Combinatoire symbolique (et analytique), séries génératrices bivariées, et distribution de paramètres sur des objets discrets

Mathilde Bouvel (CNRS, Loria, Nancy)

EJCIM à Nantes, juin 2024

Quelques mots d'introduction

Parcours :

- 2006-2010 : Thèse au Liafa (maintenant Irif), Paris ; algorithmique et combinatoire
- 2010-2013 : Chercheuse CNRS au Labri, Bordeaux : combinatoire énumérative
- 2013-2020 : Détachement à l'institut de maths de l'université de Zurich (Suisse) : combinatoire et probabilités
- depuis 2020 : Chercheuse CNRS au Loria, Nancy : combinatoire et probabilités, continued

Sujet de recherche :

- Objets d'étude principaux : Permutations (et autres objets, p.ex. graphes) évitant des sous-structures.
- Questions étudiées : combien y en a-t-il ? sont-elles bien structurées ? à quoi ressemblent-elles lorsque leur taille tend vers l'infini ?

Combinatoire symbolique (et analytique), séries génératrices bivariées, et distribution de paramètres sur des objets discrets

Mathilde Bouvel (CNRS, Loria, Nancy)

EJCIM à Nantes, juin 2024

Loria
Laboratoire lorrain de recherche en informatique et ses applications

Un exemple “jouet”

Notons $\mathcal{W}_n$ l'ensemble des mots de longueur n sur l'alphabet $\{a, b\}$, et $\mathcal{W} = \cup_{n\geq 0}\mathcal{W}_n$.

- Il y a $w_n = 2^n$ mots de longueur n dans $\mathcal{W}$.
- La série génératrice ordinaire de $\mathcal{W}$ est $W(z) = \sum_{n\geq 0} 2^n z^n = \frac{1}{1-2z}$.
- Il y a $w_{n,k} = \binom{n}{k}$ mots de longueur n avec k lettres a dans $\mathcal{W}$.
- La série génératrice bivariée de $\mathcal{W}$, pour le paramètre “nombre de a”, est $W(z, u) = \sum_{n,k\geq 0} \binom{n}{k} z^n u^k = \frac{1}{1-z(1+u)}$.
- La moyenne du nombre de a dans un mot de longueur n est $n/2$.

- La variance du nombre de a dans un mot de longueur n est $n/4$.
- Une fois divisé par n, le nombre de a dans un mot de longueur n suit une distribution asymptotiquement gaussienne.

Objectifs du mini-cours

- Description de familles d'objets discrets par des spécifications combinatoires
- "Dictionnaire" de ces spécifications vers des équations pour les séries génératrices associées
- Énumération (asymptotique) des objets à partir de ces équations (très très brièvement)

Objectifs du mini-cours

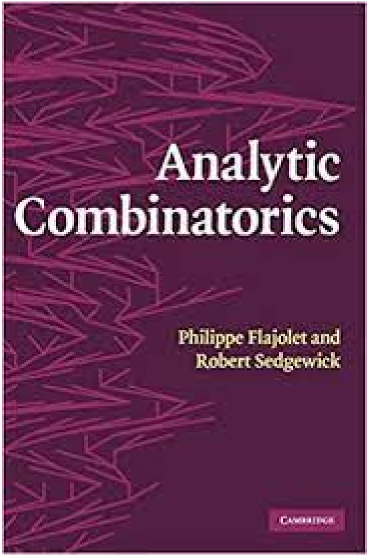
- Description de familles d'objets discrets par des spécifications combinatoires
- "Dictionnaire" de ces spécifications vers des équations pour les séries génératrices associées
- Énumération (asymptotique) des objets à partir de ces équations (très très brièvement)

- Raffinement des spécifications combinatoires pour le suivi de paramètres . . .
- . . . et les séries génératrices bivariées associées
- Expression de l'espérance, la variance, les moments du paramètre à partir de ces séries

Objectifs du mini-cours

- Description de familles d'objets discrets par des spécifications combinatoires
- "Dictionnaire" de ces spécifications vers des équations pour les séries génératrices associées
- Énumération (asymptotique) des objets à partir de ces équations (très très brièvement)
- Raffinement des spécifications combinatoires pour le suivi de paramètres ...
- ... et les séries génératrices bivariées associées
- Expression de l'espérance, la variance, les moments du paramètre à partir de ces séries
- Distribution limite de paramètres lorsque la taille des objets tend vers l'infini (sans entrer dans les détails)

Ouvrage de référence



- plus de 800 pages
- de très nombreux exemples
- des théorèmes généraux ("boîtes noires"), mais aussi les principes qui permettent d'étudier les exemples qui ne rentrent pas dans ces "boîtes noires"

- Couvre l'intégralité de ce mini-cours, et beaucoup beaucoup plus !

Combinatoire symbolique
(et analytique)
pour l'énumération

Classes combinatoires

Une classe combinatoire est un ensemble $\mathcal{C}$ d'objets discrets,
muni d'une notion de taille $|.| : \mathcal{C} \to \mathbb{N}$,
et tel que pour tout n le nombre d'objets de taille n dans $\mathcal{C}$ est fini.

Exemples :

- Mots binaires, taille $=$ nombre de lettres
- Graphes, taille $=$ nombre de sommets
- ~~Graphes, taille = nombre d'arêtes~~
- Graphes connexes, taille $=$ nombre de sommets, ou d'arêtes
- Arbres plans, taille $=$ nombre de sommets, ou de feuilles

Notations :

- $\mathcal{C}_n =$ ensemble des objets de taille n dans $\mathcal{C}$
- $c_n = \#\mathcal{C}_n =$ nombre d'objets de taille n dans $\mathcal{C}$

Énumération, séries génératrices

$(c_n)_{n \geq 0 \text{ ou } 1}$ est la séquence d'énumération de $\mathcal{C}$.

Résoudre le problème d'énumération de $\mathcal{C}$ peut vouloir dire

- trouver une formule close pour c_n (ex: $c_n = \frac{1}{n+1}\binom{2n}{n}$)
- trouver une récurrence pour les c_n (ex: $c_n = \sum_{k=0}^{n-1} c_k c_{n-k-1}$)
- donner un équivalent asymptotique de c_n (ex: $c_n \sim \frac{1}{\sqrt{\pi}} n^{-3/2} 4^n$)

Énumération, séries génératrices

$(c_n)_{n \geq 0 \text{ ou } 1}$ est la séquence d'énumération de $\mathcal{C}$.

Résoudre le problème d'énumération de $\mathcal{C}$ peut vouloir dire

- trouver une formule close pour c_n (ex: $c_n = \frac{1}{n+1}\binom{2n}{n}$)
- trouver une récurrence pour les c_n (ex: $c_n = \sum_{k=0}^{n-1} c_k c_{n-k-1}$)
- donner un équivalent asymptotique de c_n (ex: $c_n \sim \frac{1}{\sqrt{\pi}} n^{-3/2} 4^n$)
- décrire la série génératrice de $\mathcal{C}$, par une forme explicite ou une équation (ex: $C(z) = \frac{1-\sqrt{1-4z}}{2z}$ ou $C(z) = 1 + zC(z)^2$)

La série génératrice (ordinaire) de $\mathcal{C}$ est la série formelle (en la variable z)

$$C(z) = \sum_{\gamma \in \mathcal{C}} z^{|\gamma|} = \sum_{n \geq 0} c_n z^n.$$

On note $[z^n]C(z) = c_n$.

C'est un objet formel, on ne se préoccupe pas de convergence ici (sauf pour l'expression explicite d'une série, qu'on sous-entend autour de 0).

Énumération, séries génératrices

$(c_n)_{n \geq 0 \text{ ou } 1}$ est la séquence d'énumération de $\mathcal{C}$.

Résoudre le problème d'énumération de $\mathcal{C}$ peut vouloir dire

- trouver une formule close pour c_n (ex: $c_n = \frac{1}{n+1}\binom{2n}{n}$)
- trouver une récurrence pour les c_n (ex: $c_n = \sum_{k=0}^{n-1} c_k c_{n-k-1}$)
- donner un équivalent asymptotique de c_n (ex: $c_n \sim \frac{1}{\sqrt{\pi}} n^{-3/2} 4^n$)
- décrire la série génératrice de $\mathcal{C}$, par une forme explicite ou une équation (ex: $C(z) = \frac{1-\sqrt{1-4z}}{2z}$ ou $C(z) = 1 + zC(z)^2$)

La série génératrice (ordinaire) de $\mathcal{C}$ est la série formelle (en la variable z)

$$C(z) = \sum_{\gamma \in \mathcal{C}} z^{|\gamma|} = \sum_{n \geq 0} c_n z^n. \qquad \text{On note } [z^n]C(z) = c_n.$$

C'est un objet formel, on ne se préoccupe pas de convergence ici (sauf pour l'expression explicite d'une série, qu'on sous-entend autour de 0).

Spécifications combinatoires (cadre non-étiqueté)

Idée : se donner un petit nombre de constructeurs simples, qui permettent en les combinant de décrire de nombreuses classes combinatoires.

Une description d'une classe combinatoire à partir de tels constructeurs s'appelle une spécification combinatoire.

Intérêt : les constructeurs correspondent à des opérations simples sur les séries génératrices, à travers un "dictionnaire".
Ainsi, une spécification combinatoire est un bon point de départ pour résoudre le problème d'énumération.

Spécifications combinatoires (cadre non-étiqueté)

Idée : se donner un petit nombre de constructeurs simples, qui permettent en les combinant de décrire de nombreuses classes combinatoires.

Une description d'une classe combinatoire à partir de tels constructeurs s'appelle une spécification combinatoire.

Intérêt : les constructeurs correspondent à des opérations simples sur les séries génératrices, à travers un "dictionnaire".
Ainsi, une spécification combinatoire est un bon point de départ pour résoudre le problème d'énumération.

Ce qui nous attend :

- présentation des constructeurs et cas immédiats du dictionnaire
- quelques exemples de spécifications combinatoires
- le dictionnaire, suite et fin
- retour sur des exemples, pour résoudre leur énumération

Constructeurs (1/3)

- La classe neutre, notée $\mathcal{E}$.
  Elle contient un unique objet, de taille 0. On le note parfois ε.

  Sa série génératrice est $E(z) = 1$.

- La classe neutre, notée $\mathcal{E}$.
  Elle contient un unique objet, de taille 0. On le note parfois ε.

  Sa série génératrice est $E(z) = 1$.

- La classe atomique, notée $\mathcal{Z}$.
  Elle contient un unique objet, de taille 1. On l'appelle parfois atome.

  Sa série génératrice est $Z(z) = z$.

Ces constructeurs sont plutôt les briques de base dans les spécifications.

