

# **PROGRAMOZÁSI FELADAT 3**

## **Véletlen partíciók generálása**

**Vincze Erika**

# Tartalomjegyzék

<b>1. Bevezetés</b>	<b>3</b>
<b>2. A kínai étterem reprezentáció</b>	<b>5</b>
2.1. Konzisztens véletlen permutációk . . . . .	5
2.2. Egyszerű véletlen ülésrend . . . . .	5
2.3. A határeloszlások . . . . .	6
2.4. Általánosítás . . . . .	6
2.5. Véletlen ülésrend egy felcserélhető partícióhoz . . . . .	6
2.6. Véletlen ülésrend egy parciálisan felcserélhető partícióhoz . . . . .	7
2.7. A ciklusok számának együttes eloszlása . . . . .	7
<b>3. A kétparaméteres modell</b>	<b>9</b>
3.1. $(\alpha, \theta)$ ültetési rend . . . . .	9
<b>4. Véletlen egész számok és prímtényezős felbontásuk</b>	<b>11</b>
4.1. Prím faktorok együttes eloszlása . . . . .	11
4.2. Prím faktorok marginális eloszlása . . . . .	11
4.3. Prím faktorok határeloszlása . . . . .	12
<b>5. A kapott programokról</b>	<b>13</b>
5.1. A ChinesePartition program . . . . .	13
5.1.1. A program . . . . .	13
5.1.2. A leghosszabb ciklus mérete . . . . .	13
5.1.3. A ciklusok száma . . . . .	13
5.1.4. Adott hosszúságú ciklusok száma . . . . .	13
5.2. Véletlen ültetési rend adott valószínűségekhez . . . . .	13
5.3. A kétparaméteres modell . . . . .	14
5.3.1. A kétparaméteres program . . . . .	14
5.4. Számok prímtényezős felbontása . . . . .	14
5.4.1. Véletlen számok típusa . . . . .	14
5.4.2. Prímek multiplicitása . . . . .	14
5.5. További megjegyzés . . . . .	14

# 1. Bevezetés

A dolgozatban a véletlen partíciók generálásával foglalkozom. Ehhez  $[n] = \{1, 2, \dots\}$  egyenletes véletlen permutációját használom fel. A véletlen permutációk generálásának egyszerű módja a Kínai étterem folyamat. Főként ennek vizsgálatával foglalkozom.

Másik megközelítést ad a véletlen egész számok prímtényező felbontása, melynek során egy olyan  $c_i^{(n)}$  sorozathoz jutunk, amelyre  $C_1^{(n)}p_1 + C_2^{(n)}p_2 + \dots = n$ . Ekkor a kapott  $(C_1^{(n)}, C_2^{(n)}, \dots, C_n^{(n)})$  vektort vizsgálhatjuk.

A témához továbbá szorosan kapcsolódik a kétparaméteres modell, melynek megszorítása a Kínai étterem folyamat. A témakörben szereplő folyamatok határeloszlása különböző paraméterű Poisson-Dirichlet, illetve Griffith-Engen-McCloskey eloszlások.

Bizonyos urnamodellekhez, illetve utódlási folyamatokhoz is háttér ad az elmélet. A különböző leképezések elméletét Aldous vizsgálta, melynek eredményeképp a  $\theta = 1/2$  paraméterű Poisson-Dirichlet eloszlást kapta. Továbbá kapcsolat fedezető fel a Ewens-Sampling formulával, amely  $\theta = 1$  paraméterre az egyenletes eloszlású véletlen permutációkat jellemzi.

A ChinesePartition néven megírt program a Kínai étterem folyamatot reprezentálja. Ellenőrzöm rá a kívánt határeloszlástételek teljesülését. Érdekes lehet megfigyelni az eltéréseket a Mathematica beépített RandomPartition parancsával szemben. Szembeötlő, hogy a ChinesePartition lényegesen gyorsabban fut, nagyságrendekkel nagyobb számokra alkalmazható, ami a két parancs működési elvéből egyértelműen következik. A RandomPartition létrehozza az összes lehetséges permutációt, majd ebből választ véletlenszerűen, míg a ChinesePartition az új elemek bevitelével folyamatosan generálja a permutációt, így a partíciót is. A vizsgált szempontok (ciklusok száma, leghosszabb ciklus mérete, egy hosszú ciklusok száma) esetében is látszik a különbség a két módszerrel kapott partíciók között. A témakört érintő további megközelítésekhez is írtam megfelelő, őket generáló programot, így vizsgálható a kétparaméteres modell, illetve a prímtényező által létrejövő ciklusok tulajdonságai is megfigyelhetők.

