Limiting Distributions in Product Schemas

Michèle Soria
LITP-IBP, Université Paris 6

June 6, 1994

[summary by Michèle Soria]

Abstract

We study the limiting distribution of a parameter in product schemas of the type $y(u,x) = g(x)F(uw(x))$, related to classical combinatorial constructions. When $g(x)$ is "negligible", the limiting distribution in $y(u,x)$ is the same as in $F(uw(x))$. On the other hand, when g is of "dominant importance", the limiting distribution is shown to be exclusively discrete, Gaussian or Gamma, according to F and w.

1. Introduction

An important trend in asymptotic combinatorics is to classify limiting distributions appearing in combinatorial schemas according to structural and analytic characteristics of combinatorial constructions. The analysis of functional composition $F(uw(x))$, that translates into generating functions the combinatorial operation of substitution, has been largely investigated [1, 2, 6, 5]. It leads to discrete, or normal, or special distributions, according to analytic properties of F and w.

We shall here consider the case of product schemas $y(x,u) = g(x)F(uw(x))$ studied by Drmota and Soria [3]. Consider for example the two classical results on permutations: the number of cycles of fixed length l is asymptotically Poisson distributed, whereas the limiting distribution of the number of cycles of length $\neq l$ is Gaussian. Let's analyze these results from a "combinatorial schema" standpoint: starting with the construction of permutations as *Sets of Cycles of points* leads to the bivariate schema $y(x,u) = \exp(u \log \frac{1}{1-x})$, hence [5] the Gaussian distribution of the number of cycles (with no restriction) in a random permutation. Marking only cycles of fixed length l gives the product schema $y(x,u) = \exp(\log \frac{1}{1-x} - \frac{x^l}{l} + u\frac{x^l}{l})$, where the factor of dominant importance $\frac{1}{1-x}\exp(-\frac{x^l}{l})$ implies the discrete nature of the distribution, and the *Set* construction in $\exp(u\frac{x^l}{l})$ determines that it is Poisson. On the other hand, the bivariate series for cycles of length $\neq l$ is $y(x,u) = \exp(\frac{x^l}{l}).\exp(u(\log\frac{1}{1-x} - \frac{x^l}{l}))$; here the first factor is negligible, and the limiting distribution is Gaussian as in the unrestricted case.

Given a bivariate series $y(x,u) = \sum y_{nk}x^n u^k$, consider the random variables X_n satisfying $\Pr(X_n = k) = y_{nk}/\sum_k y_{nk}$. We are interested in the asymptotic density of X_n (when $n \to \infty$) in a range around the expected value ($k = \mathrm{E}\,X_n + x\sqrt{\mathrm{Var}\,X_n}$).

For a product schema $y(x,u) = g(x)F(uw(x))$, we thus have to evaluate

$$\frac{y_{nk}}{y_n} = \frac{f_k\,[x^n]g(x)w(x)^k}{[x^n]g(x)F(w(x))}$$

where $f_k = [z^k]F(z)$.

We always assume that the coefficients of the Taylor expansions of $g(x)$, $w(x)$ and $F(w(x))$ can be evaluated, by saddle point method or singularity analysis, and in any case the asymptotic behaviour of the coefficients depends on exact or approximate saddle points.

There are cases where the factor $g(x)$ has no influence on the limiting distribution, i.e. the limiting distribution of $y(x,u)$ is the same as the limiting distribution of $F(uw(x))$. Conversely, $g(x)$ may be of dominant importance, and dictates the limiting distribution of $y(x,u)$ to be either Gaussian or Gamma or discrete. We also investigate some interesting cases where neither $g(x)$ dominates nor is dominated in $y(x,u)$.

The notion of dominance can be formulated in terms of asymptotic behaviour of saddle points. In order to be more precise we introduce the notion of dominance in a product of functions.

DEFINITION 1. Let $f(x)$, $g(x)$ be convergent generating functions with non-negative coefficients. We say that $f(x)$ dominates $g(x)$ if $[x^n]f(x)g(x) \sim g(\zeta_n)[x^n]f(x)$, $(n \to \infty)$, where ζ_n is the saddle point of $f(x)$ defined by $\zeta_n f'(\zeta_n) = nf(\zeta_n)$.