On se donne deux classes combinatoires $\mathcal{A}$ et $\mathcal{B}$.

- On peut construire l'union disjointe de $\mathcal{A}$ et $\mathcal{B}$, notée $\mathcal{C} = \mathcal{A} \uplus \mathcal{B}$.

  Ce constructeur est utile lorsqu'une classe combinatoire contient deux types d'objets, le premier type correspondant exactement aux objets de $\mathcal{A}$, et le deuxième aux objets de $\mathcal{B}$.

  Taille dans $\mathcal{C}$: héritée des tailles dans $\mathcal{A}$ et $\mathcal{B}$, selon le type de l'objet.

Constructeurs (2/3)

On se donne deux classes combinatoires $\mathcal{A}$ et $\mathcal{B}$.

- On peut construire l'union disjointe de $\mathcal{A}$ et $\mathcal{B}$, notée $\mathcal{C} = \mathcal{A} \uplus \mathcal{B}$.

  Ce constructeur est utile lorsqu'une classe combinatoire contient deux types d'objets, le premier type correspondant exactement aux objets de $\mathcal{A}$, et le deuxième aux objets de $\mathcal{B}$.

  Taille dans $\mathcal{C}$: héritée des tailles dans $\mathcal{A}$ et $\mathcal{B}$, selon le type de l'objet.

- On peut aussi construire le produit cartésien, noté $\mathcal{C} = \mathcal{A} \times \mathcal{B}$.

  Ce constructeur est utile lorsque tout objet de $\mathcal{C}$ se décompose canoniquement en deux parties, la première étant un objet générique de $\mathcal{A}$ et la seconde un objet générique de $\mathcal{B}$.

  Taille de $\gamma = (\alpha, \beta) \in \mathcal{C} = \mathcal{A} \times \mathcal{B}$: somme des tailles de α (comme objet de $\mathcal{A}$) et de β (comme objet de $\mathcal{B}$).

On se donne une classe combinatoire $\mathcal{A}$ **qui ne contient pas d'objet de taille** 0.

- On peut construire une séquence, notée $\mathcal{C} = \text{Seq}(\mathcal{A})$.

  Un objet de $\text{Seq}(\mathcal{A})$ est une suite finie (de longueur arbitraire) d'objets de $\mathcal{A}$. Ses éléments sont donc ordonnés.

  Taille = somme des tailles des objets de $\mathcal{A}$ dans la suite finie.

On se donne une classe combinatoire $\mathcal{A}$ **qui ne contient pas d'objet de taille** 0.

- On peut construire une séquence, notée $\mathcal{C} = \text{Seq}(\mathcal{A})$.

  Un objet de $\text{Seq}(\mathcal{A})$ est une suite finie (de longueur arbitraire) d'objets de $\mathcal{A}$. Ses éléments sont donc ordonnés.

  Taille $=$ somme des tailles des objets de $\mathcal{A}$ dans la suite finie.

  Autrement dit : $\text{Seq}(\mathcal{A}) = \mathcal{E} \uplus \mathcal{A} \uplus (\mathcal{A} \times \mathcal{A}) \uplus (\mathcal{A} \times \mathcal{A} \times \mathcal{A}) \uplus \dots$
  Ou encore : $\text{Seq}(\mathcal{A}) = \{(\alpha_1, \dots \alpha_\ell) : \ell \geq 0, \alpha_i \in \mathcal{A} \ \forall i \in \{1, \dots, \ell\}\}$
  avec la convention que $(\alpha_1, \dots \alpha_\ell) = \varepsilon$ lorsque $\ell = 0$.

On se donne une classe combinatoire $\mathcal{A}$ **qui ne contient pas d'objet de taille** 0.

- On peut construire une séquence, notée $\mathcal{C} = \text{Seq}(\mathcal{A})$.

  Un objet de $\text{Seq}(\mathcal{A})$ est une suite finie (de longueur arbitraire) d'objets de $\mathcal{A}$. Ses éléments sont donc ordonnés.

  Taille $=$ somme des tailles des objets de $\mathcal{A}$ dans la suite finie.

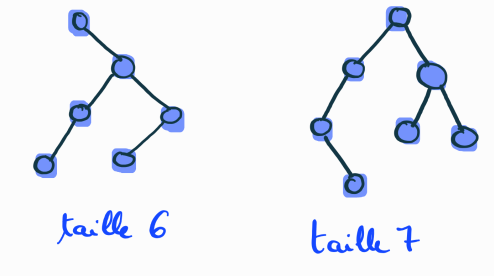
  Autrement dit : $\text{Seq}(\mathcal{A}) = \mathcal{E} \uplus \mathcal{A} \uplus (\mathcal{A} \times \mathcal{A}) \uplus (\mathcal{A} \times \mathcal{A} \times \mathcal{A}) \uplus \dots$

  Ou encore : $\text{Seq}(\mathcal{A}) = \{(\alpha_1, \dots \alpha_\ell) : \ell \geq 0, \alpha_i \in \mathcal{A} \ \forall i \in \{1, \dots, \ell\}\}$ avec la convention que $(\alpha_1, \dots \alpha_\ell) = \varepsilon$ lorsque $\ell = 0$.

- Variantes selon le nombre de composantes dans la suite finie :
  - $\text{Seq}_{\geq k}(\mathcal{A}) = \{(\alpha_1, \dots \alpha_\ell) : \ell \geq k, \alpha_i \in \mathcal{A} \ \forall i \in \{1, \dots, \ell\}\}$
  - En particulier, $\text{Seq}_{\geq 1}(\mathcal{A})$ décrit les séquences non vides d'objets de $\mathcal{A}$
  - $\text{Seq}_{=\ell}(\mathcal{A}) = \underbrace{\mathcal{A} \times \cdots \times \mathcal{A}}_{\ell}$

- Mots binaires (finis) sur l'alphabet $\{a, b\}$, taille $=$ longueur : $\mathcal{W} = \text{Seq}(\mathcal{Z}_a \uplus \mathcal{Z}_b)$ où $\mathcal{Z}_a$ et $\mathcal{Z}_b$ sont deux classes atomiques, correspondant aux lettres a et b respectivement.

- Mots binaires (finis) sur l'alphabet $\{a, b\}$, taille = longueur : $\mathcal{W} = \text{Seq}(\mathcal{Z}_a \uplus \mathcal{Z}_b)$ où $\mathcal{Z}_a$ et $\mathcal{Z}_b$ sont deux classes atomiques, correspondant aux lettres a et b respectivement.
- Arbres binaires (enracinés, plans), taille = nombre total de nœuds : $\mathcal{B} = \mathcal{E} \uplus \mathcal{Z} \times \mathcal{B} \times \mathcal{B}$, en distinguant les cas d'un arbre vide et d'un arbre non-vide (que l'on décompose autour de sa racine)



- Mots binaires (finis) sur l'alphabet $\{a, b\}$, taille $=$ longueur : $\mathcal{W} = \text{Seq}(\mathcal{Z}_a \uplus \mathcal{Z}_b)$ où $\mathcal{Z}_a$ et $\mathcal{Z}_b$ sont deux classes atomiques, correspondant aux lettres a et b respectivement.
- Arbres binaires (enracinés, plans), taille $=$ nombre total de nœuds : $\mathcal{B} = \mathcal{E} \uplus \mathcal{Z} \times \mathcal{B} \times \mathcal{B}$, en distinguant les cas d'un arbre vide et d'un arbre non-vide (que l'on décompose autour de sa racine)

Remarques :

- Il s'agit d'une spécification récursive, ce qui est typique pour les familles d'arbres (entre autres).
- On peut aussi considérer des variantes des arbres binaires où taille $=$ nombre de feuilles, ou nombre de nœuds internes, ce qui donne d'autres spécifications.

Une composition d'un entier n est une suite $(n_1, n_2, \ldots, n_k)$ avec chaque $n_i \geq 1$ telle que $n_1 + n_2 + \cdots + n_k = n$.

Exemples : $(1, 3, 1)$ est une composition de 5, souvent notée 1+3+1.
Il y a 4 compositions de 3, qui sont 3, 2+1, 1+2 et 1+1+1.

Une composition d'un entier n est une suite $(n_1, n_2, \ldots, n_k)$ avec chaque $n_i \geq 1$ telle que $n_1 + n_2 + \cdots + n_k = n$.

Exemples : $(1, 3, 1)$ est une composition de 5, souvent notée 1+3+1. Il y a 4 compositions de 3, qui sont 3, 2+1, 1+2 et 1+1+1.

L'ensemble $\mathcal{D}$ des compositions d'entiers est une classe combinatoire. Une spécification en est donnée par le système de deux équations

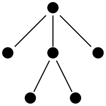
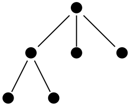
$$\begin{cases} \mathcal{I} = \mathrm{Seq}_{\geq 1}(\mathcal{Z}) \\ \mathcal{D} = \mathrm{Seq}(\mathcal{I}). \end{cases}$$

$\mathcal{I}$ représente la classe des entiers naturels non-nuls (où taille $=$ valeur).

Exemples (3/3)

Un arbre de Catalan est

- un arbre enraciné
- où les nœuds internes ont un nombre d'enfants (=arité) arbitraire,
- où les enfants de tout nœud interne sont ordonnés de gauche à droite.

Exemple :  $\neq$  .

Taille = nombre total de nœuds.

La classe combinatoire $\mathcal{A}$ des arbres de Catalan (dont on impose que la taille soit au moins 1) est décrite par la spécification

$$\mathcal{A} = \mathcal{Z} \times \mathrm{Seq}(\mathcal{A}).$$

Le dictionnaire. . .

. . . ou comment traduire les spécifications en équations sur les séries génératrices d'énumération.

Rappels :

- La série génératrice de $\mathcal{C}$ est $C(z) = \sum_{n\geq 0} c_n z^n = \sum_{\gamma\in\mathcal{C}} z^{|\gamma|}$.
- La série génératrice de la classe neutre $\mathcal{E}$ est $E(z) = 1$.
- La série génératrice de la classe atomique $\mathcal{Z}$ est $Z(z) = z$.

Le dictionnaire…

…ou comment traduire les spécifications en équations sur les séries génératrices d'énumération.

Rappels :

- La série génératrice de $\mathcal{C}$ est $C(z) = \sum_{n \geq 0} c_n z^n = \sum_{\gamma \in \mathcal{C}} z^{|\gamma|}$.
- La série génératrice de la classe neutre $\mathcal{E}$ est $E(z) = 1$.
- La série génératrice de la classe atomique $\mathcal{Z}$ est $Z(z) = z$.

