An introduction to asymptotic enumeration

Part 1. Why are asymptotic results interesting?

Part 2. Asymptotics of sequences $a(n)$

Part 3. Limit laws

The study of additional statistics

$$\begin{array}{rccc} size : & \mathcal{A} & \to & \mathbb{N} \\ & a & \mapsto & |a| \end{array} \qquad \begin{array}{rccc} s : & \mathcal{A} & \to & \mathbb{Z} \\ & a & \mapsto & s(a) \end{array}$$

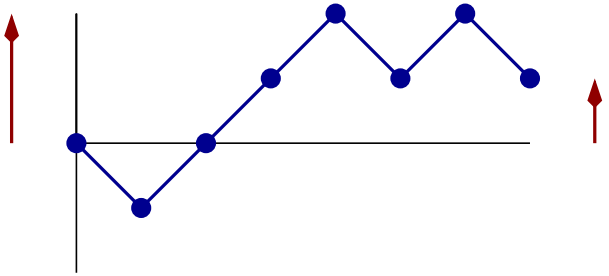
When objects are taken uniformly in $\mathcal{A}_n$, the statistic s becomes a random variable S_n:

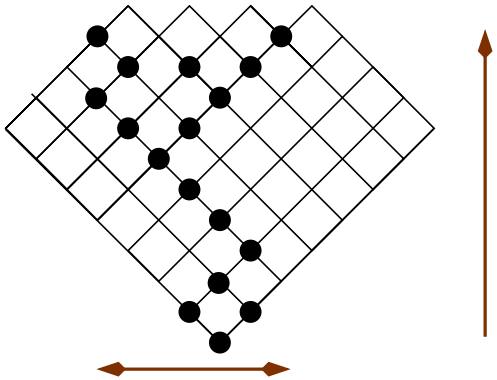
$$\mathbb{P}(\mathsf{S}_n = k) = \frac{a(n,k)}{a(n)}$$

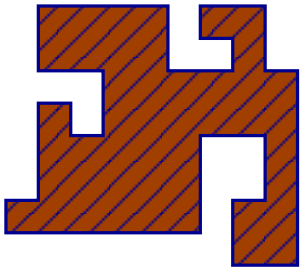
where $a(n,k)$ is the number of objects of size n for which the additional statistic s equals k.

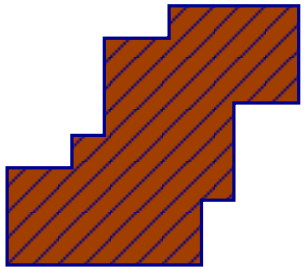
Examples of additional statistics

- the final height of a 1D walk with n steps,
- the maximal height of such a walk,
- the width/height of a directed animal with n cells,
- the area enclosed by a SAP of perimeter $2n$,
- the area enclosed by a staircase polygon of perimeter $2n$...





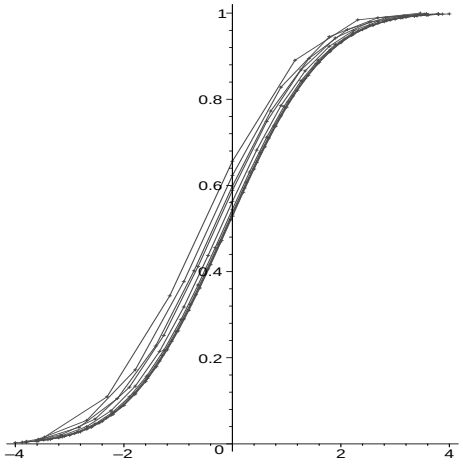




Convergence in law (or: in distribution)

Definition. The random variables S_n converge in law to the random variable S, having distribution function $F(x)$, if for all x [where F is continuous],