## A kínai étterem reprezentáció

Ez a folyamat egy  $\sigma_n$  véletlen permutációt definiál az  $[n] := \{1, \dots, n\}$  halmazon úgy, hogy a  $\Pi_n$  véletlen partíciókat  $\sigma_n$  ciklusai generálják. Így a pozitív egész számok legáltalánosabb véletlen partícióihoz jutunk.

## A kétparaméteres modell

Ez a bekezdés legfőbbképp a természetes számok véletlen partícióval foglalkozik, melyet valós számok egy  $(\alpha, \theta)$  párjával paraméterezünk. Ezzel a  $GEM(\alpha, \theta)$ (Griffith-Engen-McCloskey), illetve sorba rendezett esetben a  $PD(\alpha, \theta)$  (Poisson-Dirichlet) elos-

zásokhoz jutunk.

### **Véletlen egészek prímtényezős felbontása**

Ha  $N(n)$  az  $[n] = \{1, 2, \dots, n\}$  halmaz egy véletlen választott eleme, és  $N(n) = p_1^{c_1} p_2^{c_2} \dots$  felírásában  $p_1 < p_2 < \dots$  prímtényezőik  $c_i$  kitevői vizsgálhatók.

## 2. A kínai étterem reprezentáció

### 2.1. Konzisztens véletlen permutációk

Tekintsük a  $(\sigma_n, n = 1, 2, \dots)$  véletlen permutációk egy sorozatát a következőképpen:

1.  $\sigma_n$  egyenletes véletlen permutációja  $[n]$ -nek, minden  $n$ -re.
2. Ha  $\sigma_n$  adott a ciklusokból minden  $n$ -re, akkor  $\sigma_{n-1}$ -et kapjuk, ha  $\sigma_n$  ciklusából töröljük az  $n$ . elemet.

Például, a sztenderd jelöléseket használva:

ha  $\sigma_5 = (134)(25)$  akkor  $\sigma_4 = (134)(2)$ ;

ha  $\sigma_5 = (134)(2)(5)$  akkor  $\sigma_4 = (134)(2)$ .

Könnyen belátható, hogy a fenti eljárással a következő eljárásnak megfelelő eloszlású  $\sigma_n$  sorozathoz jutunk:

Egy kezdetben üres étteremben korlátlan számú kerek asztal áll rendelkezésre,  $1; 2; \dots$  számozással, és minden egyes asztalhoz korlátlan számú ember ülhet. A vendégek is  $1; 2; \dots$  számozásúak, és egyenként foglalnak helyet a következők szerint:

### 2.2. Egyszerű véletlen ülésrend

Az első ember az első asztalhoz ül. Az  $n + 1$ -hez tegyük fel, hogy már  $n$  ember helyet foglalt valamelyik asztalnál úgy, hogy minden  $j$ -re  $(1 \leq j \leq k)$ , ahol  $k$  az olyan asztalok száma, amelynél az első  $n$  ember ül. Az  $n + 1$ . vendég a következő  $n + 1$  hely közül egyenlő valószínűséggel választ: a  $j$ . vendég bal oldalára  $(1 \leq j \leq n)$ , vagy egyedül ül a  $k + 1$ . asztalhoz.

Legyen  $\sigma_n : [n] \rightarrow [n]$  a következő: ha az  $n$ . vendég érkezése után az  $i$ . és a  $j$ . vendég egy asztalnál ül,  $i$ . a  $j$ . bal oldalán, akkor  $\sigma_n(i) = j$ , és ha az  $i$ . vendég egyedül ül egy asztalnál, akkor  $\sigma_n(i) = i$ .

$\sigma_n$  ekkor kielégíti az (1) illetve a (2) feltételeket.