Dictionnaire pour les constructeurs $\uplus$, $\times$ et Seq :

$\mathcal{C} = \mathcal{A} \uplus \mathcal{B}$	$C(z) = A(z) + B(z)$
$\mathcal{C} = \mathcal{A} \times \mathcal{B}$	$C(z) = A(z) \cdot B(z)$
$\mathcal{C} = \mathrm{Seq}(\mathcal{A})$	$C(z) = \frac{1}{1-A(z)}$

en supposant pour la séquence que $\mathcal{A}$ ne contient pas d'objet de taille 0, c'est-à-dire que $[z^0]A(z) = 0$.

Preuves du dictionnaire

- Cas de l'union disjointe :

$$C(z) = \sum_{\gamma \in \mathcal{A} \uplus \mathcal{B}} z^{|\gamma|} = \sum_{\gamma \in \mathcal{A}} z^{|\gamma|} + \sum_{\gamma \in \mathcal{B}} z^{|\gamma|} = A(z) + B(z)$$

Preuves du dictionnaire

- Cas de l'union disjointe :

$$C(z) = \sum_{\gamma \in \mathcal{A} \uplus \mathcal{B}} z^{|\gamma|} = \sum_{\gamma \in \mathcal{A}} z^{|\gamma|} + \sum_{\gamma \in \mathcal{B}} z^{|\gamma|} = A(z) + B(z)$$

- Cas du produit cartésien:

$$C(z) = \sum_{\gamma \in \mathcal{A} \times \mathcal{B}} z^{|\gamma|} = \sum_{(\alpha,\beta) \in \mathcal{A} \times \mathcal{B}} z^{|\alpha|+|\beta|} = \sum_{\alpha \in \mathcal{A}} z^{|\alpha|} \cdot \sum_{\beta \in \mathcal{B}} z^{|\beta|} = A(z) \cdot B(z).$$

Preuves du dictionnaire

- Cas de l'union disjointe :

$$C(z) = \sum_{\gamma \in \mathcal{A} \uplus \mathcal{B}} z^{|\gamma|} = \sum_{\gamma \in \mathcal{A}} z^{|\gamma|} + \sum_{\gamma \in \mathcal{B}} z^{|\gamma|} = A(z) + B(z)$$

- Cas du produit cartésien:

$$C(z) = \sum_{\gamma \in \mathcal{A} \times \mathcal{B}} z^{|\gamma|} = \sum_{(\alpha,\beta) \in \mathcal{A} \times \mathcal{B}} z^{|\alpha|+|\beta|} = \sum_{\alpha \in \mathcal{A}} z^{|\alpha|} \cdot \sum_{\beta \in \mathcal{B}} z^{|\beta|} = A(z) \cdot B(z).$$

- Case de la séquence : de $\mathrm{Seq}(\mathcal{A}) = \uplus_{k \geq 0} \underbrace{\mathcal{A} \times \cdots \times \mathcal{A}}_{k}$, on déduit

$$C(z) = \sum_{k \geq 0} A(z)^k = \frac{1}{1 - A(z)}$$

Preuves du dictionnaire

- Cas de l'union disjointe :

$$C(z) = \sum_{\gamma \in \mathcal{A} \uplus \mathcal{B}} z^{|\gamma|} = \sum_{\gamma \in \mathcal{A}} z^{|\gamma|} + \sum_{\gamma \in \mathcal{B}} z^{|\gamma|} = A(z) + B(z)$$

- Cas du produit cartésien:

$$C(z) = \sum_{\gamma \in \mathcal{A} \times \mathcal{B}} z^{|\gamma|} = \sum_{(\alpha,\beta) \in \mathcal{A} \times \mathcal{B}} z^{|\alpha|+|\beta|} = \sum_{\alpha \in \mathcal{A}} z^{|\alpha|} \cdot \sum_{\beta \in \mathcal{B}} z^{|\beta|} = A(z) \cdot B(z).$$

- Case de la séquence : de $\mathrm{Seq}(\mathcal{A}) = \uplus_{k \geq 0} \underbrace{\mathcal{A} \times \cdots \times \mathcal{A}}_{k}$, on déduit

$$C(z) = \sum_{k \geq 0} A(z)^k = \frac{1}{1 - A(z)}$$

- Variante pour la séquence non-vide $\mathcal{C} = \mathrm{Seq}_{\geq 1}(\mathcal{A})$: $C(z) = \frac{A(z)}{1 - A(z)}$

- Mots binaires, $\mathcal{W} = \text{Seq}(\mathcal{Z}_a \uplus \mathcal{Z}_b)$

$\Rightarrow$ $W(z) = \frac{1}{1-(z+z)} = \frac{1}{1-2z}$

Remarque : On retrouve bien que le nombre de mots binaires de taille n est 2^n. En effet, $\frac{1}{1-2z} = \sum_{n\geq 0} 2^n z^n$

- Mots binaires, $\mathcal{W} = \mathrm{Seq}(\mathcal{Z}_a \uplus \mathcal{Z}_b)$

$\Rightarrow$ $W(z) = \frac{1}{1-(z+z)} = \frac{1}{1-2z}$

Remarque : On retrouve bien que le nombre de mots binaires de taille n est 2^n. En effet, $\frac{1}{1-2z} = \sum_{n\geq 0} 2^n z^n$

- Compositions, $\mathcal{D} = \mathrm{Seq}(\mathcal{I})$ avec $\mathcal{I} = \mathrm{Seq}_{\geq 1}(\mathcal{Z})$

$\Rightarrow$ $I(z) = \frac{z}{1-z}$ et $D(z) = \frac{1}{1-I(z)} = \frac{1-z}{1-2z}$

Conséquence : Il y a donc 2^{n-1} compositions de tout entier $n \geq 1$, car $\frac{1-z}{1-2z} = \frac{1}{1-2z} - \frac{z}{1-2z} = \sum_{n\geq 0} 2^n z^n - \sum_{n\geq 1} 2^{n-1} z^n = 1 + \sum_{n\geq 1} 2^{n-1} z^n$

- Mots binaires, $\mathcal{W} = \text{Seq}(\mathcal{Z}_a \uplus \mathcal{Z}_b)$

$\Rightarrow \; W(z) = \frac{1}{1-(z+z)} = \frac{1}{1-2z}$

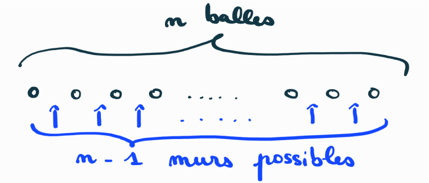
Remarque : On retrouve bien que le nombre de mots binaires de taille n est 2^n. En effet, $\frac{1}{1-2z} = \sum_{n\geq 0} 2^n z^n$

- Compositions, $\mathcal{D} = \text{Seq}(\mathcal{I})$ avec $\mathcal{I} = \text{Seq}_{\geq 1}(\mathcal{Z})$

$\Rightarrow \; I(z) = \frac{z}{1-z}$ et $D(z) = \frac{1}{1-I(z)} = \frac{1-z}{1-2z}$

Conséquence : Il y a donc 2^{n-1} compositions de tout entier $n \geq 1$, car $\frac{1-z}{1-2z} = \frac{1}{1-2z} - \frac{z}{1-2z} = \sum_{n\geq 0} 2^n z^n - \sum_{n\geq 1} 2^{n-1} z^n = 1 + \sum_{n\geq 1} 2^{n-1} z^n$

Preuve bijective :



- Arbres binaires, $\mathcal{B} = \mathcal{E} \uplus \mathcal{Z} \times \mathcal{B} \times \mathcal{B}$

$\Rightarrow$ $B(z) = 1 + zB(z)^2$

- Arbres binaires, $\mathcal{B} = \mathcal{E} \uplus \mathcal{Z} \times \mathcal{B} \times \mathcal{B}$

$\Rightarrow$ $B(z) = 1 + zB(z)^2$

Résolution :
La série $B(z)$ est donc l'inconnue dans une équation de degré 2.
A priori, cette équation a deux solutions : $B_{\pm}(z) = \frac{1 \pm \sqrt{1-4z}}{2z}$.
Pourtant, $B(z) = 1 + zB(z)^2$ définie une unique série formelle $B(z)$.

- Arbres binaires, $\mathcal{B} = \mathcal{E} \uplus \mathcal{Z} \times \mathcal{B} \times \mathcal{B}$

$\Rightarrow$ $B(z) = 1 + zB(z)^2$

Résolution :
La série $B(z)$ est donc l'inconnue dans une équation de degré 2.
A priori, cette équation a deux solutions : $B_{\pm}(z) = \frac{1 \pm \sqrt{1-4z}}{2z}$.
Pourtant, $B(z) = 1 + zB(z)^2$ définie une unique série formelle $B(z)$.

Trouver "la bonne" solution : développer $B_{\pm}(z)$ en série :

- $B_{+}(z) = 1/z - 1 - z - 2z^2 - 5z^3 - 14z^4 + O(z^5)$
- $B_{-}(z) = 1 + z + 2z^2 + 5z^3 + 14z^4 + 42z^5 + O(z^6)$

Donc $B(z) = B_{-}(z) = \frac{1-\sqrt{1-4z}}{2z}$.

- Arbres binaires, $\mathcal{B} = \mathcal{E} \uplus \mathcal{Z} \times \mathcal{B} \times \mathcal{B}$

$\Rightarrow$ $B(z) = 1 + zB(z)^2$

Résolution :
La série $B(z)$ est donc l'inconnue dans une équation de degré 2.
A priori, cette équation a deux solutions : $B_{\pm}(z) = \frac{1 \pm \sqrt{1-4z}}{2z}$.
Pourtant, $B(z) = 1 + zB(z)^2$ définie une unique série formelle $B(z)$.

Trouver "la bonne" solution : développer $B_{\pm}(z)$ en série :

- $B_{+}(z) = 1/z - 1 - z - 2z^2 - 5z^3 - 14z^4 + O(z^5)$
- $B_{-}(z) = 1 + z + 2z^2 + 5z^3 + 14z^4 + 42z^5 + O(z^6)$

Donc $B(z) = B_{-}(z) = \frac{1-\sqrt{1-4z}}{2z}$.
Le développement en série permet d'écrire $B(z) = \sum_{n \geq 0} \frac{1}{n+1} \binom{2n}{n} z^n$,
où apparaissent les fameux nombres de Catalan $Cat_n = \frac{1}{n+1} \binom{2n}{n}$.