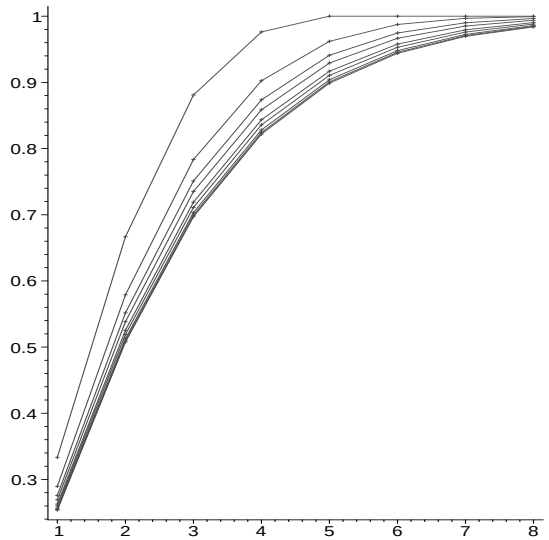
$$\mathbb{P}(\mathsf{S}_n \leq x) = F_n(x) \to F(x) = \mathbb{P}(S \leq x).$$



A discrete limit law: contacts in a Dyck path

Convergence of the distribution function of S_n:

$$\mathbb{P}(\mathsf{S}_n \leq \ell + 1) \to \sum_{k=1}^{\ell} \frac{k}{2^{k+1}}$$

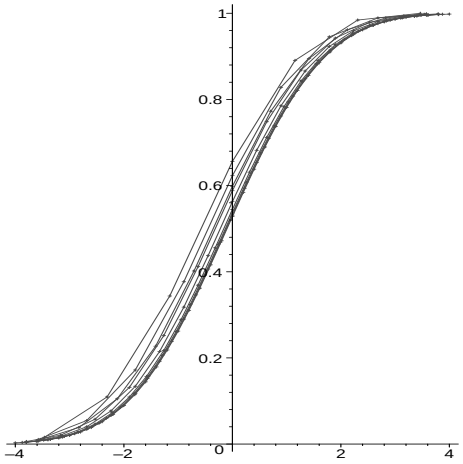


A continuous limit law: The final height of a random walk

As $\mathbb{E}(S_n^2) = n$, we consider the random variable $S_n/\sqrt{n}$.
Its distribution function satisfies

$$\mathbb{P}\left(\frac{S_n}{\sqrt{n}} \leq x\right) \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} dt$$

$\Rightarrow$ Convergence to a Gaussian law



How to prove convergence in law

- From the distribution function
- Pumping moments
- Discrete limit laws: pointwise convergence of the probability generating function
- Continuous limit laws: pointwise convergence of the Laplace or Fourier transform

1. Convergence of the moments and convergence in law

Theorem. For S_n to converge in law to S, it suffices that

- all moments converge : $\mathbb{E}(\mathrm{S}_n{}^j) \to \mathbb{E}(S^j)$ for all j
- the Laplace transform

$$\mathbb{E}(e^{xS}) = \sum_{j \geq 0} \mathbb{E}(S^j) \frac{x^j}{j!}$$

has a positive radius of cv.

Where do we find the moments of the statistic s?

Bivariate generating function:

$$A(t,u) = \sum_{a \in \mathcal{A}} t^{|a|} u^{s(a)} = \sum_n t^n \left(\sum_k a(n,k) u^k \right)$$

Specialization $u = 1$:

$$A(t,1) = \sum_{a \in \mathcal{A}} t^{|a|} = \sum_n a(n) t^n$$

- Average value of S_n:

$$u \frac{\partial}{\partial u} A(t,u) = \sum_n t^n \left(\sum_k k a(n,k) u^k \right) \Rightarrow u \frac{\partial}{\partial u} A(t,u) \Big|_{u=1} = \sum_n t^n a(n) \mathbb{E}(\mathsf{S}_n)$$

Where do we find the moments of the statistic s?

Bivariate generating function:

$$A(t,u) = \sum_{a \in \mathcal{A}} t^{|a|} u^{s(a)} = \sum_{n,k} a(n,k) t^n u^k$$

Specialization $u = 1$:

$$A(t,1) = \sum_{a \in \mathcal{A}} t^{|a|} = \sum_{n} a(n) t^n$$

- Average value of S_n:

$$u\frac{\partial}{\partial u} A(t,u) = \sum_n t^n \left(\sum_k k a(n,k) u^k \right) \Rightarrow u\frac{\partial}{\partial u} A(t,u)\Big|_{u=1} = \sum_n t^n a(n) \mathbb{E}(\mathsf{S}_n)$$

- Further moments:

$$\left(u\frac{\partial}{\partial u}\right)^i A(t,u) = \sum_n t^n \left(\sum_k k^i a(n,k) u^k \right) \Rightarrow \left(u\frac{\partial}{\partial u}\right)^i A(t,u)\Big|_{u=1} = \sum_n t^n a(n) \mathbb{E}(\mathsf{S}_n{}^i)$$

Where do we find the moments of the statistic s?