Ebből a konstrukcióból az egyenletes véletlen permutációk sok tulajdonsága kiolvasható. Az  $n$ . vendég érkezése után az asztaloknál ülő emberek száma a következőképpen alakul:

$$K_n = \{\sigma_n \text{ ciklusai}\} = Z_1 + \dots + Z_n$$

ahol  $Z_j$  az indikátora annak az eseménynek, hogy a  $j$ . vendég új asztalhoz ül. A konstrukció miatt a  $Z_j$ -k független Bernoulli( $1/j$ ) valószínűségi változók, ezért:

$$\frac{K_n}{\log n} \rightarrow 1 \text{ majdnem biztosan, és } \frac{K_n - \log n}{(\log n)^{1/2}} \xrightarrow{d} B_1,$$

ahol  $B_1$  a standard Gauss eloszlás.

Legyen  $\Pi_n$  a  $\sigma_n$  ciklusai által generált véletlen partíciója  $[n]$ -nek. Ekkor  $\Pi_n$  az  $[n]$  felcserélhető véletlen partíciója, és  $\Pi_n$  konzisztens, ahogy  $n$  változik. Így a  $\Pi_\infty = (\Pi_n)$  sorozat az  $N$  egy felcserélhető véletlen partíciója. Legyen  $X_n$  annak az eseménynek az indikátora, hogy az  $n + 1$ . vendég az első asztalhoz ül. Ekkor az  $(X_n)_{n \geq 1}$  egy felcserélhető sorozat, ami a Pólya-féle urnamodellnek az eredménye,  $a = b = 1$ -re. Ezért  $S_n = X_1 + \dots + X_n$  egyenletes eloszlású a  $\{0; 1; \dots; n\}$ -on. Ugyanígy az  $S_n + 1$  méretű ciklusa az  $1$ -et tartalmazó  $\sigma_{n+1}$ -nek egyenletes eloszlású a  $\{1; \dots; n+1\}$ -en. Az aszimptotikus gyakorisága a  $\sigma_n$   $1$ -et tartalmazó osztályainak, majdnem biztosan korlátos,  $S_n = n$ -nel, ami természetesen egyenletes eloszlású a  $[0; 1]$ -en.

### 2.3. A határeloszlások

Legyen  $(N_{n,1}; \dots; N_{n,K_n})$  a  $\Pi_n$  blokkjainak mérete, az elemek megjelenése szerinti sorrendben. Az éttermes modell szerint  $N_{i,n}$  az  $i$ . asztalnál ülők száma, amikor  $n$  ember van az étteremben. Az előzőek szerint  $N_{n,1}$  egyenletes eloszlású  $[n]$ -en. Hasonlóan, adott  $N_{n,1} = n_1 < n$ , esetén  $N_{n,2}$  is egyenletes eloszlású  $[n-n_1]$ -en, és így tovább. Ez a diszkrét egyenletes "stick-breaking" séma, melynek eloszlása, ha  $n \rightarrow \infty$ , a relatív gyakorisága  $\sigma_n$  ciklusainak a méretének  $(N_{n,i}/n; i \geq 1)$  tart a folyamatos egyenletes "stick-breaking" folyamathoz eloszlásban.

$$(\tilde{P}_1, \tilde{P}_2, \dots) = (U_1, \bar{U}_1 U_2, \dots),$$

ahol  $U_i$ -k független  $U[0, 1]$  változók, és  $\bar{U} = 1 - U$ . Ekkor a hozzá tartozó EPPF:

$$p_{0,1}(n_1, \dots, n_k) := \frac{1}{n!} \prod_{i=1}^k (n_i - 1)!.$$

**2.1. Megjegyzés:.** *EPPF (exchangeable partition probability function):*

Az  $[n]$  bármely  $\{A_1, \dots, A_k\}$  partíciójára

$$\mathbb{P}(\pi_n = \{A_1, \dots, A_k\}) = p(|A_1|, \dots, |A_k|)$$

az  $n$   $(n_1, \dots, n_k)$  kompozíciójának valamely  $p$  szimmetrikus függvénye. Ezt a  $p$  függvényt nevezzük a  $\pi_n$  EPPF-jének.