Retour sur les exemples (3/3)

- Arbres de Catalan, $\mathcal{A} = \mathcal{Z} \times \mathrm{Seq}(\mathcal{A})$

$\Rightarrow$ $A(z) = \frac{z}{1-A(z)}$, ou encore $A(z)^2 - A(z) + z = 0$

- Arbres de Catalan, $\mathcal{A} = \mathcal{Z} \times \mathrm{Seq}(\mathcal{A})$

$\Rightarrow$ $A(z) = \frac{z}{1-A(z)}$, ou encore $A(z)^2 - A(z) + z = 0$

Résolution : Avec la même méthode que pour les arbres binaires, on trouve $A(z) = \frac{1-\sqrt{1-4z}}{2}$.

- Arbres de Catalan, $\mathcal{A} = \mathcal{Z} \times \text{Seq}(\mathcal{A})$

$\Rightarrow$ $A(z) = \frac{z}{1-A(z)}$, ou encore $A(z)^2 - A(z) + z = 0$

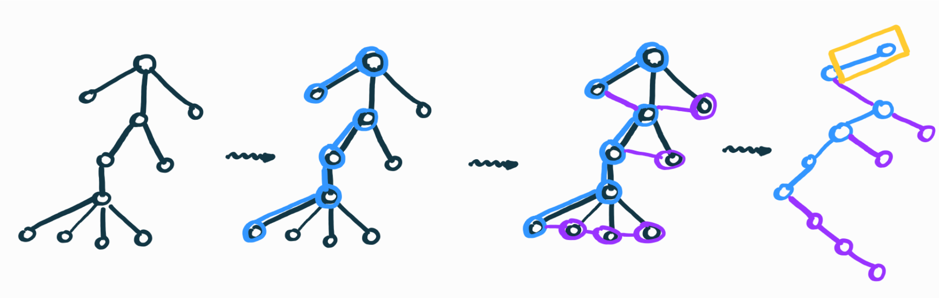
Résolution : Avec la même méthode que pour les arbres binaires, on trouve $A(z) = \frac{1-\sqrt{1-4z}}{2}$.

Coïncidence : Et donc $z \cdot B(z) = A(z)$???

- Cela signifie qu'il y a autant d'arbres binaires à n nœuds que d'arbres de Catalan à $n+1$ nœuds (pour tout $n \geq 0$).
- Ici : c'est une conséquence des expressions obtenues pour les séries génératrices
- Autre possibilité : bijection parfois appelée "fils aîné – frère droit"

La bijection “fils aîné – frère droit”

Il y a autant arbres de Catalan à $n+1$ nœuds que d'arbres binaires à n nœuds



Cadre non-étiqueté et cadre étiqueté

Jusqu'ici, les objets considérés sont composés d'atomes indistinguables a priori : rien ne différencie deux lettres a dans un mot binaire (ou deux nœuds dans un arbre), une fois isolés du reste de l'objet.

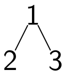
On parle de cadre non-étiqueté.

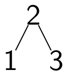
Jusqu'ici, les objets considérés sont composés d'atomes indistinguables a priori : rien ne différencie deux lettres a dans un mot binaire (ou deux nœuds dans un arbre), une fois isolés du reste de l'objet.

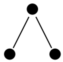
On parle de cadre non-étiqueté.

On peut "décorer" (= étiqueter) les atomes d'un objet de taille n par les entiers de 1 à n. Ceci indique un ordre total sur les atomes qui composent l'objet, qui sont donc distingués les uns des autres par leurs étiquettes.

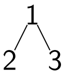
On parle de cadre étiqueté.

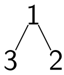
Exemple : Les arbres enracinés 

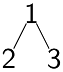
et 

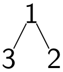
sont des objets étiquetés différents, alors qu'ils correspondent au même arbre non-étiqueté .

Exemple, suite :

- Si on considère les arbres étiquetés plongés dans le plan (par exemple, 

$\neq$ 

), il y a $n!$ étiquetages possibles d'un arbre de Catalan de taille n, qui correspondent tous à des objets différents de la classe d'arbres étiquetés. Cela donne $n!a_n$ arbres étiquetés de taille n (avec $a_n = Cat_{n-1}$).
- Dans le cas des arbres étiquetés mais non-plongés dans le plan (par exemple, 

$=$ 

), il y a moins de $n!a_n$ arbres de taille n. Mais $n!$ reste le "bon ordre de grandeur".

Exemple, suite :

- Si on considère les arbres étiquetés plongés dans le plan (par exemple, $\begin{array}{c}1\\ 2\quad 3\end{array} \neq \begin{array}{c}1\\ 3\quad 2\end{array}$), il y a $n!$ étiquetages possibles d'un arbre de Catalan de taille n, qui correspondent tous à des objets différents de la classe d'arbres étiquetés. Cela donne $n!a_n$ arbres étiquetés de taille n (avec $a_n = Cat_{n-1}$).
- Dans le cas des arbres étiquetés mais non-plongés dans le plan (par exemple, $\begin{array}{c}1\\ 2\quad 3\end{array} = \begin{array}{c}1\\ 3\quad 2\end{array}$), il y a moins de $n!a_n$ arbres de taille n. Mais $n!$ reste le "bon ordre de grandeur".

À une classe combinatoire étiquetée $\mathcal{C}$, on associe sa série génératrice exponentielle (c_n étant toujours le nombre d'objets de taille n dans $\mathcal{C}$) :

$$C(z) = \sum_{n\geq 0} \frac{c_n}{n!} z^n = \sum_{\gamma\in\mathcal{C}} \frac{1}{|\gamma|!} z^{|\gamma|}.$$

Classe étiquetée des permutations

C'est l'exemple le plus simple de classe combinatoire étiquetée.

- Une permutation de taille n est une bijection de $\{1, 2, \ldots, n\}$ dans lui-même.
- On peut voir une permutation de taille n comme une suite de n nombres, où chaque nombre entre 1 et n apparaît exactement une fois.
- Exemple : $\sigma = 3\ 1\ 5\ 2\ 6\ 4$ est telle que $\sigma(1) = 3, \ldots, \sigma(6) = 4$.
- Il y a $n!$ permutations de taille n.

On note $\mathcal{P}$ la classe combinatoire des permutations.

La série génératrice exponentielle de $\mathcal{P}$ est

$$P(z) = \sum_{n \geq 0} \frac{n!}{n!} z^n = \sum_{n \geq 0} z^n = \frac{1}{1 - z}.$$

Spécifications, constructeurs et dictionnaire en étiqueté

- Comme dans le cadre non-étiqueté, on peut décrire des classes combinatoires étiquetées au moyen de spécifications combinatoires utilisant des constructeurs.
- Il y a un autre dictionnaire de traduction des constructeurs étiquetés vers des opérateurs sur les séries génératrices exponentielles.

Spécifications, constructeurs et dictionnaire en étiqueté

- Comme dans le cadre non-étiqueté, on peut décrire des classes combinatoires étiquetées au moyen de spécifications combinatoires utilisant des constructeurs.
- Il y a un autre dictionnaire de traduction des constructeurs étiquetés vers des opérateurs sur les séries génératrices exponentielles.

Les “briques de base” restent la classe neutre $\mathcal{E}$ (contenant un unique objet, de taille 0), et la classe atomique $\mathcal{Z}$ (contenant un unique objet, de taille 1, dont l'unique atome est étiqueté par 1).

Comme $0! = 1$ et $1! = 1$, leurs séries génératrices exponentielles sont les mêmes que leurs séries génératrices ordinaires : $E(z) = 1$ et $Z(z) = z$.

Union disjointe et produit cartésien

Soient $\mathcal{A}$ et $\mathcal{B}$ deux classes combinatoires, dont les séries génératrices exponentielles sont $A(z)$ et $B(z)$.

- L'union disjointe $\mathcal{C} = \mathcal{A} \uplus \mathcal{B}$ est bien une classe combinatoire étiquetée, de série génératrice exponentielle $C(z) = A(z) + B(z)$, car

$$C(z) = \sum_{\gamma \in \mathcal{A} \uplus \mathcal{B}} \frac{1}{|\gamma|!} z^{|\gamma|} = \sum_{\gamma \in \mathcal{A}} \frac{1}{|\gamma|!} z^{|\gamma|} + \sum_{\gamma \in \mathcal{B}} \frac{1}{|\gamma|!} z^{|\gamma|} = A(z) + B(z).$$

- Le produit cartésien $\mathcal{A} \times \mathcal{B}$ n'est pas une classe combinatoire étiquetée. Le problème vient des "doublons d'étiquettes". Autrement dit, il faut construire un ordre total sur les atomes de $(\alpha, \beta) \in \mathcal{A} \times \mathcal{B}$ à partir des ordres totaux sur les atomes de α et de β.

Produit étiqueté : définition

Le produit étiqueté $\mathcal{A} \star \mathcal{B}$ est défini ainsi. Un objet de $\mathcal{C} = \mathcal{A} \star \mathcal{B}$ de taille n est donné par :

- un objet α de $\mathcal{A}$ de taille $k \leq n$
- un objet β de $\mathcal{B}$ de taille $n - k$
- et un réétiquetage de $\{1, \ldots, k\}$ en un sous-ensemble de k éléments dans $\{1, \ldots, n\}$, disons $\{i_1 < i_2 < \cdots < i_k\}$

Les k atomes de α sont alors réétiquetés (en préservant l'ordre naturel) par $\{i_1, i_2, \ldots, i_k\}$, et les $(n-k)$ atomes de β par $\{1, \ldots, n\} \setminus \{i_1, i_2, \ldots, i_k\}$.

$\mathcal{C} = \mathcal{A} \star \mathcal{B}$ est bien une classe combinatoire étiquetée, et sa série génératrice exponentielle est $C(z) = A(z) \cdot B(z)$.

Cette formule indique que le produit étiqueté est le bon analogue du produit cartésien dans le cadre étiqueté.

Produit étiqueté : preuve

Théorème : La série génératrice de $\mathcal{C} = \mathcal{A} \star \mathcal{B}$ est $C(z) = A(z) \cdot B(z)$.

Notons a_n (resp. b_n, c_n) le nombre d'objets de taille n dans $\mathcal{A}$ (resp. $\mathcal{B}$, $\mathcal{C}$).

Alors $c_n = \sum_{k=0}^{n} a_k \cdot b_{n-k} \cdot \binom{n}{k}$, le coefficient binomial tenant compte du réétiquetage.