Obviously, if the radius of convergence of the first factor is smaller than that of the second one, the first factor usually dominates. But if both factors have the same radius of convergence, the situation is more involved. However this definition agrees with classical scalings: for example $\exp(\frac{1}{1-x})$ dominates $\frac{1}{(1-x)^\alpha}\log^\beta\frac{1}{1-x}$, and also $\frac{1}{(1-x)^\alpha}$ dominates $\log^\beta\frac{1}{1-x}$, etc.

In a bivariate schema $y(x,u) = g(x)F(uw(x))$, according as the limiting distribution is dictated by the first or second factor, we shall say that g is *dominating* or *dominated* in $y(x,u)$

2. g is dominated in $y(x,u)$

When g is regular at the singular curve of $F(uw(x))$, the limiting distribution is shown to be either Gaussian or discrete. But the situation of simultaneous singularities is more difficult to handle. (In the following we shall refer to *admissible* functions in the sense of Hayman [9], and to *alg-log* functions in the sense of Flajolet and Odlyzko [4].)

THEOREM 1. *Suppose that $F(x)$ is an admissible or alg-log function, with finite radius of convergence R. Let $w(r) = R$ and assume that $w(x)$ and $g(x)$ are regular at $x = r$ (i.e. the radii of convergence of $w(x)$ and $g(x)$ are greater than r). Then $g(x)$ is dominated in $y(x,u) = g(x)F(uw(x))$, and the limiting distribution is Gaussian with mean value $\sim \mu_w(R)^{-1}n$ and variance $\sim \sigma_w^2(R)\mu_w(R)^{-3}n$.*

THEOREM 2. *Suppose that $w(x)$ is an admissible or alg-log function, with finite radius of convergence r such that $w(r) = R$ is finite. Furthermore assume that $F(y)$ is regular at $y = w(r)$ and $g(x)$ is regular at $x = r$. Then $g(x)$ is dominated in $y(x,u) = g(x)F(uw(x))$ and has a discrete limiting distribution, with $Pr[X_n = k] \sim kf_kR^{k-1}/F'(R)$.*

Apart from these two simple cases, the situation where two functions are singular is more complex, and there is probably no general criterion to decide a-priori whether $g(x)$ is dominated or not. Nevertheless we can treat many special cases. For example, if $g(x)$ and $w(x)$ are alg-log functions and $F(w(x))$ has an essential singularity at $x = r$ then $g(x)$ is usually dominated -e.g. if $g(x) = w(x) = \frac{1}{1-x}$ and $F(z) = e^z$, $g(x)$ is dominated in $y(x,u)$)-. Yet, by scaling singularities with the notion of dominance in Definition 1, we can expect the following "general" rule.

RULE 1. Suppose that $g(x)$, $w(x)$, and $F(w(x))$ are admissible or alg-log functions. If $F(w(x))$ dominates $g(x)$, then $g(x)$ is (usually) dominated in $g(x)F(uw(x))$.

3. g is dominating in $y(x,u)$

Alternatively to the previous section, where the factor $g(x)$ has actually no influence on the asymptotic limit distribution of $y(x,u)=g(x)F(uw(x))$, this section is devoted to the case where $g(x)$ is of dominant importance. This means that the saddle point ζ_n of $g(x)$, given by $\zeta_n g'(\zeta_n)= ng(\zeta_n)$, can be used instead of the exact saddle points in the evaluation of the mean, variance and probability. Hence we get, in the range of interest $k \gg\ll \mathrm{E}\,X_n$:

$$\frac{y_{nk}}{y_n} \sim \frac{f_k w(\zeta_n)^k}{F(w(\zeta_n))}.$$

Once again, the case of different radii of convergence is easy to handle.

THEOREM 3. *Suppose that $g(x)$ is admissible or alg-log and has finite radius of convergence r. If $w(x)$ and $F(w(x))$ are regular at $x=r$, then $g(x)$ is dominating in $g(x)F(uw(x))$ and has a discrete limiting distribution with* $\Pr[X_n=k]\sim f_k w(r)^k/F(w(r))$.

In general we can expect a rule of the following kind, which is the converse statement of Rule 1.