Bivariate generating function:

$$A(t,u) = \sum_{a \in \mathcal{A}} t^{|a|} u^{s(a)} = \sum_{n,k} a(n,k) t^n u^k$$

Specialization $u = 1$:

$$A(t,1) = \sum_{a \in \mathcal{A}} t^{|a|} = \sum_{n} a(n) t^n$$

- Average value of S_n:

$$u\frac{\partial}{\partial u} A(t,u) = \sum_n t^n \left(\sum_k k a(n,k) u^k \right) \Rightarrow u\frac{\partial}{\partial u} A(t,u)\Big|_{u=1} = \sum_n t^n a(n) \mathbb{E}(\mathsf{S}_n)$$

- Further moments:

$$\left(u\frac{\partial}{\partial u}\right)^i A(t,u) = \sum_n t^n \left(\sum_k k^i a(n,k) u^k \right) \Rightarrow \left(u\frac{\partial}{\partial u}\right)^i A(t,u)\Big|_{u=1} = \sum_n t^n a(n) \mathbb{E}({\mathsf{S}_n}^i)$$

- The exponential moments:

$$\frac{\partial^i A}{\partial u^i}(t,1) = \sum_n t^n a(n)\ \mathbb{E}(\mathsf{S}_n(\mathsf{S}_n - 1)\cdots(\mathsf{S}_n - i + 1))$$

Contacts in bilateral Dyck paths

- GF of paths counted by their half-length (t) and number of visits at 0 (u):

$$A(t,u) = \frac{u}{1-2uP(t)} \qquad \text{with} \qquad P(t) = \frac{1-\sqrt{1-4t}}{2}$$

$$\Rightarrow A(t,1) = \frac{1}{\sqrt{1-4t}} = \sum_{n\geq 0} \binom{2n}{n} t^n = \sum_n a(n) t^n$$

Contacts in bilateral Dyck paths

- GF of paths counted by their half-length (t) and number of visits at 0 (u):

$$A(t,u) = \frac{u}{1-2uP(t)} \qquad \text{with} \qquad P(t) = \frac{1-\sqrt{1-4t}}{2}$$

$$\Rightarrow A(t,1) = \frac{1}{\sqrt{1-4t}} = \sum_{n\geq 0} \binom{2n}{n} t^n = \sum_n a(n) t^n$$

- Take derivatives w.r.t. u

$$\frac{\partial^i A}{\partial u^i}(t,1) = \frac{i!\ (2P(t))^i}{(1-4t)^{(i+1)/2}} = \sum_n t^n a(n)\ \mathbb{E}(\mathsf{S}_n(\mathsf{S}_n-1)\dots(\mathsf{S}_n-i+1))$$

Contacts in bilateral Dyck paths

- GF of paths counted by their half-length (t) and number of visits at 0 (u):

$$A(t,u) = \frac{u}{1-2uP(t)} \quad \text{with} \quad P(t) = \frac{1-\sqrt{1-4t}}{2}$$

$$\Rightarrow A(t,1) = \frac{1}{\sqrt{1-4t}} = \sum_{n\geq 0} \binom{2n}{n} t^n = \sum_n a(n) t^n$$

- Take derivatives w.r.t. u

$$\frac{\partial^i A}{\partial u^i}(t,1) = \frac{i!\ (2P(t))^i}{(1-4t)^{(i+1)/2}} = \sum_n t^n a(n)\ \mathbb{E}(\mathsf{S}_n(\mathsf{S}_n-1)\dots(\mathsf{S}_n-i+1))$$

- Singularity analysis:

$$\mathbb{E}(\mathsf{S}_n(\mathsf{S}_n-1)\dots(\mathsf{S}_n-i+1)) \sim \frac{i!\Gamma(1/2)}{\Gamma((i+1)/2)}\ n^{i/2}$$

$$\Rightarrow \mathbb{E}\left(\left(\frac{\mathsf{S}_n}{\sqrt{n}}\right)^i\right) \sim \frac{i!\Gamma(1/2)}{\Gamma((i+1)/2)}$$