Ainsi,

$$\begin{aligned} C(z) &= \sum_{n\geq 0} \frac{c_n}{n!} z^n = \sum_{n\geq 0} \sum_{k=0}^{n} a_k \cdot b_{n-k} \cdot \frac{n!}{k!(n-k)!} \frac{1}{n!} z^n \\ &= \sum_{n\geq 0} \sum_{k=0}^{n} \frac{a_k}{k!} z^k \cdot \frac{b_{n-k}}{(n-k)!} z^{n-k} = A(z) \cdot B(z). \qquad \square \end{aligned}$$

Constructeur de séquence

On le définit à partir de $\uplus$ et $\star$, de manière similaire au cadre non-étiqueté :

$$\begin{aligned}\mathrm{Seq}(\mathcal{A}) &= \mathcal{E} \;\uplus\; \mathcal{A} \;\uplus\; \mathcal{A}\star\mathcal{A} \;\uplus\; \mathcal{A}\star\mathcal{A}\star\mathcal{A} \;\uplus \dots \\ &= \biguplus_{k\geq 0} \underbrace{\mathcal{A}\star\cdots\star\mathcal{A}}_{k}.\end{aligned}$$

Comme avant, il est nécessaire que $\mathcal{A}$ ne contienne pas d'objet de taille 0 !

Le dictionnaire dans le cadre étiqueté donne que la série génératrice exponentielle pour $\mathcal{C} = \mathrm{Seq}(\mathcal{A})$ est $C(z) = \frac{1}{1-A(z)}$.

Mêmes variantes $\mathrm{Seq}_{\geq k}$ et $\mathrm{Seq}_{=\ell}$ quand on restreint le nombre de composantes.

Constructeur d'ensemble

On se donne une classe combinatoire étiquetée $\mathcal{A}$, sans objet de taille 0, et $A(z)$ sa série génératrice exponentielle.

- Pour tout $k \geq 0$, on définit $\mathrm{Set}_{=k}(\mathcal{A})$ la classe étiquetée dont les éléments sont les ensembles de k objets de $\mathcal{A}$.
- C'est-à-dire que $\mathrm{Set}_{=k}(\mathcal{A})$ est une variante de $\mathrm{Seq}_{=k}(\mathcal{A})$ où l'on oublie l'ordre entre les k objets de $\mathcal{A}$ qui composent un objet.
- La série génératrice exponentielle est donc $\mathrm{Set}_{=k}(\mathcal{A})$ est $\frac{1}{k!}A(z)^k$.
- On définit ensuite $\mathrm{Set}(\mathcal{A}) = \biguplus_{k \geq 0} \mathrm{Set}_{=k}(\mathcal{A})$.
- Sa série génératrice exponentielle de $\mathcal{C} = \mathrm{Set}(\mathcal{A})$ est $C(z) = \sum_{k \geq 0} \frac{1}{k!}A(z)^k = \exp(A(z))$.

On peut considérer la variante $\mathrm{Set}_{\geq \ell}$ et ajuster la série génératrice en conséquence (soustraction des premiers termes).

Constructeur de cycle

Partant toujours de $\mathcal{A}$ sans objet de taille 0, et $A(z)$ sa série génératrice exponentielle.

- Pour tout $k \geq 1$, $\mathrm{Cyc}_{=k}(\mathcal{A})$ est la variante de $\mathrm{Seq}_{=k}(\mathcal{A})$ où les objets sont considérés à permutation cyclique près de leurs composantes.
- Cela signifie que $(\alpha_1, \alpha_2, \alpha_3)$, $(\alpha_2, \alpha_3, \alpha_1)$ et $(\alpha_3, \alpha_1, \alpha_2)$ représentent toutes le même objet de $\mathrm{Cyc}_{=3}(\mathcal{A})$.
- La série génératrice exponentielle de $\mathrm{Cyc}_{=k}(\mathcal{A})$ est $\frac{1}{k}A(z)^k$.
- On définit ensuite $\mathrm{Cyc}(\mathcal{A}) = \biguplus_{k\geq 1} \mathrm{Cyc}_{=k}(\mathcal{A})$.
- La série génératrice exponentielle de $\mathcal{C} = \mathrm{Cyc}(\mathcal{A})$ est $C(z) = \sum_{k\geq 1} \frac{1}{k}A(z)^k = \log\left(\frac{1}{1-A(z)}\right)$.
- Remarque : on ne considère par les cycles contenant 0 éléments.

Et toujours la variante $\mathrm{Cyc}_{\geq \ell}$.

Rappel : la série génératrice exponentielle des permutations est $P(z) = \frac{1}{1-z}$. On peut le retrouver de deux manières.

- Point de vue “mots” :
  permutation de taille n = suite ordonnée des symboles $\{1, 2, \ldots, n\}$
- Spécification correspondante : $\mathcal{P} = \mathrm{Seq}(\mathcal{Z})$

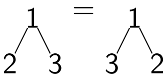
$\Rightarrow$ Série génératrice : $\frac{1}{1-z} = P(z)$ par le dictionnaire étiqueté

- Point de vue algébrique :
  permutation de taille n = ensemble de cycles à supports disjoints (dont on fait le produit dans le groupe symétrique)
- Spécification correspondante : $\tilde{\mathcal{P}} = \mathrm{Set}(\mathrm{Cyc}(\mathcal{Z}))$

$\Rightarrow$ Série génératrice : $\exp(\log(\frac{1}{1-z})) = \frac{1}{1-z} = P(z)$ par le dictionnaire étiqueté

Arbres de Cayley = arbres enracinés, nœuds étiquetés par les entiers de 1 à n, et pas considérés plongés dans le plan.

Exemple :



Spécification pour les arbres de Cayley à au moins un nœud :

$$\mathcal{T} = \mathcal{Z} \star \mathsf{Set}(\mathcal{T})$$

La série génératrice exponentielle associée satisfait

$$T(z) = z \cdot \exp(T(z))$$

Interlude

Jusqu'ici, on a vu :

- les classes combinatoires non-étiquetées, et leurs séries génératrices ordinaires
- les spécifications combinatoires et le dictionnaire (des constructeurs vers des opérateurs sur les séries génératrices)
- la même chose en étiqueté (avec séries exponentielles)

Ce qu'on fera demain : Raffinement pour l'étude de paramètres sur ces objets.

Mais avant : Un peu de combinatoire analytique ; ou comment utiliser les équations obtenues sur les séries génératrices pour faire de l'énumération.

Des spécifications à l'énumération sur nos exemples

En non-étiqueté :

- Mots binaires, $W(z) = \frac{1}{1-(z+z)} = \frac{1}{1-2z}$ $\Rightarrow w_n = 2^n$
- Compositions, $D(z) = \frac{1-z}{1-2z}$ $\Rightarrow d_n = 2^{n-1}$ pour $n \geq 1$
- Arbres binaires, $B(z) = 1 + zB(z)^2$ $\Rightarrow b_n = Cat_n = \frac{1}{n+1}\binom{2n}{n}$
  en passant par $B(z) = B_-(z) = \frac{1-\sqrt{1-4z}}{2z}$ et le développement en série de la racine

En étiqueté :

- Permutations, $P(z) = \frac{1}{1-z}$ $\Rightarrow p_n = n!$
- Arbres de Cayley, $T(z) = z \cdot \exp(T(z))$ $\Rightarrow n^{n-1}$
  peut être obtenu par inversion de Lagrange

Dans ces cas faciles, on a toujours une résolution explicite du problème énumération.

Approche de la combinatoire analytique

Objectif : Même si on ne peut pas résoudre explicitement l'énumération ($c_n = \dots$), trouver un équivalent asymptotique de c_n lorque $n \to \infty$

- On considère $C(z)$ comme une fonction de la variable complexe z.
- Par construction, cette fonction possède un développement en série entière au voisinage de $z = 0$: $C(z) = \sum_{n \geq 0} \tilde{c}_n z^n$.
- Dans les bons cas (souvent !), ce développement s'étend dans un voisinage macroscopique de $z = 0$, et on peut étudier $C(z)$ lorsqu'on approche de la frontière de ce domaine d'analycité.
- On peut ainsi relier le comportement asymptotique des coefficients $[z^n]C(z)$ lorsque $n \to \infty$ au comportement de la série $C(z)$ lorsque z est au voisinage de sa singularité dominante ($=$ le z de plus petit module tel que $C(z)$ n'est plus analytique en z, supposé unique ici).

Les deux grands principes de la combinatoire analytique

1. La localisation de la singularité dominante de $C(z)$ détermine l'ordre de croissance exponentielle de $[z^n]C(z)$.
   Plus précisément, si ρ est la singularité dominante de $C(z)$, alors $[z^n]C(z) \sim \theta(n) \cdot \rho^{-n}$, où $\theta(n)$ est sous-exponentielle en n.
2. La nature de la singularité dominante de $C(z)$ détermine le facteur sous-exponentiel $\theta(n)$.

Exemple des arbres binaires :

- $B(z) = \frac{1-\sqrt{1-4z}}{2z}$ a sa singularité dominante en $1/4$, et elle est de "type racine".
- Explique que $[z^n]B(z) \sim \frac{1}{\sqrt{\pi}} n^{-3/2} 4^n$

Exemple des compositions :

- $D(z) = \frac{1-z}{1-2z}$ a sa sing. dominante en $1/2$, et c'est un pôle simple
- Le comportement de $[z^n]D(z) = 2^{n-1}$ est bien 2^n pour leur ordre exponentiel, et le terme sous-exponentiel est $\theta(n) = O(1)$ dans ce cas.

Séries génératrices bivariées et étude de paramètres

Paramètres et séries bivariées

Pour une classe combinatoire $\mathcal{C}$, un paramètre est une fonction $\chi : \mathcal{C} \to \mathbb{N}$ qui associe une valeur entière à tout objet de $\mathcal{C}$.

Exemples :

- Nombre de a dans un mot binaire
- Degré de la racine dans un arbre
- Nombre de points fixes dans une permutation

La série génératrice (ordinaire) bivariée de la classe (non-étiquetée) $\mathcal{C}$ pour le paramètre χ est

$$C(z,u) = \sum_{n \geq 0} \sum_{k \geq 0} c_{n,k} z^n u^k \quad \text{avec} \quad c_{n,k} = \#\{\gamma \in \mathcal{C} : |\gamma| = n \text{ et } \chi(\gamma) = k\},$$

où z et u sont des variables formelles.

Remarque : $C(z) = C(z,1)$ est la série génératrice ordinaire de $\mathcal{C}$.

Exemple “jouet”

On considère $\mathcal{W}$ = classe des mots binaires sur l'alphabet $\{a, b\}$, avec taille = nombre de lettres, et pour le paramètre χ = nombre de a

- Il y a $w_{n,k} = \binom{n}{k}$ mots de longueur n avec k lettres a dans $\mathcal{W}$.
- La série génératrice bivariée de $\mathcal{W}$, pour le paramètre “nombre de a”, est

$$W(z,u) = \sum_{n,k \geq 0} \binom{n}{k} z^n u^k = \frac{1}{1 - z(1+u)}, \quad \text{car}$$

$$W(z,u) = \sum_{n \geq 0} \left(\sum_{k=0}^{n} \binom{n}{k} u^k \right) z^n = \sum_{n \geq 0} (1+u)^n z^n = \frac{1}{1 - z(1+u)}.$$

Pour $u = 1$, on retrouve bien $W(z,1) = \frac{1}{1-2z}$.

