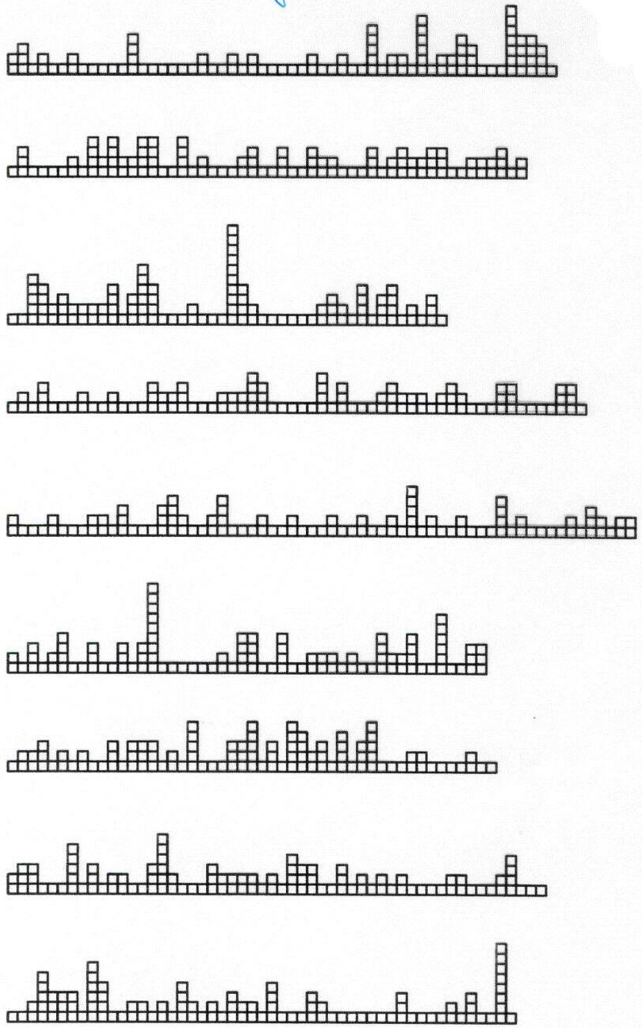
RANDOM STRUCTURES

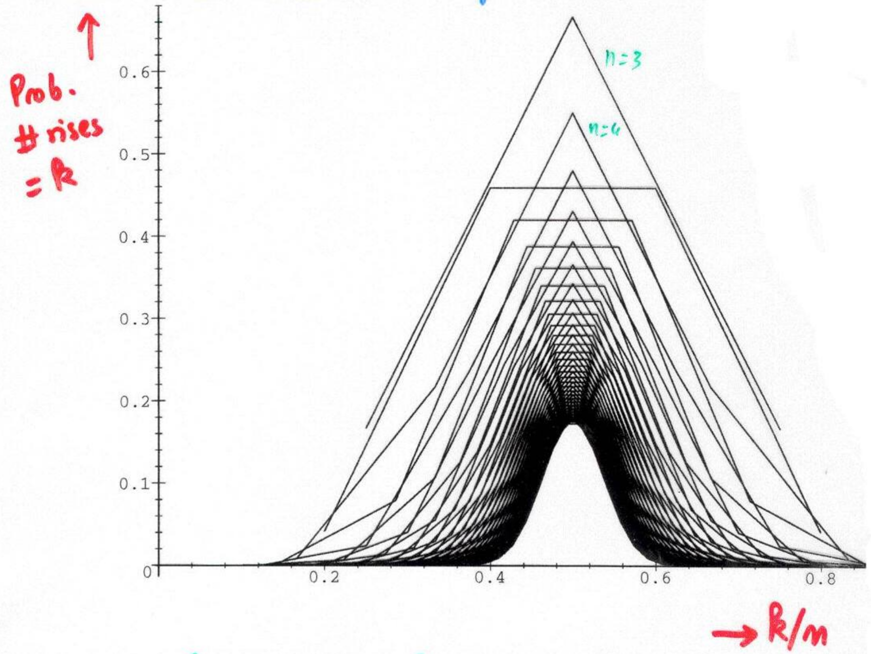
Chapter 9: Multivariate asymptotics and limit distributions.

RANDOM COMPOSITIONS of $n=100$.



⇝ Concentration of distribution ?