These are the moments of the Rayleigh law of density $x/2\exp(-x^2/4)$ for $x>0$

$\Rightarrow$ Convergence in law to a Rayleigh distribution

The area of staircase polygons

- GF of staircase polygons by half-perimeter (t) and area (u):

$$P(t,u) = t^2u + 2tuP(t,u) + P(t,u)P(tu,u)$$

- Repeated differentiations of the equation

$\Rightarrow \frac{\partial^i P}{\partial u^i}(t,1)$ is always algebraic and

$$\frac{\partial^j P}{\partial u^j}(t,1) = \sum_n t^n a(n)\ \mathbb{E}(\mathsf{S}_n(\mathsf{S}_n-1)\dots(\mathsf{S}_n-j+1)) \sim \frac{c_j}{(1-4t)^{(3j-1)/2}}$$

with

$$c_0 = -1/2, \quad c_j = \frac{j(3j-4)}{8}c_{j-1} + \sum_{i=1}^{j-1}\binom{j}{i}c_i c_{j-i}$$

- Singularity analysis

$$\mathbb{E}\left(\left(\frac{\mathsf{S}_n}{n^{3/2}}\right)^j\right) \sim \frac{c_j}{c_0}\ \frac{\Gamma(-1/2)}{\Gamma((3j-1)/2)}$$

Moments of the Airy distribution $\Rightarrow$ Convergence in law

2. Continuity theorem for discrete limit laws

Theorem. For S_n to converge in law to S, it suffices that the probability generating function of S_n converges pointwise to that of S:

$$G_n(u) := \sum_k \mathbb{P}(S_n = k)u^k \to G(u) := \sum_k \mathbb{P}(S = k)u^k$$

for u in a set Ω having an accumulation point of modulus < 1.

Example: $\Omega = [0, 1/2]$

Where do we find the probability g.f. of the statistic s?

- Probability GF:

$$G_n(u) := \sum_k \mathbb{P}(S_n = k)u^k = \frac{1}{a(n)} \sum_k a(n,k)u^k$$

- Bivariate generating function:

$$A(t,u) = \sum_n t^n \left(\sum_k a(n,k)u^k \right) = \sum_n t^n a(n) G_n(u)$$

Specialization $u = 1$:

$$A(t,1) = \sum_{a \in \mathcal{A}} t^{|a|} = \sum_n a(n)t^n$$

$\Rightarrow$ Estimate the coefficient of t^n in $A(t,u)$, where u is a fixed real or complex number

$\Rightarrow$ Recycle the methods used for the analysis of $A(t,1)$!

Contacts in Dyck paths

- Let $A(t,u)$ count Dyck paths by the half-length (t) and contacts (u):

$$A(t,u) = \frac{u}{1-u(1-\sqrt{1-4t})/2}$$

- Fix u in $[0,1]$. Then, as $t \to 1/4$:

$$A(t,u) = \frac{2}{2-u}\left(1 - \frac{u}{2-u}\sqrt{1-4t}\right) + O(1-4t)$$

Contacts in Dyck paths

- Let $A(t,u)$ count Dyck paths by the half-length (t) and contacts (u):

$$A(t,u) = \frac{u}{1-u(1-\sqrt{1-4t})/2}$$

- Fix u in $[0,1]$. Then, as $t \to 1/4$:

$$A(t,u) = \frac{2}{2-u}\left(1 - \frac{u}{2-u}\sqrt{1-4t}\right) + O(1-4t)$$

- Singularity analysis:

$$[t^n]A(t,u) = a(n)G_n(u) \sim -\frac{2u}{(2-u)^2}\frac{4^n}{\Gamma(-1/2)}n^{-3/2}$$

$$G_n(u) \sim \frac{u}{(2-u)^2} = \sum_{k\geq 1}\frac{k}{2^{k+1}}\,u^k = G(u)$$

Convergence in law to the distribution given by

$$\mathbb{P}(S=k) = \frac{k}{2^{k+1}}$$

The degree of the root in a Cayley tree

- Exponential GF of labelled rooted trees by vertices (t) and degree of the root (u)