Des séries bivariées à la moyenne

On considère une classe combinatoire $\mathcal{C}$, avec un paramètre χ.
On note $C(z)$ et $C(z,u)$ ses séries génératrices univariée et bivariée.
On note aussi c_n (resp. $c_{n,k}$) = nombre d'objets de taille n dans $\mathcal{C}$ (et qui ont pour valeur du paramètre k)

La moyenne (ou espérance) de χ sur les objets de taille n est notée $\mathbb{E}_n[\chi]$.
Par définition,

$$\mathbb{E}_n[\chi] = \frac{\sum_{\gamma \in \mathcal{C}_n} \chi(\gamma)}{\#\mathcal{C}_n} = \frac{\sum_k k \cdot c_{n,k}}{c_n}.$$

- dénominateur : $c_n = [z^n]C(z) = [z^n]C(z,1)$
- numérateur : $\sum_k k \cdot c_{n,k} = [z^n]\frac{dC(z,u)}{du}|_{u=1}$, car
$$\frac{dC(z,u)}{du} = \sum_{n \geq 0} \sum_{k \geq 1} c_{n,k} \cdot k u^{k-1} z^n.$$

Expressions comme quotients de coef. de séries bivariées

- $c_n = [z^n]C(z) = [z^n]C(z,1)$ est le nombre d'objets de taille n dans $\mathcal{C}$
- On a vu que la moyenne est

$$\mathbb{E}_n[\chi] = \frac{[z^n]\partial_u C(z,u)|_{u=1}}{[z^n]C(z,1)}.$$

Notation : ∂_u désigne la dérivée partielle par rapport à u

- Généralisation aux moments d'ordre supérieurs (variance, ect), pour raffiner notre compréhension de χ :

$$\mathbb{E}_n[\chi(\chi-1)\dots(\chi-r+1)] = \frac{[z^n]\partial_u^r C(z,u)|_{u=1}}{[z^n]C(z,1)}.$$

car $\frac{d^r C(z,u)}{du^r} = \sum_{n\geq 0}\sum_{k\geq r} c_{n,k}\cdot k(k-1)\dots(k-r+1)u^{k-r}z^n$.

Moments et moments factoriels

- Moment "classique" d'ordre r de χ : $\mathbb{E}_n[\chi^r]$
- Moment factoriel d'ordre r de χ : $\mathbb{E}_n[\chi(\chi-1)\dots(\chi-r+1)]$.

Leurs études sont équivalentes, car les moments "classiques" d'ordre inférieur ou égal à r peuvent être retrouvés à partir des moments factoriels d'ordre inférieur ou égal à r (et vice versa), comme combinaison linéaire.

Exemple : la variance, définie par $\mathbb{V}_n(\chi) = \mathbb{E}_n[\chi^2] - \mathbb{E}_n[\chi]^2$, s'exprime ainsi :

$$\mathbb{V}_n(\chi) = \mathbb{E}_n[\chi(\chi-1)] - \mathbb{E}_n[\chi](\mathbb{E}_n[\chi]-1).$$

Intérêt pour la variance, car c'est bon indicateur de la concentration (ou non) des valeurs d'un paramètre autour de sa valeur moyenne :

si l'écart type $\sqrt{\mathbb{V}_n(\chi)}$ est négligeable devant $\mathbb{E}_n[\chi]$, alors il y a concentration de χ autour de sa moyenne, pour des objets de grande taille.

Suivre un paramètre dans une spécification : les marqueurs

- Marqueur (noté μ) $\approx$ décoration accrochée sur certains atomes ou constructeurs dans un objet combinatoire.
- Il ne modifie pas la taille d'un objet.
- Si le paramètre compte certaines sous-structures d'un objet, un marqueur identifie chacune des sous-structures à compter.
  Ainsi, paramètre $=$ nombre total de marqueurs dans l'objet.

Dictionnaire : identique au cas univarié, sauf que la variable supplémentaire u est associée aux marqueurs μ (de la même façon que la variable z est associée aux atomes $\mathcal{Z}$).

Retour sur les mots binaires : spécification

- Spécification “classique” : $\mathcal{W} = \mathrm{Seq}(\mathcal{Z}_a \uplus \mathcal{Z}_b)$, où $\mathcal{Z}_a$ et $\mathcal{Z}_b$ sont deux classes atomiques, correspondant aux lettres a et b.
- Introduction d'un marqueur μ pour chaque lettre a:

$$\mathcal{W} = \mathrm{Seq}(\mu \mathcal{Z}_a \uplus \mathcal{Z}_b)$$

- Nombre de marqueurs μ = nombre de lettres a
- Équation pour la série bivariée $W(z, u)$ de $\mathcal{W}$, pour le paramètre χ = nombre de a :

$$W(z, u) = \frac{1}{1 - (zu + z)} \stackrel{=}{} \frac{1}{1 - (u + 1)z},$$

comme vu précédemment.

Retour sur les mots binaires : moyenne et variance

- On peut retrouver la moyenne $\mathbb{E}_n[\chi] = n/2$. En effet,

$$\partial_u W(z,u)|_{u=1} = \frac{z}{(1-2z)^2} = \sum_{n\geq 1} n2^{n-1}z^n,$$

ce qui donne $$\mathbb{E}_n[\chi] = \frac{[z^n]\partial_u W(z,u)|_{u=1}}{[z^n]W(z,1)} = \frac{n2^{n-1}}{2^n} = \frac{n}{2}.$$

- On peut aussi trouver la variance en calculant (pour $n \geq 2$) :

$$[z^n]\partial_u^2 W(z,u)|_{u=1} = [z^n]\frac{2z^2}{(1-2z)^3} = n(n-1)2^{n-2},$$

ce qui donne

$$\mathbb{E}_n[\chi(\chi-1)] = \frac{[z^n]\partial_u^2 W(z,u)|_{u=1}}{[z^n]W(z,1)} = \frac{n(n-1)2^{n-2}}{2^n} = \frac{n(n-1)}{4};$$

dont on déduit

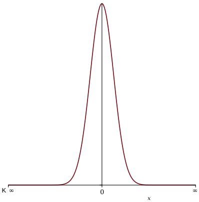
$$\mathbb{V}_n(\chi) = \mathbb{E}_n[\chi(\chi-1)] - \mathbb{E}_n[\chi](\mathbb{E}_n[\chi]-1) = \frac{n(n-1)}{4} - \frac{n}{2}\left(\frac{n}{2}-1\right) = \frac{n}{4}.$$

Conclusion de l'exemple sur les mots binaire

- Moyenne : $n/2$
- Variance : $n/4$
- Donc écart type $= \sqrt{n}/2$ négligeable devant la moyenne lorsque $n \to \infty$

Donc le nombre de lettres a est typiquement concentré autour de sa moyenne pour des objets de grande taille.

En fait, on verra que lorsque n est grand, le nombre de a est distribué à la limite comme une loi normale (= gaussienne).



Suite : des exemples !

Plus long préfixe de a dans un mot binaire 1/2

Spécification alternative pour $\mathcal{W}$:

$$\mathcal{W} = \text{Seq}(\mathcal{Z}_a) \times (\mathcal{E} \uplus \mathcal{Z}_b \times \text{Seq}(\mathcal{Z}_a \uplus \mathcal{Z}_b))$$

On peut la raffiner avec un marqueur associé au nombre de a dans le plus grand préfixe de a :

$$\mathcal{W} = \text{Seq}(\mu\mathcal{Z}_a) \times (\mathcal{E} \uplus \mathcal{Z}_b \times (\text{Seq}(\mathcal{Z}_a \uplus \mathcal{Z}_b)))$$

qui donne la série bivariée

$$V(z,u) = \left(1 + \frac{z}{1-2z}\right) \cdot \frac{1}{1-uz} = \frac{1-z}{1-2z} \cdot \frac{1}{1-uz}.$$

Remarque : on a bien toujours $V(z,1) = \frac{1}{1-2z}$

$$\partial_u V(z,u)|_{u=1} = \frac{z(1-z)}{(1-2z)(1-uz)^2}|_{u=1} = \frac{z}{(1-2z)(1-z)},$$

dont le développement en séries donne $\partial_u V(z,u)|_{u=1} = \sum_{n\geq 1}(2^n-1)z^n$

Donc la longueur moyenne du plus long préfixe de a dans un mot binaire de taille n est :

$$\frac{[z^n]\partial_u V(z,u)|_{u=1}}{[z^n]V(z,1)} = \frac{2^n-1}{2^n} = 1 - \frac{1}{2^n} \to 1 \text{ quand } n \to \infty$$

Calcul semblable pour la variance, qui tend vers 2 lorsque n tend vers l'infini.

Ici, on n'a pas concentration autour de la moyenne.

On verra qu'il y a convergence vers une loi géométrique de paramètre $1/2$.

Rappel : arbres de Catalan = arbres plans, enracinés, avec arité arbitaire.

La spécification $\mathcal{A} = \mathcal{Z} \times \mathrm{Seq}(\mathcal{A})$ peut être raffinée en

$$\tilde{\mathcal{A}} = \mathcal{Z} \times \mathrm{Seq}(\mu\mathcal{A}), \text{ où } \mathcal{A} \text{ est la classe définie ci-dessus,}$$

pour étudier le paramètre χ = degré de la racine.

La série bivariée correspondante est donc

$$A(z,u) = \frac{z}{1 - uA(z,1)}, \text{ où } A(z,1) = A(z) = \frac{1-\sqrt{1-4z}}{2}$$

De $A(z,u) = \frac{z}{1-uA(z,1)}$ avec $A(z,1) = A(z) = \frac{1-\sqrt{1-4z}}{2}$, on obtient

$$\partial_u A(z,u)|_{u=1} = \frac{2z(1-\sqrt{1-4z})}{(1+\sqrt{1-4z})^2},$$

qui se développe sous la forme

$$\partial_u A(z,u)|_{u=1} = \sum_{n\geq 2} \frac{3(2n-2)!}{(n+1)!(n-2)!} z^n.$$

Ainsi, le degré moyen de la racine parmi les arbres de Catalan de taille n est (pour $n \geq 2$)

$$\mathbb{E}_n[\chi] = \frac{\frac{3(2n-2)!}{(n+1)!(n-2)!}}{\frac{1}{n}\binom{2n-2}{n-1}} = 3\frac{n-1}{n+1},$$

Ce degré moyen est 3 à la limite lorsque n tend vers l'infini.