### 2.4. Általánosítás

A kínai étterem folyamat könnyen általánosítható, ha a  $\sigma_n$  permutációk sorozatához tartozó  $\pi_\infty := (\pi_n)$  a legáltalánosabb felcserélhető partíciója az egész számoknak. A hozzá tartozó EPPF:  $p(n_1, \dots, n_k)$  megadja minden  $(n_1, \dots, n_k)$ -ra annak a valószínűségét, hogy  $\pi_n$  az  $[n]$ -et éppen  $(n_1, \dots, n_k)$  méretű halmazokra osztja. A kínai éttermes megfontolásból most csak a  $\sigma_n$  ciklusok által generált partíciókkal foglalkozunk. Így  $1 \leq i \leq K_n$ -re az állítás, miszerint az  $n + 1$ . látogató az  $i$ . asztalhoz ül, azt jelenti, hogy  $\pi_{n+1}$  az  $[n + 1]$  olyan partíciója, melynek  $[n]$ -re való megszorítása  $\pi_n$ , és  $n + 1$  a  $\pi_n$   $i$ . osztályához tartozik. Hasonlóan, ha az  $n + 1$ . vendég új asztalhoz ül, akkor  $n_k$ . Így  $1 \leq i \leq K_n$ -re az állítás, miszerint az  $n + 1$ -edik látogató az  $i$ . asztalhoz ül, azt jelenti, hogy  $\pi_{n+1}$  az  $[n + 1]$  olyan partíciója, melynek  $[n]$ -re való megszorítása  $\pi_n$  egy  $\{n + 1\}$  külön blokkal. Adott végtelen EPPF-hez tartozó felcserélhető random partíciója a természetes számoknak konstruálható a következőképpen:

### 2.5. Véletlen ülésrend egy felcserélhető partícióhoz

Az első vendég az első asztalhoz ül. Ha  $n$ -ig adott  $\pi_n$  partíció, az első  $n$  ember helyével,  $k$  asztallal, akkor az  $n + 1$ . vendég:

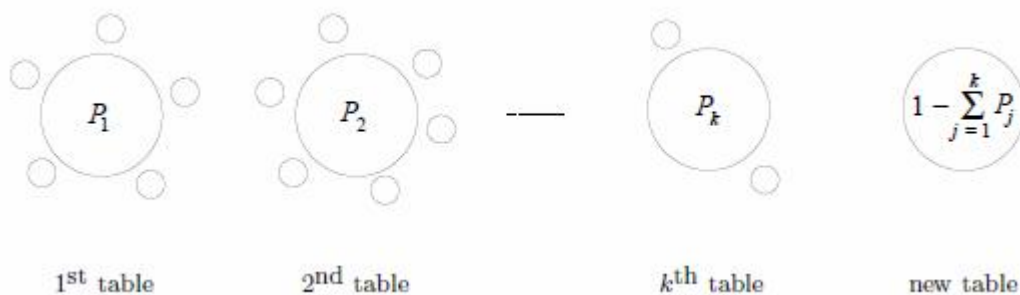
- a  $j$ . asztalhoz ül  $p(\dots, n_j + 1, \dots)/p(n_1, \dots, n_k)$  valószínűséggel
- új asztalhoz ül  $p(n_1, \dots, n_k, 1)/p(n_1, \dots, n_k)$  valószínűséggel.

Ez megfelel a feltételeknek. Az ültetési rendek legtöbbje által generált  $\pi_\infty$  nem felcserélhető. Kingman elmélete szerint annak, hogy felcserélhető partíciót kapjunk, szükséges feltétele, hogy minden  $i$ -re létezik egy majdnem biztos határeloszlás az  $i$ . asztalnál ülők számára. Ennek elérésének a legegyszerűbb módja a következő:

## 2.6. Véletlen ülésrend egy parciálisan felcserélhető partícióhoz

Legyen  $P_i, i = 1, 2, \dots$  nem negatív véletlen változók tetszőleges sorozata úgy, hogy  $\sum_{i=1}^k P_i \leq 1$ . Adott  $P_i, i = 1, 2, \dots$  sorozathoz az első ember üljön az első asztalhoz, és ha  $n$ -ig adott az ülésrend  $k$  asztallal, akkor az  $n + 1$ . ember:

- a  $j$ . helyre ül  $P_j$  valószínűséggel
- új asztalhoz ül  $1 - \sum_{i=1}^k P_i$  valószínűséggel.



