m = 0, 1, \ldots, N$$

where $z \geq 0$, $p_0 > 0$, $p_N > 0$ and

$$P(z) = \sum_{m=0}^{N} p_m z^m$$

(I have left out the dependence of X and p_m on Λ_N and have assumed $p_0 > 0$ to avoid dealing with zeros of $P(z)$ at $z = 0$.)

z is called the fugacity in statistical mechanics and $P(z)$, with the proper definition of the p_m, is the grand canonical partition function. In some cases we shall consider, z does not enter into the probabilities and we simply have $\text{Prob}(X = m) = p_m/P(1)$.

Using the fundamental theorem of algebra we write

$$P(z) = C \prod_{j=1}^{N} (z + \zeta_j)$$

where $\{-\zeta_j\}$ are the zeroes of $P(z)$. Note that for any finite N none of the zeroes can be on the positive real axis. They must be on the negative real axis or come in complex conjugate pairs.

It is clear from the definitions that

$$\langle X_\Lambda \rangle = \mu(z, \Lambda_N) = z \frac{\mathrm{d}}{\mathrm{d}z} \log P(z) = \sum_{j=1}^{N} \frac{z}{\zeta_j + z}.$$

Letting $\zeta_j = x_j + iy_j$ we can write

$$\mu(z, \Lambda_N) = \sum^{(0)} \frac{z}{x_j + z} + \sum^{(1)} \frac{2zx_j + 2z^2}{(z + x_j)^2 + y_j^2}$$

where the first sum is over all those zeros which lie on the negative real z-axis, $x_j > 0$, and the second sum is over all complex pairs. We also have

$$\mathrm{Var}(X_{\Lambda_N}) = \sigma^2(z, \Lambda_N) = z\frac{\mathrm{d}}{\mathrm{d}z}\langle X_{\Lambda_N}\rangle = \sum_j \frac{z}{z + \zeta_j}\left(1 - \frac{z}{z + \zeta_j}\right).$$

I am doing this separation for later convenience.

We shall be interested in the asymptotic normality of the fluctuations of the number of objects in Λ_N when $N \to \infty$. We say that our system satisfies a CLT if

$$\sup_{x\in\mathbb{R}} |F(x) - G(x)| \to 0, \quad \text{for } N \to \infty,$$

where

$$F(x) \equiv \frac{1}{P(z)} \sum_{m=0}^{\mu+x\sigma} p_m z^m = \text{Prob}\left\{\frac{X-\mu}{\sigma} < x\right\}$$

$$G(x) = (2\pi)^{-1/2} \int_{-\infty}^{x} e^{-u^2/2} \mathrm{d}u.$$

We say that our system satisfies the stronger local CLT (LCLT) if

$$\sup_{x\in\mathbb{R}}\left|F_\ell(x)-\frac{\mathrm{d}}{\mathrm{d}x}G(x)\right|\to 0,\quad \text{for} N\to\infty,$$

where

$$F_\ell(x)=\sigma\mathrm{Prob}\left(\frac{X-\mu}{\sigma}=x\right)$$

and the dependence on z and Λ_N has been suppressed.

Also suppressed is the fact that X can have only non-negative integer values.

I will now state three theorems which give sufficient conditions for a central limit theorem (CLT) or a stronger local CLT (LCLT) based on the location of the zeros of $P(z : \Lambda_N)$.

Theorem 1. (Harper 67, Bender 73, Canfield 80, Pitmann 97)

Suppose that the zeros of $P(z : \Lambda_N)$ are all on the negative real axis, i.e. $\zeta_j > 0$, and that for some $z_0 > 0$, $\sigma(z_0 : \Lambda_N) \to \infty$, when $N \to \infty$. Then the system satisfies a LCLT for $z = z_0$.

Extension: We actually prove a LCLT when $\text{Im}\zeta_j/\text{Re}\zeta_j \leq \sqrt{3}$ (Pittel, et al., 2014).