$$A(t,u) = t\exp(uT(t)) \quad \text{with} \quad T(t) = t\exp(T(t))$$

- For $u > 0$, $A(t,u)$ has a unique dominant singularity at $t = 1/e$, in the vicinity of which

$$A(t,u) = e^{u-1} - ue^{u-1}\sqrt{2}\sqrt{1-te}\,(1+o(1))$$

- Singularity analysis:

$$[t^n]A(t,u) = \frac{a(n)}{n!}G_n(u) \sim -ue^{u-1}\sqrt{2}\frac{e^n}{\Gamma(-1/2)}n^{-3/2}$$

$$G_n(u) \sim ue^{u-1} = \sum_{k\geq 1}\frac{u^k}{e(k-1)!} = G(u)$$

Distribution function of a Poisson random variable

3. Continuity theorem for continuous limit laws

Theorem 1. For the sequence S_n to converge in law to S, it suffices that the Fourier transform (or characteristic function) converges pointwise

$$\mathbb{E}(e^{ixS_n}) \to \mathbb{E}(e^{ixS})$$

for all x in $\mathbb{R}$.

Theorem 2. For the sequence S_n to converge in law to S, it suffices that the Laplace transform converges pointwise

$$\mathbb{E}(e^{xS_n}) \to \mathbb{E}(e^{xS}))$$

for all x in a real neighborhood of 0.

Where do we find the Laplace transform of the statistic s?

Bivariate generating function:

$$A(t,u) = \sum_n t^n \left(\sum_k a(n,k) u^k \right) = \sum_n t^n a(n)\ \mathbb{E}\left(u^{\mathsf{S}_n}\right)$$

$$\Rightarrow a(n)\ \mathbb{E}\left(e^{x\mathsf{S}_n}\right) = [t^n] A(t, e^x)$$

$\Rightarrow$ Estimate the coefficient of t^n in $A(t,u)$, for u close to 1

The number of cycles in permutations

- From Gilbert's lectures:

$$A(t,u) = \sum_n \frac{t^n}{n!} \sum_{|\sigma|=n} u^{c(\sigma)} = \frac{1}{(1-t)^u}$$

- Moments:

$$\frac{\partial^i A}{\partial u^i}(t,1) = \sum_n t^n \ \mathbb{E}(\mathsf{S}_n(\mathsf{S}_n-1)\dots(\mathsf{S}_n-i+1)) = \left(\log\frac{1}{1-t}\right)^i \frac{1}{1-t}$$

- Singularity analysis:

$$\mathbb{E}(\mathsf{S}_n(\mathsf{S}_n-1)\dots(\mathsf{S}_n-i+1)) \sim (\log n)^i$$

$$\Rightarrow \mathbb{E}({\mathsf{S}_n}^i) \sim (\log n)^i$$

$$\frac{S_n}{\log n} \xrightarrow{d} 1$$

Concentration phenomenon. What are the fluctuations around $\log n$?

The number of cycles in permutations (contd.)

- Normalize S_n. As $\mathbb{E}(S_n) \sim \log n$ and $\mathbb{E}((S_n - \log n)^2) \sim \log n$, define

$$\tilde{S}_n := \frac{S_n - \log n}{\sqrt{\log n}}$$

The number of cycles in permutations (contd.)

- Normalize S_n. As $\mathbb{E}(S_n) \sim \log n$ and $\mathbb{E}((S_n - \log n)^2) \sim \log n$, define

$$\tilde{S}_n := \frac{S_n - \log n}{\sqrt{\log n}}$$

- Study the Laplace transform of $\tilde{S}_n$:

$$\mathbb{E}(e^{x\tilde{S}_n}) = e^{-x\sqrt{\log n}}\ \mathbb{E}\left((e^{x/\sqrt{\log n}})^{S_n}\right) = e^{-x\sqrt{\log n}}\ \mathbb{E}(u^{S_n})$$

with $u = e^{x/\sqrt{\log n}}$.

The number of cycles in permutations (contd.)