Eulerian distribution = rises in permutations



Gaussian limit ?

DE MOIVRE LAPLACE GAUSS

$$\Phi(x) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-w^2/2}\, dw$$

WHY is the binomial distribution asymptotically GAUSSIAN ?

- De Moivre: Stirling's approximation: $\frac{1}{2^n}\binom{n}{k}$
- {Gauss, Laplace}: As a sum of a large number of RV's

$\leadsto$ Probability generating function: $X \leadsto \mathbb{E}(u^X) = \sum_k \Pr(X=k)u^k$

- BINOMIAL $\longrightarrow$ GAUSSIAN LAW

$$f_n(u) = \left(\frac{1+u}{2}\right)^n \implies \text{Normal}$$

- LARGE POWERS $\rightarrow$ GAUSSIAN LAW

$$f_n(u) = g(u)^n \implies \text{Normal}$$

(+ $\exists$ mean, variance)

— Central Limit Theorem —

- QUASI-POWERS $\rightarrow$ GAUSSIAN LAW

$$f_n(u) \sim A(u).\, B(u)^{\beta_n} \implies \text{Normal}$$

Typical case: $f_n(u)$ arises from a bivariate generating function

$$f_n(u) = \frac{[z^n]\, C(z,u)}{[z^n]\, C(z,1)}.$$

QUASI-POWERS THEOREM ... Bender/Richmond ... HWANG!

(X_n) are RV's with probability generating function $f_n(u)$.

$u \simeq 1$

$$\boxed{f_n(u) = A(u)\, B(u)^{\beta_n} \left(1 + O\left(\frac{1}{\kappa_n}\right)\right)}$$

for u in complex neighbourhood of 1

$\beta_n, \kappa_n \to \infty$; $A(u), B(u)$ analytical 1 (+ details: $\mathrm{Var}(B(u)) > 0$)

$\mu_n \equiv \mathbb{E}(X_n) \sim \beta_n B'(1)$; $\sigma_n^2 \equiv \mathbb{V}(X_n) \sim \beta_n \cdot C^{st}$

$$\boxed{\Pr(X_n \le \mu_n + x\sigma_n) \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-w^2/2}\, dw}$$

SPEED IS $O(\kappa_n^{-1} + \beta_n^{-1/2})$.

Proof involves

- Continuity theorem for characteristic functions ≡ Lévy's Theorem.

$$\phi_X(t) := \mathbb{E}(e^{itX})$$ characteristic function

$\equiv \mathrm{PGF}(e^{it})$ for discrete RV's.

Continuity theorem: If $\phi_{X_n}(t) \to \phi_Y(t)$ pointwise for all t's, then

$$X_n \xRightarrow{D} Y$$

ie. for cumulative distribution functions [cdf]

$\forall x$: $F_{X_n}(x) \to F_Y(x)$ [at points of continuity].

+ **Berry - Esseen**: if characteristic functions are close, then cumulative distribution functions are also close. [see Feller]

The supercritical sequence $\mathcal{F} = \mathrm{Seq}(\mathcal{G})$

$$F(z,u) = \frac{1}{1-u\,G(z)}$$

Assume $G(r) > 1$ where r = radius of conv. of G.

Theorem: The # of $\mathcal{G}$-components in a large $\mathcal{F}$-structure is asympt. Normal

Proof: Let $\rho \in (0,r)$ be such that $G(\rho) = 1$.

- Equation $1-u\,G(z) = 0$ has root $\rho(u)$ where $\rho(u)$ depends analytically on u for u near 1
- $F(z,u)$ with u a param. has simple pole coeff $[z^n]\,F(z,u) \sim c(u)\,\rho(u)^{-n}$

Quasi Powers Theorem applies!

Applications:

- Integer compositions of all sorts
- Surjections aka preferential arrangements.
- Any structure "driven" by a SEQUENCE in a supercritical way (ie $G(\rho)>1$).