Nombre de feuilles dans les arbres binaires

- Arbres binaires décrits par $\mathcal{B} = \mathcal{E} \uplus \mathcal{Z} \times \mathcal{B} \times \mathcal{B}$
- Spécification alternative : $\mathcal{B} = \mathcal{E} \uplus \tilde{\mathcal{B}}$ avec $\tilde{\mathcal{B}}$ les arbres binaires non-vides, décrits par

$$\tilde{\mathcal{B}} = \mathcal{Z} \ \uplus \ \mathcal{Z} \times \tilde{\mathcal{B}} \ \uplus \ \mathcal{Z} \times \tilde{\mathcal{B}} \ \uplus \ \mathcal{Z} \times \tilde{\mathcal{B}} \times \tilde{\mathcal{B}}$$

- Raffinement pour suivre le nombre de feuilles :

$$\tilde{\mathcal{B}} = \mu\mathcal{Z} \ \uplus \ \mathcal{Z} \times \tilde{\mathcal{B}} \ \uplus \ \mathcal{Z} \times \tilde{\mathcal{B}} \ \uplus \ \mathcal{Z} \times \tilde{\mathcal{B}} \times \tilde{\mathcal{B}}$$

- Équation fonctionnelle pour $\tilde{B}(z,u) = B(z,u) - 1$:

$$\tilde{B}(z,u) = zu + 2z\tilde{B}(z,u) + z\tilde{B}(z,u)^2.$$

- On peut en déduire que le nombre moyen de feuille est asymptotiquement $n/4$

Nombre de cycles de taille 1 dans les permutations

- Un exemple dans le cadre étiqueté : permutations, et leur nombre de points fixes
- On avait $\mathcal{P} = \mathrm{Seq}(\mathcal{Z})$ et $\tilde{\mathcal{P}} = \mathrm{Set}(\mathrm{Cyc}(\mathcal{Z})) = \mathrm{Set}(\mathcal{Z} \uplus \mathrm{Cyc}_{\geq 2}(\mathcal{Z}))$
- On raffine la deuxième, avec un marqueur pour chaque point fixe (= cycle de taille 1) :

$$\tilde{\mathcal{P}} = \mathrm{Set}(\mu\mathcal{Z} \uplus \mathrm{Cyc}_{\geq 2}(\mathcal{Z}))$$

- la série génératrice (exponentielle) associée est $P(z,u) = \frac{\exp(z(u-1))}{1-z}$

En dérivant par rapport à u, on obtient

$$\partial_u A(z,u)|_{u=1} = \frac{z}{1-z} \quad \text{et} \quad \partial_u^2 A(z,u)|_{u=1} \; \frac{z^2}{1-z},$$

qui donne $\mathbb{E}_n[\chi] = 1$ et $\mathbb{V}_n(\chi) = 1$ pour $\chi =$ nombre de points fixes.

Comprendre le comportement d'un paramètre sur une classe combinatoire : Vers les lois limites

Pourquoi la moyenne ne suffit pas

Rappel du cadre de travail :

- une classe combinatoire $\mathcal{C}$ et un paramètre $\chi : \mathcal{C} \to \mathbb{N}$
- la série bivariée $C(z, u)$ permet d'exprimer la moyenne, la variance, et tous les moments (factoriels) de χ sur $\mathcal{C}_n$, pour tout n

La moyenne $\mathbb{E}_n[\chi]$ nous renseigne très partiellement sur le comportement de χ.

Exemple : Les deux cas suivants donnent la même moyenne, tout en ayant des comportements très différents (distributions différentes) :

- Classe $\mathcal{W}$ des mots binaires, $\chi =$ nombre de a

$\Rightarrow$ $\mathbb{E}_n[\chi] = n/2$, et χ peut prendre toutes les valeurs entre 0 et n

- Classe $\mathcal{Y} = \{a^n, b^n : n \geq 0\}$, $\chi =$ nombre de a

$\Rightarrow$ $\mathbb{E}_n[\chi] = n/2$, et χ prend seulement les valeurs 0 et n

Et pourquoi la variance, c'est mieux

Rappel : Une "petite" variance ($\sqrt{\mathbb{V}_n(\chi)} = o(\mathbb{E}_n[\chi])$) implique une concentration de χ autour de sa moyenne.

- Pour $\mathcal{W}$, on a vu qu'il y a ce phénomène de concentration.
- Pour $\mathcal{Y}$, clairement pas !

Plus formellement :

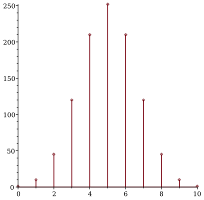
- $X_n =$ v. a. discrète qui décrit la distribution de χ sur $\mathcal{C}_n$
- Autrement dit : pour tout k, $\quad \mathbb{P}(X_n = k) = \frac{c_{n,k}}{c_n}$
- $\mu_n = \mathbb{E}_n[\chi]$ et $\sigma_n = \sqrt{\mathbb{V}_n(\chi)}$
- Sous l'hypothèse que $\lim_{n\to\infty} \frac{\sigma_n}{\mu_n} = 0$,
- on a, pour tout $\varepsilon > 0$, $$\lim_{n\to\infty} \mathbb{P}\left(1-\varepsilon \leq \frac{X_n}{\mu_n} \leq 1+\varepsilon\right) = 1.$$

Preuve : conséquence immédiate de l'inégalité de Bienaymé-Tchebychev

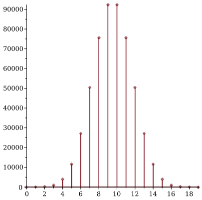
Au delà : convergence des distributions

On regarde l'histogramme de X_n, et on décrit sa "forme typique" lorsque n devient grand.

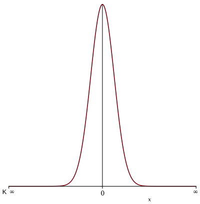
Exemple :
histogramme associé aux coefficients binomiaux $\binom{n}{k}$ (pour k de 0 à n)



$n = 10$



$n = 19$



$n \to \infty$

Il y a convergence de cette distribution (renormalisée) vers une loi normale.

Fin du programme

- Présentation de quelques distributions classiques (rappel, ou non)

- Principes pour démontrer la convergence en distribution en utilisant la combinatoire analytique
(Pas de détails, et pas de preuves)

- Des exemples !

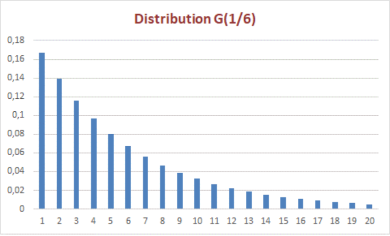
- Selon le temps disponible, un aperçu de mes recherches proches de ce thème

Loi géométrique

Étant donné un paramètre $p \in (0,1)$, la loi géométrique de paramètre p est définie par

$$\text{pour tout entier } k \geq 1, \quad \mathbb{P}(X = k) = (1-p)^{k-1} p.$$

Interprétation classique :
X désigne le nombre de tirages de Bernoulli indépendants, de probabilité de succès p, nécessaires pour obtenir le premier succès.



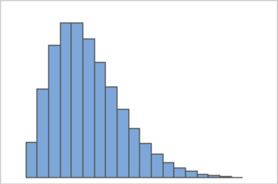
Loi binomiale négative

C'est une généralisation (d'une variante) de la loi géométrique.

La loi binomiale négative, de paramètres $p \in (0,1)$ et $n \geq 1$ (un entier fixé), est définie par

$$\text{pour tout entier } k \geq 0, \quad \mathbb{P}(X = k) = \binom{k+n-1}{k}(1-p)^k p^n.$$

Interprétation :
Ici, X compte le nombre d'échecs dans des tirages de Bernoulli indépendants, de probabilité de succès p, jusqu'à obtenir n succès.

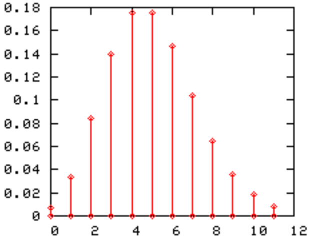


Loi de Poisson

La loi de Poisson, de paramètre λ un réel strictement positif, est définie par

$$\text{pour tout entier } k \geq 0, \quad \mathbb{P}(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}.$$

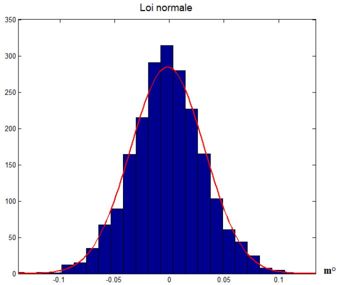
Interprétation :
Cette loi décrit le nombre d'événements "rares" dans un intervalle de temps donné (avec λ le nombre moyen d'événements se produisant dans cet intervalle de temps).



Loi normale = loi de Gauss = loi gaussienne

La loi normale (centrée réduite, ou standard) est la loi de probabilité sur $\mathbb{R}$ dont la densité est $f(t) = \frac{1}{\sqrt{2\pi}} e^{-t^2/2}$. C'est-à-dire qu'elle correspond à une variable aléatoire X telle que, pour tout $x \in \mathbb{R}$,

$$\mathbb{P}(X \leq x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} dt$$



Remarques :

- C'est la loi de probabilité associée à la fameuse courbe en cloche.
- C'est une distribution continue, alors que nos trois premiers exemples étaient des distributions discrètes.

Distributions limites discrètes et continues

Cadre de travail :

- $\mathcal{C}$ une classe combinatoire, χ un paramètre
- X_n = v. a. qui décrit la distribution de χ sur $\mathcal{C}_n$
- Autrement dit : pour tout k, $\mathbb{P}(X_n = k) = \frac{c_{n,k}}{c_n}$

Pour chaque n fixé, X_n est une v. a. discrète, et même à support fini.

Lorsque $n \to \infty$, la distribution limite de X_n peut être :

- soit discrète (avec support infini mais dénombrable)

$\rightsquigarrow$ souvent lorsque la moyenne (et la variance) de X_n est (sont) finie(s) (= indépendante(s) de n) à la limite

- soit continue (avec support indénombrable), pour une version renormalisée de la variable aléatoire X_n

$\rightsquigarrow$ souvent lorsque la moyenne est une fonction de n, tendant vers l'infini avec n

D'abord, une définition : série génératrice de probabilité

- La série bivariée associée à $\mathcal{C}$ et χ est $C(z,u) = \sum_{n\geq 0}\sum_{k\geq 0} c_{n,k} z^n u^k$
- Le polynôme $c_n(u) = \sum_{k\geq 0} c_{n,k} u^k = [z^n]C(z,u)$ caractérise la distribution de X_n. En effet,

$$\frac{[z^n]C(z,u)}{[z^n]C(z,1)} = \sum_{k\geq 0} \frac{c_{n,k}}{c_n} u^k = \sum_{k\geq 0} \mathbb{P}(X_n = k) u^k \stackrel{\text{def}}{=} \mathbb{E}_{\mathcal{C}_n}[u^\chi].$$

- On appelle $\mathbb{E}_{\mathcal{C}_n}[u^\chi]$ la série génératrice de probabilité du paramètre χ sur $\mathcal{C}_n$.