- Normalize S_n. As $\mathbb{E}(S_n) \sim \log n$ and $\mathbb{E}((S_n - \log n)^2) \sim \log n$, define

$$\tilde{S}_n := \frac{S_n - \log n}{\sqrt{\log n}}$$

- Study the Laplace transform of $\tilde{S}_n$:

$$\mathbb{E}(e^{x\tilde{S}_n}) = e^{-x\sqrt{\log n}}\ \mathbb{E}\left((e^{x/\sqrt{\log n}})^{S_n}\right) = e^{-x\sqrt{\log n}}\ \mathbb{E}(u^{S_n})$$

with $u = e^{x/\sqrt{\log n}}$.

- Generating function:

$$A(t,u) = \sum_n t^n\ \mathbb{E}(u^{S_n}) = \frac{1}{(1-t)^u} \quad \Rightarrow \mathbb{E}(u^{S_n}) = \frac{\Gamma(n+u)}{n!\Gamma(u)}$$

$$\mathbb{E}(e^{x\tilde{S}_n}) \sim \exp(x^2/2)$$

Laplace transform of the gaussian distribution $\mathcal{N}(0,1)$

Conclusion:

$$\tilde{S}_n := \frac{S_n - \log n}{\sqrt{\log n}} \xrightarrow{d} \mathcal{N}(0,1)$$

Patterns and universality in limit laws

"If the position and nature of the singularity do not change when u varies around 1, one expects a discrete limit law"

Sub-critical composition schema for alg-log series

If $A(t,u) = C(uB(t))$ with $B(\rho_B) < \rho_C$, the first singularity of $A(t,u)$ occurs at ρ_B and we have a discrete limit law with probability g.f.:

$$G(u) = \sum \mathbb{P}(S=k)u^k = \frac{uB'(\tau u)}{B'(\tau)}$$

where $\tau = B(\rho_B)$.

Examples:

- Contacts in Dyck paths

$$A(t,u) = \frac{1}{1-u(1-\sqrt{1-4t})/2}$$

- When C is the Sequence species ($C(z) = 1/(1-z)$), convergence to a negative binomial law
- When C is the Set species ($C = \exp$), convergence to a Poisson law

Patterns and universality in limit laws

"If the position or nature of the singularity changes smoothly around $u = 1$, one expects a gaussian limit law"

The quasi-power theorem

Assume that, uniformly in u taken in a neighborhood of 1, the probability g.f. satisfies

$$G_n(u) = \frac{1}{a(n)}[t^n]A(t,u) = \kappa(u)(\mu(u))^{\beta_n}\,(1 + O(\epsilon_n))$$

where $\beta_n \to \infty$ and $\epsilon_n \to 0$, $\kappa(u)$ and $\mu(u)$ are analytic at 1 (+ variability condition)

Let $\mu_n = \mathbb{E}(S_n)$ and $\sigma_n^2 = \mathbb{E}((S_n - E(S_n))^2)$. Then the normalized variable

$$\frac{S_n - \mu_n}{\sigma_n}$$

converges to a gaussian law $\mathcal{N}(0,1)$

Example: Number of cycles in a permutation

$$G_n(u) \sim \frac{n^{u-1}}{\Gamma(u)} \sim \frac{1}{\Gamma(u)}\,(\exp(u-1))^{\log n}$$

Patterns and universality in limit laws

The most interesting cases are those such that the position or nature of the singularity of $A(t,u)$ changes non-smoothly around $u = 1$.

Example. The area of staircase polygons

$$A(t,u) = t^2u + 2tuA(t,u) + A(t,u)A(tu,u)$$

	Type of singularity	
$0 < u < 1$	simple pole	
$u = 1$	square root	
$u > 1$	radius 0	

Patterns and universality in limit laws

The most interesting cases are those such that the position or nature of the singularity of $A(t,u)$ changes non-smoothly around $u=1$.

Example. The area of staircase polygons

$$A(t,u) = t^2u + 2tuA(t,u) + A(t,u)A(tu,u)$$

	Type of singularity	Average area in polygons weighted by $u^{a(p)}$
$0<u<1$	simple pole	κn
$u=1$	square root	$\kappa n^{3/2}$
$u>1$	radius 0	$n^2/4$

Random staircase polygons of perimeter $2n$ with probability

$$\text{prob}(p) = \frac{1}{Z_n(u)} u^{a(p)}$$

These cases precisely correspond to phase transitions in statistical mechanics