For a large collection of combinatorial classes & parameters, we have a functional equation

$$\Phi(z, y, u) = 0$$

In the counting case $(u=1)$ get a singular expansion

$$y(z,1) = \cdots (1 - z/\rho)^{\alpha} + \cdots\cdots$$

A PERTURBATION of u near 1 will often induce a smooth perturbation of the expansion of $y(z,u)$, e.g.,

movable singularity $y(z,u) = \cdots (1 - z/\rho(u))^{\alpha} + \cdots$

movable exponent $y(z,u) = \cdots (1 - z/\rho)^{\alpha(u)} + \cdots$

with $\rho(u)$ or $\alpha(u)$ analytic at 1

$\Longrightarrow$ Asymptotic normality $\left\{ \begin{array}{l} \text{by singularity analysis} \\ \text{+ Quasi Powers} \end{array} \right.$

- The path-in-graphs (aka automata) framework $\Rightarrow$ | singularity moves analytically, exponent remains -1

$\frac{A(z,u)}{B(z,u)}$ ($\rightarrow$ cf. also Markov chain theory).

e.g. # occurrences in words of a fixed pattern

- For algebraic systems, e.g., simple families of trees and local parameters (e.g., # leaves) $\Rightarrow$ | singularity moves, exponent $\equiv \frac{1}{2}$

Drmota-Lalley-Woods Theorem

$\Rightarrow$ Asymptotic Normality –

$\boxed{\text{Local configurations in random structures are almost always NORMALLY DISTRIBUTED}}$

B. Vallée = extension to dynamical analysis

THE EXP-LOG SCHEME

F+ Soria, 1990 ≡ fixed singularity movable exponent

THEOREM $\mathcal{F} = \mathrm{Set}(\bullet\, \mathcal{G}) \Longrightarrow$ (Normality of the number of components)

EXP

LOGARITHMIC SINGULARITY

- EXAMPLE Cycles in Permutations [Gončarov 1943]

$$F(z,u) = (1-z)^{-u}$$
$$= \exp\left(u \log \frac{1}{1-z}\right)$$

By singularity analysis

$$[z^n]\, F(z,u) = \frac{n^{u-1}}{\Gamma(u)}\left(1 + O\left(\frac{1}{n}\right)\right)$$

POINTWISE for EACH u, but also $\boxed{\text{UNIFORMLY}}$

$$[z^n]\, F(z,u) = \frac{1}{\Gamma(u)} \left(e^{u-1}\right)^{\log n} \left[1 + O\left(\frac{1}{n}\right)\right]$$

= A QUASI-POWERS APPROXIMATION, $\beta_n = \log n$

$\Longrightarrow$ Limiting Gaussian law

— 2.7 —

$\boxed{\text{EXAMPLE}}$) Polynomials over finite fields

— are a sequence of coefficients

$\Longrightarrow$ GF has a pole

Coeffs grow like q^n

— are a set of irreducibles (primes)

$\Longrightarrow$ GF has a LOG sing.

Coeffs grow like $\dfrac{q^n}{n}$

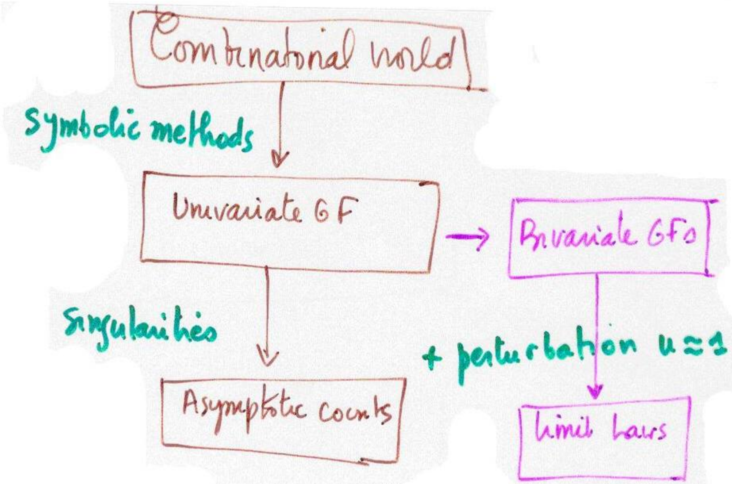
A Prime Number Theorem: *The density of irreducibles is* $\dfrac{1}{n}$.

— Bivariate Analytic Schema

$\exp(u \log)$

$\Longrightarrow$ Gaussian law

An Erdős-Kac Theorem: *The number of irreducibles is asymptotically Gaussian.*



Read
Buy
Advertise

ANALYTIC COMBINATORICS

by P. Flajolet + R. Sedgewick
(2007).