Theorem 2. (Pittel,. . . , 2014)

Suppose that $\text{Re}\zeta_j \geq 0$ and that $\sigma(z_0, \Lambda_N) \to \infty$. Then the system satisfies a CLT as $N \to \infty$ for $z = z_0$.

Extension: Under some additional conditions we prove a LCLT for the case $\text{Re}\zeta_j > 0$.

Theorem 3. (Iaglonitzer and Souillard, 79)

Let $\delta(z_0)$ be a disc of radius $\delta > 0$ centered around $z_0 > 0$ such that there are no zeros of $P(z)$ in $\delta(z_0)$. Then if $\sigma^2(z_0, \Lambda_N)/\mu^{2/3}(z_0, \Lambda_N) \to \infty$, the system satisfies a CLT for $z = z_0$.

Example 1: Determinantal point processes

Determinantal point processes are prominent structures in the theory of random matrices as well as in many other contexts. These are processes for which the k-point correlation function can be written as a $k \times k$ determinant,

$$\rho_{(k)}(x_1, \ldots, x_k) = \det[K(x_j, x_l)]_{j,l=1,\ldots,k},$$

where $K(x, y)$ — referred to as the correlation kernel — is independent of k. Here K (viewed as the kernel for an integral operator supported on Λ) is Hermitian, and all its eigenvalues are discrete and lie between zero and one.

The probability that there are exactly k points in Λ is given here bp $p_k/P(1)$. In terms of the correlation functions, the mean and variance $\mu(1,\Lambda)$ and $\sigma^2(1,\Lambda)$,

$$\begin{aligned}\mu(\Lambda) &= \int_\Lambda \rho_{(1)}(x)\,dx,\\ \sigma^2(1,\Lambda) &= \int_\Lambda dx_1 \int_\Lambda dx_2\,(\rho_{(2)}(x_1,x_2)\\ &\quad - \rho_1(x_1)\rho_1(x_2) + \rho_{(1)}(x_1)\delta(x_1-x_2)).\end{aligned}$$

It can be shown, using Mobius formula to compute Prob$(X_\Lambda = m)$ from the correlation functions (see e.g. Forrester book, 2010), that for any determinantal process with Hermitian kernel:

$$P(z;\Lambda) = C\prod_{l=0}^{\infty}(\zeta_l + z), \qquad \zeta_l = \frac{1-\lambda_l(J)}{\lambda_l(J)}$$

where the $\lambda_l(\Lambda)$ are eigenvalues of the integral operator K supported on Λ, with $\lambda_\ell \in (0,1)$ so the ζ_l are all on the negative real axis. Theorem 1 then gives a LCLT for all such a determinantal point processes (Forrester and L, 2014).

Random Matrices

One of the best known examples of a determinantal point process is given by the eigenvalues of the random matrices specified by the Gaussian Unitary Ensemble (GUE): a Gaussian probability measure on the space of complex $N \times N$ Hermitian matrices M with $\text{Prob}(M) \sim \exp(-\frac{1}{2}\text{Tr}MM^{\perp})$ which is unitary invariant and thus unchanged by conjugation by unitary matrices.

The density of eigenvalues $\{\lambda_j\}$ of M is given by

$$\prod(\lambda_j - \lambda_k)^{\beta} \exp\left(-\frac{\beta}{4}\sum_{n=1}^{N}\lambda_n^2\right),$$

where $\beta = 2$. By scaling the eigenvalues so that the mean density is unity and taking $N \to \infty$, one obtains a translation invariant determinantal point process specified by the so-called sine kernel $K(x,y) = \sin \pi(x-y)/\pi(x-y)$.