A konstrukció és a nagy számok törvénye miatt így léteznek a határeloszlások, melyek éppen a  $P_i$ -kel egyeznek meg. Az adott  $P_i$  sorozatra meghatározható annak a valószínűsége, hogy  $[n]$  éppen  $(n_1, \dots, n_k)$  méretű blokkokra bomlik. Ez a következő formulával fejezhető ki:

$$p(n_1, \dots, n_k) = \mathbb{E}\left[\prod_{i=1}^k P_i^{n_i-1} \prod_{i=1}^{k-1} (1 - \sum_{j=1}^i P_j)\right] \quad (1)$$

**2.1. Tétel.** Legyen  $(P_i)$  nem negatív véletlen változók sorozata úgy, hogy  $\sum_{i=1}^k P_i \leq 1$ .  $p(n_1, \dots, n_k)$ -t pedig az előző formulával határozzuk meg.

1.  $\exists \pi_\infty$  felcserélhető véletlen partíciója a természetes számoknak, melynek blokkjainak sűrűsége megjelenési sorrendben  $(\tilde{P}_i)$  eloszlása  $(P_i)$  akkor és csak akkor, ha  $p(n_1, \dots, n_k)$  függvény szimmetrikus függvénye  $(n_1, \dots, n_k)$ -nek minden  $k$ -ra.
2. Ha  $\pi_\infty$  ilyen felcserélhető véletlen partíciója a természetes számoknak, akkor a  $\pi_\infty$ -hez tartozó EPPF:  $p(n_1, \dots, n_k)$  a 1. képlet alapján adódik  $P_i = \tilde{P}_i$ -vel.

## 2.7. A ciklusok számának együttes eloszlása

Egy permutáció ciklusait megadhatjuk a következőképpen:  $c = (c_1, \dots, c_n) \in Z^+$ ; ahol  $c_j$  a  $j$  méretű ciklusok számát jelöli. Ekkor  $c$ -t a permutáció típusának nevezzük. A  $c$

szerint adott  $c_i$  méretű ciklusokkal rendelkező permutációk számát jelöljük  $N(n, c)$ -vel!  $N(n, c)$  kiszámítására a Cauchy-formulát használhatjuk:

$$N(n, c) = \mathbf{1} \left\{ \sum_{j=1}^n j c_j = n \right\} \prod_{j=1}^n \left( \frac{1}{j} \right)^{c_j} \frac{1}{c_j!},$$

Ha a permutációt egyenletesen és véletlenül választottuk a lehetséges  $n!$  számú  $S_n$  permutációból, akkor a  $j$  hosszú ciklusok  $C_j^{(n)}$  számai függenek egymástól. Ekkor az együttes eloszlás szintén a Cauchy formulából következik, és a következőképp adható meg:

$$\mathbb{P}[C^{(n)} = c] = \frac{1}{n!} N(n, c) = \mathbf{1} \left\{ \sum_{j=1}^n j c_j = n \right\} \prod_{j=1}^n \left( \frac{1}{j} \right)^{c_j} \frac{1}{c_j!},$$

valamely  $c$  pozitív egész számra.

**2.2. Tétel.** *A ciklusok számlálója eloszlásban a természetes számokon értelmezett  $j^{-1}$  várható értékű Poisson folyamathoz tart. Vagyis, ha  $n \rightarrow \infty$ ,*

$$(C_1^{(n)}, C_2^{(n)}, \dots) \xrightarrow{d} (Z_1, Z_2, \dots),$$

ahol  $Z_j$ ,  $j = 1, 2, \dots$  független Poisson eloszlású véletlen változók  $\frac{1}{j}$  várható értékkel.

