Special Limit Distributions for Combinatorial Structures

Michèle Soria
LITP-IBP and INRIA

November 8, 1993

[summary by Joris van der Hoeven]

Abstract

The problem of obtaining asymptotic information about parameters of algorithms can often be reduced to the computation of limit distributions of combinatorial structures. Michèle Soria and Philippe Flajolet have shown that for a large number of combinatorial schemes these limit distributions are normal [1]. However, one also frequently encounters discrete limit distributions. In certain degenerate cases it is even possible to obtain continuous special limit distributions. Here, we are interested in these latter two cases and give some examples. More precisely, we shall study special limit distributions arising from a bivariate generating function of the form $F(uC(z))$.

1. Introduction

Consider the generating function $C(z)$ of some combinatorial structure C. Let $F(C)$ be a new combinatorial structure, obtained by applying a combinatorial construction F to C. Then the information about the number of C-structures "in" an $F(C)$-structure is contained in the bivariate generating function $F(uC(z))$. Usually we take one of the following constructors, which are listed with their associated exponential and ordinary generating functions.

Constructor	Labelled(e.g.f.)	Unlabelled(o.g.f.)
Sequence	$\dfrac{1}{1-uC(z)}$	$\dfrac{1}{1-uC(z)}$
Cycle	$\log\dfrac{1}{1-uC(z)}$	$\displaystyle\sum_{k\geq 1}\frac{\phi(k)}{k}\log\frac{1}{1-u^kC(z^k)}$
Set	$\exp(uC(z))$	$\displaystyle\exp\left(\sum_{k\geq 1}\frac{(-1)^{k+1}}{k}u^kC(z^k)\right)$

We will study

$$\Pr(X_n = k) = \frac{[u^k]F(u)[z^n]C^k(z)}{[z^n]F(C(z))},$$

when n tends to infinity. Here X_n is the random variable giving the number of C-structures in an $F(C)$ structure of size n. We will take k of the form $k = \mu_n + x\sigma_n$, where μ_n and σ_n denote respectively the mean value and the standard deviation of X_n. We will also note $r = \rho_C$ and $R = \rho_F$ the convergence radii of C and F.

Several cases should be distinguished, depending on the sign of $C(r) - R$:

- Sub-critical case ($C(r) < R$), leads to a discrete limit distribution;

– Critical case ($C(r) = R$), leads to a continuous special limit distribution, if $R < \infty$;
– Super-critical case ($C(r) > R$), leads to a normal limit distribution.

In the case of continuous limit distributions, we have different types of theorems, depending on what kind of information we wish to obtain. Classically the following types are distinguished:

– Continuity theorem: $X_n \to X \iff \chi_{X_n}(t) \to \chi_X(t)$;
– Integral limit theorem (GLT): $\Pr(x < (X_n - \mu_n)/\sigma_n < y) \to \int_x^y \omega(t)dt$;
– Local limit theorem (LLT): $\sigma_n \Pr(X_n = \lfloor \mu_n + \lambda\sigma_n \rfloor) \to \omega(\lambda)$;
– Exponential tails: $M_{X_n}(t)$ is uniformly bounded in any interval around 0.

Here $\chi_{X_n} = E(e^{itX_n})$ is the characteristic function of X_n and $M_{X_n} = E(e^{tX_n})$ its moment generating function. As generating functions arising from combinatorial problems are often quite regular, one can usually systematically obtain the four types of theorems by using familiar analytical techniques such as the saddle point method and singularity analysis.

2. The sub-critical case

THEOREM 1. *Suppose that $C(z)$ has an algebraic aperiodic singularity*

$$C(z) = \tau - \gamma(1 - z/r)^\lambda + \cdots,$$

with $0 < \lambda < 1$ and $\tau < R$. Then we have a discrete limit distribution,

$$\Pr(X_n = k) \sim \frac{kf_k\tau^{k-1}}{F'(\tau)} \qquad \text{and} \qquad \mu_n \sim 1 + \frac{\tau F''(\tau)}{F'(\tau)}.$$

The proof runs as follows. The condition $\tau < R$ gives

$$F(C(z)) = F(\tau) - F'(\tau)\gamma(1 - z/r)^\lambda + \cdots.$$

The n-th coefficient $[z^n]F(C(z))$ is obtained by singularity analysis. If k is a constant, $[z^n]C^k(z)$ can also be computed by singularity analysis, when $n \to \infty$. Combining these computations, we obtain

$$\Pr(X_n = k) = \frac{[u^k]F(u)[z^n]C^k(z)}{[z^n]F(C(z))} = \frac{f_k k\tau^{k-1}}{F'(\tau)}.$$

Finally, by the usual formula for μ_n, we have

$$\mu_n = \frac{[z^n]C(z)F'(C(z))}{[z^n]F(C(z))} \sim 1 + \frac{\tau F''(\tau)}{F'(\tau)}.$$