Costin and L. studied $X(\Lambda)$, $\Lambda \subset \mathbb{R}$, for the particular determinantal point process corresponding to the eigenvalues of the GUE in the limit $N \to \infty$, scaled so that the average scaling is 1 (bulk scaling limit) and thus specified by the sine kernel. They proved the CLT for this system. This was done by showing that as a consequence of the property that $\sigma_\Lambda \to \infty$ as $|\Lambda| \to \infty$, all cumulants of the characteristic function beyond the second vanish for $|\Lambda| \to \infty$.

In fact the proof makes no explicit use of the particular determinantal point process under consideration, requiring only that the corresponding kernel be locally trace class and self-adjoint, and that the variance tends to infinity. Theorem 1 then shows that the system actually satisfies a LCLT.

The proof by Costin and L. of a CLT for Gaussian Orthogonal Ensembles (GOE), $\beta = 1$, and Gaussian Symplectic Ensembles (GSE), $\beta = 4$, neither of whose eigenvalues are determinantal processes has been strengthened to a LCLT for the GSE but not for the GOE.

The eigenvalues of the Ginibre ensemble of non-Hermitian matrices with standard Gaussian complex entries give an example of a determinantal point process with a complex Hermitian kernel: in the limit $N \to \infty$ this is given by $K(w,z) = \frac{1}{\pi} e^{-(|w|^2+|z|^2)/2} e^{w\bar{z}}$ where z and w are complex. Here $\Lambda \subset \mathbb{R}^2$.

Spin polarized free fermions in dimension d provide examples of determinantal point processes in higher dimensions. For J a sphere of radius R, σ_J^2 / R^{d-1} is proportional to $\log R$ in the limit $R \to \infty$, and in particular σ_J^2 diverges in this limit so that we have a LCLT. The kernel $K(x,y)$ for each of these systems is proportional to the Bessel function $J_{d/2}(x-y)/|x-y|^{d/2}$.

There are other well known examples of determinantal processes in statistical physics and mathematics. A possible example, conditional upon the validity of the Montgomery-Odlyzko law, is the set of Riemann zeros for large modulus. The Montgomery-Odlyzko law states that certain statistical properties of the latter, upon appropriate scaling, coincide with the bulk scaled GUE and thus might form a determinantal point process. A proof that these zeros satisfy a LCLT is an open question, while the weaker statement of a CLT has now been proven (Bourgade, et al.) even without the Riemann hypothesis.

Example 2: Graph counting polynomials

Let G be a connected graph with a set E of edges each of which can be occupied by a particle or empty. We say that a configuration of occupied edges is admissible if the number of occupied edges attached to any vertex v, $d(v)$, is restricted to some set of natural numbers so that the maximal total number of occupied edges is N. When $d(v) = \{0, 1\}$, $P(z)$ corresponds to the "matching polynomial" of the graph. It was shown by Godsil (1981) that when $d(v) = \{0, 1\}$ then the system with Prob$(X = m) = p_m/P(1)$, the number of permissible configurations in which there are exactly m occupied edges, satisfies a LCLT when $N \to \infty$.

The proof is based on the result by Heilmann and Lieb (1972) that all the zeros of $P(z)$ lie on the negative real axis. Godsil's result then applies to $\text{Prob}(X = m) = p_m z_0^m / P(z_0)$ for any $z_0 > 0$. The occupied bonds are called in statistical mechanics dimers while the vertices to which no occupied edges are attached are called monomers. The variances of the number of occupied edges can be proven by the statistical mechanical methods described below to have the behavior $\sigma_N \geq CN$.

In current work by B. Pittel, L., D. Ruelle and E. Speer we consider the case of graph counting polynomials which arise when the restriction of $d(v) \in \{0, 1\}$ above is generalized to $d(v)$ some set of non-negative integers. We obtain a LCLT when $d(v) = \{0, 1, 2\}$ corresponding in statistical mechanics to "unbranched" polymers. The result makes use of the fact that all the zeros of $P(z)$ have negative real values* so that a CLT follows from Theorem 2. To obtain a LCLT one has to use some new ideas and verify some additional properties on the location of the zeros. (You should ask Boris to give a seminar on this.)