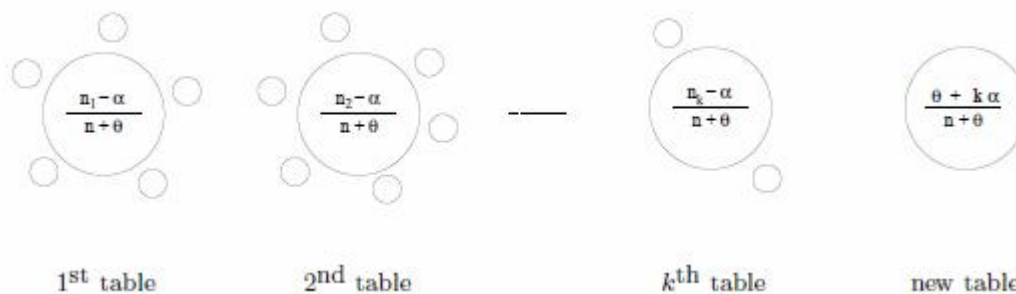


### 3. A kétparaméteres modell

#### 3.1. $(\alpha, \theta)$ ültetési rend

Az  $n$ . lépésben  $k$  elfoglalt asztallal az  $i$ . asztalnál ülők száma  $n_i$ . A következő vendég

- az  $i$ . asztalhoz ül  $(n_i - \alpha)/(n + \theta)$  valószínűséggel
- új asztalhoz ül  $(\theta + k\alpha)/(n + \theta)$ .



A valószínűségek tulajdonságainak megfelelően  $\alpha$  és  $\theta$  a következők szerint választhatók.

- $\alpha = -\kappa < 0$  és  $\theta = m\kappa$  valamely  $m = 1, 2, \dots$  esetén.
- vagy  $0 \leq \alpha \leq 1$  és  $\theta > -\alpha$ .

Ekkor a következő eseteket vizsgáljuk:

$(\alpha = -\kappa < 0)$  és  $\theta = m\kappa$  valamely  $m = 1, 2, \dots$  esetén:  
Ekkor  $\pi_\infty$  eloszlása  $m\kappa$  paraméterű szimmetrikus Dirichlet.

$\alpha = -\kappa$  és  $\theta > 0$  esetén:

Ez az előző eset gyenge limesze, ha  $\kappa \rightarrow 0$  és  $m\kappa \rightarrow \theta$ . A Blackwell-MacQueen urnamodellhez hasonlóan  $\pi_\infty$  eloszlása  $\theta$  paraméterű Dirichlet folyamatból vett mintáknak felel meg.

$\alpha = 0$  és  $\theta = 1$  esetén:

Az előző modellt adja.

$0 < \alpha < 1$  és  $\theta > -\alpha$  esetén:

Az  $\alpha$  indexű "stable subordinator"-ral rokon esethez vezet.

**3.1. Tétel.:** Tegyük fel, hogy a természetes számok egy  $\pi_\infty$  felcserélhető véletlen partíció blokkjainak  $\bar{P}_j$  a sűrűsége (az elemek megjelenési sorrendjében) úgy, hogy  $0 < \bar{P}_1 < 1$  majnem biztosan, és igaz rá a következők valamelyike:

1.  $\pi_\infty [n]$ -re való megszorítása  $\pi_n \text{Gibbs}_{[n]}(v_\bullet, w_\bullet)$  partíció, ami azt jelenti, hogy a hozzá tartozó EPPF szorzat alakú, nem negatív  $v_\bullet$  és  $w_\bullet$  sorozatok párjaira, vagy
2. a  $\bar{P}_j$  sűrűségek szorzat alakúak, valamely  $W_i$  véletlen változókra.

Ekkor  $\pi_\infty$  eloszlását vagy a kétparaméteres modell determinálja valamely  $(\alpha, \theta)$  párra, vagy "coupon collectors partition" valamely  $m = 2, 3, \dots$ -ra.

## 4. Véletlen egész számok és prímtényező felbontásuk

A véletlen permutációk témája megközelíthető a véletlen egész számok prímtényező felbontásán keresztül is. Vizsgálhatjuk a hasonlóságot, illetve az eltéréseket a korábbi véletlen permutációkhoz képest.