Rappel de combinatoire analytique :

- on étudie $C(z,u)$ en $u=1$, dans un voisinage de sa singularité dominante z_0

$\Rightarrow$ on obtient des informations sur $c_n = [z^n]C(z,1)$ (notamment son asymptotique)

Pour passer de $c_n = c_n(1)$ à $c_n(u)$:

- étendre cette analyse en $u = 1$ dans un voisinage fixé de 1
- extraire $[z^n]C(z,u)$ en voyant u comme un paramètre

La variable u (autour de $u = 1$) introduit une "déformation" de la série univariée $C(z)$, et on étudie comment cette déformation va perturber l'analyse de singularité.

On conclut grâce à des théorèmes de continuité (sous le tapis pour nous).

Discussion sur la taille du voisinage de $u = 1$ considéré :

- $C(z, u)$ en $u = 1 \Rightarrow$ énumération de la classe $\mathcal{C}$
- dans un voisinage infinitésimal de $u = 1 \Rightarrow$ moyenne et la variance de X_n
- dans un voisinage fixé de $u = 1 \Rightarrow$ distribution de X_n

$\Rightarrow$ Plus le voisinage est grand, plus les informations obtenues sont précises.

La robustesse de la combinatoire analytique permet souvent de traiter les déformations autour de $u = 1$ dans un voisinage fixé.

On a ainsi accès aux distributions limites *via* des théorèmes de continuité.

À suivre : des exemples

- Spécification raffinée : $\mathcal{W} = \mathrm{Seq}(\mu \mathcal{Z}_a) \times (\mathcal{E} \uplus \mathcal{Z}_b(\mathrm{Seq}(\mathcal{Z}_a \uplus \mathcal{Z}_b)))$
- Série bivariée : $V(z,u) = \frac{1}{1-uz} \cdot \frac{1-z}{1-2z}$, composée de deux facteurs
  - facteur $\frac{1-z}{(1-2z)}$
    où u n'apparaît pas, avec singularité dominante est en $z_0 = 1/2$
  - facteur $\frac{1}{1-uz}$
    qui n'a pas de singularité plus petite que z_0 tant que $|u| < 2$

$\Rightarrow$ on travaille dans le voisinage de $u = 1$ donné par $\{u : |u| < 2\}$

- Pour tout u_0 dans ce domaine, on a

$$V(z,u_0) \sim_{z \to 1/2} \frac{1}{1-1/2 \cdot u_0} \frac{1-1/2}{1-2z} = \frac{1}{2-u_0} \frac{1}{1-2z}$$

- dont on déduit (combi. analytique) $[z^n]V(z,u_0) \sim_{n\to\infty} \frac{1}{2-u_0} \cdot 2^n$
- ce qui signifie, pour la série génératrice de probabilité $\mathbb{E}_{\mathcal{W}_n}[u^{\text{nb de } a}] = \frac{[z^n]V(z,u_0)}{[z^n]V(z,1)}$ que $\mathbb{E}_{\mathcal{W}_n}[u^{\text{nb de } a}] \to_{n\to\infty} \frac{1}{2-u_0}$

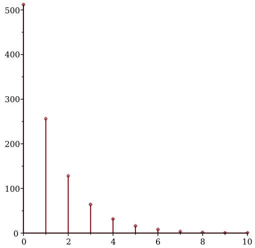
Plus long préfixe de a dans un mot binaire 2/3

- On a vu que, pour tout u_0 tel que $|u_0| < 2$, $\mathbb{E}_{\mathcal{W}_n}[u^{\text{nb de } a}] \rightarrow_{n\to\infty} \frac{1}{2-u_0}$
- On remarque que

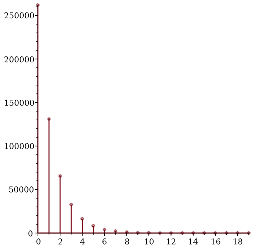
$$\frac{1}{2-u_0} = \sum_{k\geq 0} \frac{1}{2^{k+1}} u_0^k = \sum_{k\geq 0} (1-\tfrac{1}{2})^k \tfrac{1}{2} u_0^k$$

- C'est la série génératrice de probabilités associée à une loi géométrique de paramètre $1/2$
- Par théorème de continuité, la distribution de $X_n =$ longueur du plus long préfixe de a dans un mot binaire tend, lorsque n tend vers l'infini, vers une distribution géométrique de paramètre $1/2$.

Illustration sur les histogrammes :



$n = 10$



$n = 19$

Degré de la racine dans les arbres de Catalan

- Spécification raffinée : $\tilde{\mathcal{A}} = \mathcal{Z} \times \mathrm{Seq}(\mu\mathcal{A}),$ où $\mathcal{A} = \mathcal{Z} \times \mathrm{Seq}(\mathcal{A})$
- Série bivariée : $A(z,u) = \frac{z}{1-uA(z)}$, avec $A(z) = \frac{1-\sqrt{1-4z}}{2}$
- Ici, $[u^r]A(z,u) = zA(z)^r$ et on peut se passer de la série génératrice de probabilité.
- La distribution de $\chi =$ degré de la racine sur $\mathcal{A}_n$ est (par inversion de Lagrange, détails sous le tapis) :

$$\mathbb{P}_{\mathcal{A}_n}(\chi = r) = \frac{[z^{n-1}]A(z)^r}{[z^n]A(z)} = \frac{r}{n-1}\binom{2n-3-r}{n-1}\frac{n}{\binom{2n-2}{n-1}} \sim_{n\to\infty} \frac{r}{2^{r+1}}$$

- À la limite quand $n \to \infty$, χ (décalé de 1) suit une loi binomiale négative (de paramètres $(2, 1/2)$)

Nombre de cycles de taille 1 dans les permutations

- Spécification raffinée : $\tilde{\mathcal{P}} = \mathrm{Set}(\mu\mathcal{Z} \uplus \mathrm{Cyc}_{\geq 2}(\mathcal{Z}))$
- Série bivariée : $P(z,u) = \frac{\exp(z(u-1))}{1-z} = \frac{\exp(zu)\exp(-z)}{1-z}$
- Comme avant, on peut en extraire la distribution de χ = nombre de points fixes sur $\mathcal{P}_n$, en écrivant $\exp(zu) = \sum_k \frac{z^k}{k!} u^k$:

$$p_{n,k} = [z^n u^k] P(z,u) = [z^n] \frac{z^k}{k!} \frac{\exp(-z)}{1-z} = \frac{1}{k!} [z^{n-k}] \frac{\exp(-z)}{1-z}$$

- Pour finir, on écrit $\frac{\exp(-z)}{1-z} = \sum_{n \geq 0} (\sum_{j=0}^{n} \frac{(-1)^j}{j!}) z^n$ et on obtient

$$\text{pour tout } k, \quad p_{n,k} \to_{n\to\infty} \frac{1}{k!} \sum_{j \geq 0} \frac{(-1)^j}{j!} = \frac{e^{-1}}{k!}$$

$\Rightarrow$ la loi limite de χ est une loi de Poisson de paramètre 1.

Nombre de *a* dans un mot binaire, le (dernier) retour

- Spécification raffinée : $\mathcal{W} = \text{Seq}(\mu \mathcal{Z}_a \uplus \mathcal{Z}_b)$
- Série bivariée : $W(z,u) = \frac{1}{1-(u+1)z}$
- On peut pas "isoler" u comme dans les deux exemples précédents.
- Et contrairement au premier exemple, la singularité dominante varie avec u :

  pour tout u_0 fixé, la singularité dominante de $W(z,u_0)$ est en $\rho(u_0) = \frac{1}{1+u_0}$, et $[z^n]W(z,u_0) = \rho(u_0)^{-n}$.

$\Rightarrow$ La série génératrice de probabilité du nombre de a dans un mot binaire de longueur n est

$$\frac{[z^n]W(z,u_0)}{[z^n]W(z,1)} = \left(\frac{1}{2\rho(u_0)}\right)^n.$$

$\Rightarrow$ Loi limite gaussienne pour la v.a. renormalisée

Ouverture :
d’autres types de limites

Limite d'un paramètre *versus* limite d'objet

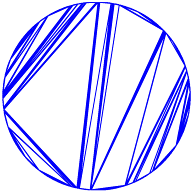
Ce qu'on a fait jusqu'ici :

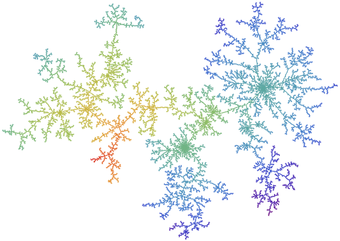
- Classe combinatoire $\mathcal{C}$, et un paramètre χ sur $\mathcal{C}$
- On cherche à répondre à la question "Comment se comporte typiquement χ sur les objets de $\mathcal{C}$ de grande taille ?"
- χ introduit un "filtre" à travers lequel on regarde les objets de $\mathcal{C}$

Autre manière d'envisager les limites d'objets combinatoires :

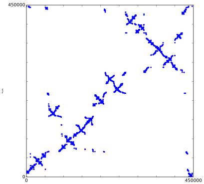
- Peut-on se passer de "filtre", et regarder directement les objets de $\mathcal{C}$?
- Peut-on décrire la "forme typique" d'un objet de $\mathcal{C}$ de grande taille ?
- Ça veut dire quoi qu'une suite d'objets (permutations, graphes, arbres, cartes, . . .) tend vers une limite ?

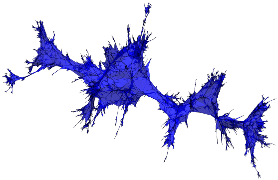
On peut donner un sens précis à ces questions. Les réponses obtenues donne une vision plus globale mais moins précise des objets étudiés.

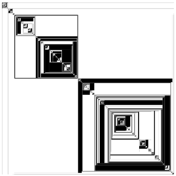




Merci !







Figures de
N. Curien, I. Kortchemski, C. Pivoteau, J. Bettinelli et M. Maazoun