Constructor	$\Pr(X_n = k) \sim$	Law	Example
Sequence	$(1-\tau)^2 k\tau^{k-1}$	Geometric δ	General trees
Cycle	$e^{-\tau}\dfrac{\tau^{k-1}}{(k-1)!}$	Poisson δ	Cayley trees
Set	$(1-\tau)\tau^k$	Geometric	
Partition	$e^{1-\tau-e^\tau}\dfrac{B_k\tau^{k-1}}{(k-1)!}$	Bell δ	
Ordered partition	$(1-\tau)^2 e^{-\frac{\tau}{1-\tau}} k f_k \tau^{k-3}$, $f_k = \sum_p \frac{1}{p!}\binom{k-1}{p-1}$	Laguerre δ	

TABLE 1. Application of Theorem 1 for some classical constructors in a labelled environment

λ	α	$\Pr(X_n = xn^\lambda)$	Law	Example
$\frac{1}{2}$	1	$\frac{xe^{-x^2/2}}{\sqrt{n}}$	Raleigh	Random mappings
$\frac{1}{2}$	2	$\sqrt{\frac{2}{\pi n}}x^2e^{-x^2/2}$	Maxwell	Pairs of random mappings
$\frac{1}{4}$	1	$\frac{\Gamma(1/4)}{n^{1/4}}P_{1/4}(x)$	Soria	Extended forests

TABLE 2. Applications of Theorem 2

3. The critical case

THEOREM 2. *Let F be an algebraic-logarithmic function: $F(t) = (1-t)^{-\alpha}\log^\beta[1/(1-t)]$. Suppose that $C(z)$ has an algebraic aperiodic singularity*

$$C(z) = 1 - \gamma(1 - z/r)^\lambda + \cdots, \quad \text{with} \quad 0 < \lambda < 1.$$

Then we have a special continuous limit distribution

$$\Pr(X_n = xn^\lambda) \sim \frac{x^{\alpha-1}\gamma^\alpha}{n^\lambda}\frac{\Gamma(\lambda\alpha)}{\Gamma(\alpha)}P_\lambda(\gamma x), \qquad \text{with} \quad P_\lambda(x) = \sum_{m\geq 0}\frac{(-x)^m}{m!\Gamma(-m\lambda)},$$

where $x = O(1)$. We also have $\mu_n \sim \mu n^\lambda$ and $\sigma_n^2 \sim \sigma^2 n^{2\lambda}$.

PROOF. We have

$$F(C(z)) \sim \frac{\lambda^\beta}{\gamma^\alpha(1 - z/r)^{\lambda\alpha}}\log^\beta\frac{1}{1 - z/r}.$$

The results for μ_n and σ_n are easily obtained by singularity analysis. We must now compute $\Pr(X_n = k)$, for $k = xn^\lambda$, where $x = O(1)$. We have

$$\Pr(X_n = k) = \frac{[u^k]F(u)[z^n]C^k(z)}{[z^n]F(C(z))}$$

and again we use singularity analysis. □

PROPOSITION 1. *$P_\lambda(x)$ is normally convergent for $|x| < x_0$ and hypergeometric, if λ is rational.*

PROOF. If $\lambda = p/q$ is rational, we can write

$$P_{p/q}(x) = \sum_{r=1}^{q-1}P_{p/q}^{(r)}(x), \qquad \text{where} \quad P_{p/q}^{(r)}(x) = \sum_{m>0}\frac{(-x)^{mq+r}}{(mq+r)!\Gamma(-pm - rp/q)}.$$

These latter functions are easily seen to be expressible as finite sums of generalized hypergeometric functions. □

As an example, extended forests have the following bivariate generating function:

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

where $E(z) = 2g(2zg(z))$, with $2g(z) = 1 - \sqrt{1-4z}$. We have $E(z) = 1 - \sqrt[4]{1-4z}$ and

$$P_{1/4}(x) = \frac{x}{4\Gamma(3/4)}{}_0F_2\left(3/4, 1/2\middle|\frac{x^4}{4^4}\right) - \frac{x^2}{4\sqrt{\pi}}{}_0F_2\left(3/4, 5/4\middle|\frac{x^4}{4^4}\right) + \frac{x^3}{8\Gamma(1/4)}{}_0F_2\left(3/2, 5/4\middle|\frac{x^4}{4^4}\right).$$

In her talk, Michèle Soria also considered the super-critical case and the critical case with $R = \infty$. Both cases have normal limit distributions. For more information on the bivariate scheme $F(uC(z))$, see [2].

Bibliography

[1] Flajolet (Philippe) and Soria (Michèle). – Gaussian limiting distributions for the number of components in combinatorial structures. *Journal of Combinatorial Theory, Series A*, vol. 53, 1990, pp. 165–182.

[2] Flajolet (Philippe) and Soria (Michèle). – General combinatorial schemas: Gaussian limit distributions and exponential tails. *Discrete Mathematics*, vol. 114, 1993, pp. 159–180.