Minden egész szám egyértelműen felbontható prímek szorzatára. Például  $m = 220$  prímfelbontásában a 2 kettő, az 5 egy és a 11 is egy kitevővel szerepel. Ekkor jellemezhetjük  $m$ -et egy  $c$  típus jelzővel, ahol  $c$  egy vektor, amely a  $(c_p)$  számokból tevődik össze. A  $(c_p)$ -ket pedig az adja meg, hogy a  $p$  prímtényező mekkora kitevővel szerepel az adott szám felbontásában. Tehát  $c = (c_2, c_3, c_5, \dots)$ , úgy, hogy  $m = \prod p^{c_p}$ . Esetünkben, az  $m = 220$  típusa a következő:  $c = (2, 0, 1, 0, 1, 0, 1, \dots)$ .

### 4.1. Prím faktorok együttes eloszlása

Adott  $c$  típusra 1 vagy 0 neki megfelelő egész létezik az  $[n] = \{1, 2, \dots, n\}$  halmazból, attól függően, hogy  $\prod_{p \leq n} p^{c_p} \leq n$  teljesül-e. Legyen  $N = N(n)$  az  $[n]$  számból egyenletesen véletlen választott egész szám. Az  $N(n)$  prímtényező felbontásában szereplő  $p$  prímek  $C_p^{(n)}$  kitevői véletlen változók. Amennyiben  $p > n$ ,  $C_p^{(n)} = 0$  adódik. Így

$$N(n) = \prod_{p \leq n} p^{C_p^{(n)}} = \prod_p p^{C_p^{(n)}}$$

egyenletes eloszlású  $[n]$ -en és  $C^{(n)} = (C_2^{(n)}, C_3^{(n)}, \dots)$  adott a következő képlettel:

$$\mathbb{P}[C^{(b)} = c] = \frac{1}{n} \mathbf{1} \left\{ \sum_p c_p \log p \leq \log n \right\}, c \in Z_+^\infty.$$

### 4.2. Prím faktorok marginális eloszlása

Egy  $k \geq 0$  számra az, hogy  $C_p^{(n)} \geq k$  azt jelenti, hogy  $N(n)$  a  $p^k$  többszöröse. Ekkor  $N(n) \in [n] \lfloor n/p^k \rfloor$  értéket vehet fel. Tehát

$$\mathbb{P}[C_p^{(n)} \geq k] = \frac{1}{n} \lfloor n/p^k \rfloor, \quad (2)$$

amiből pedig

$$\mathbb{P}[C_p^{(n)} = k] = \frac{1}{n} (\lfloor n/p^k \rfloor - \lfloor n/p^{k+1} \rfloor). \quad (3)$$

Az együttes eloszlást is jellemezhetjük: jelölje  $\pi(b)$  a  $b$ -nél kisebb vagy egyenlő prímek számát. Ekkor valamely  $c = (c_p, p \leq b) \in Z_+^{\pi(b)}$  esetén

$$\mathbb{P}[C_p^{(n)} \geq c_p, p \leq b] = \frac{1}{n} \lfloor n/d \rfloor,$$

ahol  $d = \prod_{p \leq b} p^{c_p}$ .

### 4.3. Prím faktorok határeloszlása

A 2 képletből egyértelműen következik, hogy bármely  $p$  prímre és  $k$  pozitív egész számra,

$$\mathbb{P}[C_p^{(n)} = k] \rightarrow \left(1 - \frac{1}{p}\right) \left(\frac{1}{p}\right)^k,$$

ha  $n \rightarrow \infty$ , vagyis  $C_p^{(n)}$  eloszlásban egy  $1/p$  paraméterű geometriai eloszlású  $Z_p$  véletlen változóhoz tart. Hasonlóan a 3 képletből  $c = (c_p, p \leq b) \in Z_+^{\pi(b)}$ , ha  $n \rightarrow \infty$ ,

$$\mathbb{P}[C_p^{(n)} = c_p, p \leq b] \rightarrow \prod_{p \leq b} \left(1 - \frac{1}{p}\right) \left(\frac{1}{p}\right)^{c_p},$$

ahol a jobb oldal éppen  $\mathbb{P}[Z_p = c_p, p \leq b]$ -nek felel meg, vagyis a geometriai eloszlású  $Z_p$  valószínűségi változók függetlenek. Tehát

$$(C_2^{(n)}, C_3^{(n)}, \dots) \xrightarrow{d} (Z_2, Z_3, \dots).$$

## 5. A kapott programokról

A Mathematica-val készített programokban a következő eredményekre jutottam, illetve az alábbi kérdéseket vizsgáltam:

### 5.1. A ChinesePartition program

#### 5.1.1. A program

A ChinesePartition nevű program a Kínai étterem folyamat szerint hoz létre véletlen partíciót. Ennek másik része a ChinesePartitionuj, amelyben az érkezők sorrendjét is nyomon követhetjük.

#### 5.1.2. A leghosszabb ciklus mérete

A ChinesePartitionnal kapott legnagyobb ciklus méretét a RandomPartitionnal generált permutációkéval vetem össze. Mivel a Kínai étterem folyamat esetében nagyobb az esélye, hogy az új érkező nagyobb asztalhoz ül, így a nagyobb ciklusok kialakulása valószínűbb, mint ha véletlen választunk az összes permutáció közül. A kapott ábrán ez szépen látszik.

#### 5.1.3. A ciklusok száma

A ciklusok  $K_n$  számára teljesül, hogy  $\frac{K_n}{\log n} \rightarrow 1$ , melyet a számítások is alátámasztanak. Ha több permutációt is létrehozunk és ezek átlagát vizsgáljuk, akkor egyértelműbben erre az eredményre jutunk.

A ciklusok számára is a várt adatokat kapjuk a RandomPartitionnal szemben, ugyanis a ChinesePartition esetében kisebb az esély egy új ciklus létrejöttének. Így a ciklusok száma nagyobb, ha a RandomPartitionnal hozzuk létre.

A ciklusok hosszainak összeszámlálására a ciklus nevű program használható.

#### 5.1.4. Adott hosszúságú ciklusok száma

Az 1 hosszú ciklusok száma is jelentős eltérést mutat a RandomPartitionnal szemben, hiszen a Kínai étterem folyamat esetében a konstrukcióból következik, hogy kisebb valószínűséggel jön létre új ciklus, illetve a meglévő ciklusok nagyobb valószínűséggel nőnek meg.

A végzett számítások továbbá alátámasztják, hogy a  $j$  méretű ciklusok arányának a várható értéke  $1/j$ .

### 5.2. Véletlen ültetési rend adott valószínűségekhez

Itt létrehoztam egy megfelelő sorozatot, melyek az adott asztalokhoz ülés valószínűségeit adják. Ehhez a korábban említett módszer olyan véletlen ültetési rendet hoz létre, melynek éppen az adott valószínűségek sorozata a határértéke. Erre egy-egy példát generáltam.

## **5.3. A kétparaméteres modell**

### **5.3.1. A kétparameteres program**

A kétparameteres nevű program a paraméterek függvényében létrehozza a partíciókat. Ennek határeloszlása is látható a keletkezett ábrán.

## **5.4. Számok prímtényezős felbontása**

A számok  $c$ -vel jelölt típusa kapható az itt megírt programokkal.

### **5.4.1. Véletlen számok típusa**

A randomtípus egy adott számig véletlen választott szám típusát adja meg, a típus pedig egy általunk választott egészre számol.

### **5.4.2. Prímek multiplicitása**

A multipl program egy általunk választott prím kitevőjét adja meg egy adott számig véletlen kapott szám prímtényezős felbontásában.

## **5.5. További megjegyzés**

Az általam írt ChinesePartition használható akár  $10^6$  nagyságrendre is, míg a beépített RandomPartition függvény csak ezres nagyságrendig működik. Ez jól látszik, ha megvizsgáljuk a két program futásidejét.

## Hivatkozások

- [1] Jim Pitman: Combinatorial Stochastic Processes
- [2] Richard Arratia, A.D. Barbour, Simon Tavaré : Logarithmic Combinatorial Structures: a Probabilistic Approach
- [3] S. Tavaré, W. J. Ewens: Multivariate Ewens Distribution (41. Fejezet)