RULE 2. Suppose that $g(x)$, $w(x)$, and $F(w(x))$ are admissible or alg-log functions. If $g(x)$ dominates $F(w(x))$ then $g(x)$ is (usually) dominating in $g(x)F(uw(x))$.

One very interesting thing in the dominating case is that there are only few kinds of limiting distributions which can be classified in the following way if $g(x)$ has finite radius of convergence.

THEOREM 4. *Let $g(x)$ be admissible or alg-log, with finite radius of convergence r, and saddle point ζ_n. Suppose that $g(x)$ is dominating in $y(x,u)=g(x)F(uw(x))$, then only four situations can appear:*

a) If $\lim_{x\to r-} w(x)=w(r)$ exists and $F(x)$ is regular at $w(r)$, then X_n has a discrete limiting distribution given by $\Pr[X_n=k]\sim f_k w(r)^k/F(w(r))$.

b) If $\lim_{x\to r-} w(x)=\infty$ and $F(x)$ is entire and admissible, then X_n is asymptotically normally distributed, and $\mathrm{E}\,X_n\sim w(\zeta_n)F'(w(\zeta_n))/F(w(\zeta_n))$.

c) If $\lim_{x\to r-} w(x)=w(r)$ exists and $F(x)$ is admissible and singular at $x=w(r)$, then X_n is asymptotically normally distributed with mean value expressed as in b).

d) If $\lim_{x\to r-} w(x)=w(r)$ exists and if $F(x)$ is an alg-log function (with $\alpha>0$) which is singular at $x=w(r)$, then X_n is asymptotically Gamma distributed with parameter α and $\mathrm{E}\,X_n\sim\alpha/\log\frac{w(r)}{w(\zeta_n)}$.

4. Combinatorial Schemas

Many typical schemas $y(x,u)=g(x)F(uw(x))$ occurring in combinatorial structures related to the "sequence-of" and "cycle-of" constructions, are discussed in [3]. Beside giving illustrations of sections 2 and 3, we also investigate some cases where g is neither dominated nor dominating in $y(x,u)$.

Consider the product schema $y(x,u)=g(x)\exp(uw(x))$, which underlies the examples given in section 1. In the case of cycles of length l, function g has a finite radius of convergence, and w is an entire function. Thus g is dominating, and the limit distribution is Poisson by Theorem 3-a. On the other hand g is dominated for cycles of length $\neq l$, hence the Gaussian limit law.

Now consider the schema $y(x,u)=g(x)\frac{1}{1-uw(x)}$, and let r be the radius of convergence of $F(w(x))=\frac{1}{1-w(x)}$. If g is an entire function, it is dominated in $y(x,u)$, whereas if g is exponential with radius of convergence r, it is dominating. An interesting situation (which actually happens for certain random mapping parameters [8]) arises when g is an alg-log function with radius of convergence r, and $w(r) = 1$: the limit law is shown to be hypergeometric [7, 8].

Bibliography

[1] Bender (Edward A.). – Central and local limit theorems applied to asymptotic enumeration. *Journal of Combinatorial Theory, Series A*, vol. 15, 1973, pp. 91–111.

[2] Canfield (E. Rodney). – Central and local limit theorems for the coefficients of polynomials of binomial type. *Journal of Combinatorial Theory, Series A*, vol. 23, 1977, pp. 275–290.

[3] Drmota (Michael) and Soria (Michèle). – Marking in combinatorial constructions and limit distributions. – Submitted to TCS, November 1993.

[4] Flajolet (Philippe) and Odlyzko (Andrew M.). – Singularity analysis of generating functions. *SIAM Journal on Discrete Mathematics*, vol. 3, n° 2, 1990, pp. 216–240.

[5] 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.

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

[7] Flajolet (Philippe) and Soria (Michèle). – General combinatorial schemas: Special limit distributions. – Manuscript, January 1994.

[8] Gittenberger (Bernhard). – Local limit theorems for distributions of certain random mapping parameters. – Submitted to Random Structures and Algorithms, January 1994.

[9] Hayman (W. K.). – A generalization of Stirling's formula. *Journal für die reine und angewandte Mathematik*, vol. 196, 1956, pp. 67–95.