$*$ The proof of this follows from an extension of the work of Ruelle (99).

Before going on to give applications of Theorem 3, which come at this time exclusively from statistical mechanics, let me give proofs of Th. 1 and 2. I will begin with Theorem 2 from which an extended version of Th. 1 will follow using two additional inputs.

Proof of Theorem 2

Let us split up the roots of $P(z)$ as was already done before into those which are real (and therefore negative) and those which are complex. Then we can write

$$P(z) = C \prod_{\zeta_j < 0}^{(0)} (t_j + z) \prod_{\alpha}^{(1)} (|\zeta_\alpha|^2 + 2x_\alpha z + z^2), \quad z > 0,$$

where $t_j = \zeta_j$ when $\zeta_j > 0$ and ζ_α when ζ_j is complex and $x_\alpha = \mathrm{Re}\zeta_\alpha$. Each term $t_j + z = (t_j + z)\left(\frac{t_j}{t_j+z} + \frac{z}{t_j+z}\right)$ is the generating function of a Bernoulli random variable η_j,which takes the value 0 with probability $\frac{t_j}{t_j+z}$ and the value 1 with probability $\frac{z}{t_j+z}$.

Similarly each term $\zeta_\alpha^2 + 2x_\alpha z + z^2$ is, for $x_\alpha \geq 0$, the generating function for a random variable $\eta_\alpha = \{0, 1, 2\}$ with probabilities

$$w_\alpha(0) = \frac{|\zeta_\alpha|^2}{|z + \zeta_\alpha|^2}, \quad w_\alpha(1) = \frac{2x_\alpha z}{|z + \zeta_\alpha|^2}, \quad w_\alpha(2) = \frac{z^2}{|z + \zeta_\alpha|^2}.$$

Thus X can be written as a sum of independent, non-identical, random variables η_j and η_α. It follows then from results in Feller that X will satisfy a CLT whenever $\mathrm{Var}(X_{\Lambda_N}) \to \infty$ as $N \to \infty$. This proves Theorem 2.

To prove the extension of Theorem 2 giving a LCLT when $\mathrm{Re}\zeta_j > 0$ requires additional work and conditions.

Proof of Theorem 1

To go from a CLT to a LCLT it is sufficient (Bender, 1973) that the p_m have the log concavity property, i.e. $p_m^2 \geq p_{m-1}p_{m+1}$. When all the roots are on the negative axis concavity of $P(z)$ was proven by I. Newton, which proves Theorem 1.

To prove the extension of Th. 1 we use a result by Menon (1969) that the product of concave polynomials is also concave. This is obviously true for the η_j since $p_2 = 0$ and can be checked to be true for the generating function of the random variables η_α, t_α whenever $\mathrm{Im}\zeta_\alpha/\mathrm{Re}\zeta_\alpha \leq \sqrt{3}$.

Lee-Yang zeros in Statistical Mechanics

The pressure of an equilibrium system of atoms (molecules) in a domain $\Lambda \subset \mathbb{R}^d$ ($\mathbb{Z}^d$) at a physical fugacity $z > 0$ (z is proportional to the exponential of the chemical potential) is given by

$$\Pi(\beta, z; \Lambda) = |\Lambda|^{-1} \log P(\beta, z; \Lambda)$$

where

$$P(\beta, z; \Lambda) = \sum_{m=0}^{N_{\text{max}}} z^m Q(\beta, m; \Lambda)$$

is the "grand canonical" partition function and $Q(\beta, m; \Lambda)$ is the "canonical" partition function. β is the reciprocal temperature.

Restricting ourselves for simplicity to classical lattice systems, $\Lambda \subset \mathbb{Z}$, $N_{\text{max}}(\Lambda) \leq |\Lambda|$, we have

$$Q(\beta, m; \Lambda) = \sum_{\underline{\eta}} \exp[-\beta U_{\Lambda}(\underline{\eta})]$$

where

$$\underline{\eta} = \{\eta(x); x \in \Lambda; \eta(x) = \{0, 1\}\} .$$

A typical energy function $U_{\Lambda}(\underline{\eta})$ is given by

$$U_{\Lambda}(\underline{\eta}) = - \sum_{x,y \in \Lambda} J(x - y)\eta(x)\eta(y) \quad (+ \text{ boundary terms})$$

with $J(\underline{r})$ some rapidly decaying function of the distance, e.g. $J(\underline{r}) = J\delta(|\underline{r}| - 1)$. From the definition, $Q(\beta, 0; \lambda) = 1$ and $Q(\beta, 1; \lambda) = |\Lambda|$.

As already noted earlier, for finite Λ there can be no zeros at $z \geq 0$, the physically relevant value of the fugacity. This means that $\Pi(\beta, z; \Lambda)$ is analytic for all fugacities and there can be no "phase transitions", defined as non-analycities in the pressure as a function of z.

The situation is different in the thermodynamic limit $\Lambda \to \mathbb{Z}^d$. This is the right model for a macroscopic system containing, say 10^{23} atoms, when we are not considering surface effects. In this limit the pressure is given by

$$\Pi(\beta, z) = \lim_{\Lambda \to \mathbb{Z}^d} \frac{\log P(\beta, z; \Lambda)}{|\Lambda|}.$$

The zeros of $P(\beta, z; \Lambda)$ can now approach the positive z-axis and cause singularities in the pressure.

That this actually happens for some values of β can be proven rigorously in some cases. The most famous of these cases occurs when $J(\underline{r}) < 0$, when Lee ad Yang proved that the zeros of $P(\beta, z; \Lambda)$ lie on a circle, $|z| = R$ - corresponding to zero magnetic field in Ising spin language, i.e. $|e^{-\beta h}| = 1$. When this is combined with Onsager's proof for nearest neighbor ferromagnetic interactions, of a singularity in $\Pi(\beta, z; \Lambda)$ for $\beta > \beta_{\text{c}}$ this gives a finite density of zeros at $|z| = R$.

The proof of the existence of the thermodynamic limit under quite general conditions on the interactions goes back to the middle of the last century (Ruelle, Fisher, van Hove) Ruelle, Penrose, Greenweld also show that $\prod(\beta, z, \Lambda) = \sum b_m z^m$.

It follows from the definitions that

$$\langle X_\Lambda \rangle = z \frac{\mathrm{d} \log P(\beta, z, \Lambda)}{\mathrm{d}z} = |\Lambda| z + \mathcal{O}(z^2)$$

and that

$$\frac{\mathrm{d}\langle X_\Lambda \rangle}{\mathrm{d}z} = \mathrm{Var}(X_\Lambda) \geq 0.$$

Now for small enough z, $\langle X_\Lambda \rangle \sim z|\Lambda|$ so $\langle X_\Lambda \rangle \geq c|\Lambda|$ for all $z \geq \varepsilon > 0$.

One can also prove that the variance of X_Λ goes to ∞ as $|\Lambda| \to \infty$ under quite general conditions. The proof is due to Ginibre (1967).

Let $P(X_\Lambda = m) = P_m/m!$. If $P_j = 0$ then P_{j+k} for all $k \geq 0$. Then if

$$P_{n+2}/P_{n+1} \geq P_{n+1}/P_n - A, \quad A > -1 \qquad (*)$$

one has that

$$\text{Var}(X_\Lambda) = \langle X_\Lambda \rangle/(1 + A).$$

The proof is elementary. Ginibre then shows that for pair interaction satisfying some simple conditions $(*)$ holds. In work with Boris we extended this to graph counting polynomials whenever $d(v) = \{0, 1, 2, \ldots, k\}$ (with no gaps).