

- (1) Si l'on calcule la somme des probabilités de transition à partir d'une partition π dont la liste des tailles est $(c_1, c_2, \dots, c_\ell)$, alors on obtient :

$$\begin{aligned}\sum_{k=1}^{\ell+1} P(\pi, \pi[k]) &= \frac{\theta + \alpha \ell}{\theta + n} + \left(\sum_{k=1}^{\ell} \frac{c_k - \alpha}{\theta + n} \right) \\ &= \frac{1}{\theta + n} \left(\theta + \sum_{k=1}^n c_k \right) = \frac{\theta + n}{\theta + n} = 1.\end{aligned}$$

- (2) On fait bien sûr précéder tous nos programmes des imports :

```
import numpy as np
import numpy.random as rand
import scipy.stats as scs
import matplotlib.pyplot as plt
```

Le programme suivant calcule $\pi^{(N)}$ pour la chaîne de Markov de paramètres (θ, α) :

```
def update_partition(pi, N, theta, alpha):
    # calcul du noyau de transition
    P = np.array([len(part) for part in pi] + [0]) - alpha
    P[-1] += alpha*(len(pi) + 1) + theta
    P = P / (N + theta)
    # choix du successeur
    k = rand.choice(len(pi) + 1, p=P)
    if k < len(pi):
        pi[k].append(N+1)
    else:
        pi.append([N+1])

def markov_partition(N, theta, alpha):
    pi = [[1]]
    for k in range(1, N):
        update_partition(pi, k, theta, alpha)
    return pi
```

Voici un exemple d'appel du programme :

```
>>> markov_partition(15, 1, 0.5)
[[1, 2, 3, 5, 8, 10, 13, 14, 15], [4, 12], [6], [7, 9], [11]]
```

Si l'on connaît $\pi^{(N)}$, alors par cohérence de la suite, on peut récupérer toute partition $\pi^{(n)}$ avec $n < N$: il suffit de prendre les intersections non vides des parts de $\pi^{(N)}$ avec $[1, n]$. Par exemple, si $\pi^{(15)}$ est l'exemple ci-dessus, alors $\pi^{(10)} = (\{1, 2, 3, 5, 8, 10\}, \{4\}, \{6\}, \{7, 9\})$.

- (3) Voici un programme de dessin qui donne une couleur différente à chaque part des partitions.
-

```
def dessin_partition(pi):
    fig, ax = plt.subplots()
    L = len(pi)
    n = sum([len(part) for part in pi])
    # colors est un tableau dont la k-ieme ligne contient la couleur de la
    # k-ième fonction à dessiner
    colors = np.zeros((L, 3))
    colors[:, 0] = 0.2 + 0.6*np.arange(L)/L
```

```

colors[:, 2] = 0.8 - 0.6*np.arange(L)/L
# C est un tableau dont la k-ième ligne est l'ensemble des ordonnées
# de la k-ième fonction à dessiner
C = np.zeros((n+1, L))
for p in range(L):
    for k in pi[p]:
        C[k:, p:] += 1
# dessin ; on remplit également l'espace entre deux courbes consécutives
for p in range(L-1, -1, -1):
    ax.fill_between(np.arange(n+1), np.zeros(n+1), C[:, p],
                    color = tuple(colors[p, :] + np.array([0, 0.3, 0])))
    ax.plot(np.arange(n+1), C[:, p], color = tuple(colors[p, :]))
ax.set_aspect(1)
plt.show()

```

- (4) Pour étudier les statistiques de la suite $(\ell^{(n)})_{n \geq 1}$ définie par $\ell^{(n)} = \ell(\pi^{(n)})$, on pourrait utiliser le programme `markov_partition`, mais il est plus efficace de reprogrammer la suite. En effet, réalisons les transitions de la chaîne de Markov $(\pi^{(n)})_{n \geq 1}$ en utilisant une suite $(\xi^{(n)})_{n \geq 1}$ de variables indépendantes uniformes sur $[0, 1]$, et en posant :

$$\pi^{(n+1)} = \begin{cases} \pi^{(n)}[k] & \text{si } \sum_{j=1}^{k-1} (c_j(\pi^{(n)}) - \alpha) \leq (\theta + n)\xi^{(n)} < \sum_{j=1}^k (c_j(\pi^{(n)}) - \alpha), \\ \pi^{(n)}[\ell^{(n)} + 1] & \text{si } (\theta + n)\xi^{(n)} \geq n - \ell^{(n)}\alpha. \end{cases}$$

On vérifie aisément que cette réalisation donne les bonnes probabilités de transition. Remarquons alors qu'avec cette réalisation, on a : $\ell^{(n+1)} = \ell^{(n)} + 1_{(\xi^{(n)} \geq \frac{n - \ell^{(n)}\alpha}{\theta + n})}$.

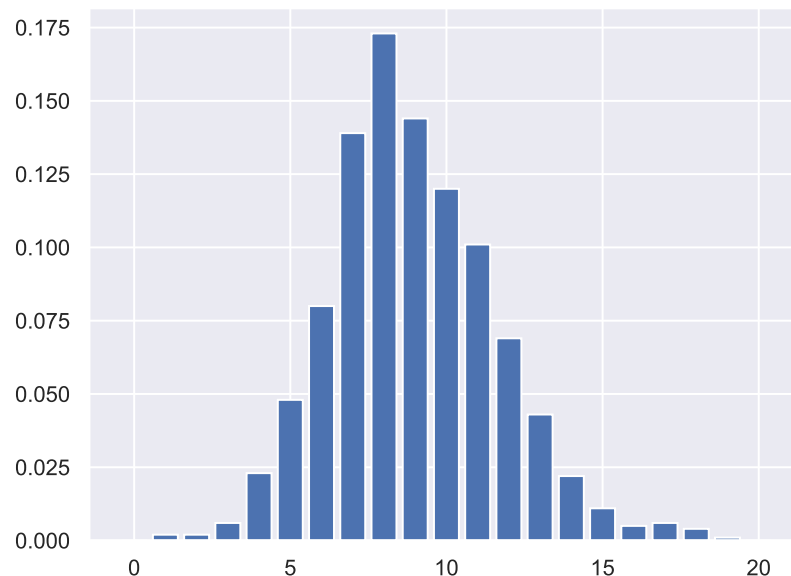
Le programme suivant calcule donc très efficacement un tirage de $\ell^{(N)} = \ell(\pi^{(N)})$ si $(\pi^{(n)})_{n \geq 1}$ est une chaîne de paramètres (θ, α) :

```

def length_partition(N, theta, alpha):
    res = 1
    for k in range(1, N):
        res += scs.bernoulli((theta + alpha * res)/(theta + k)).rvs()
    return res

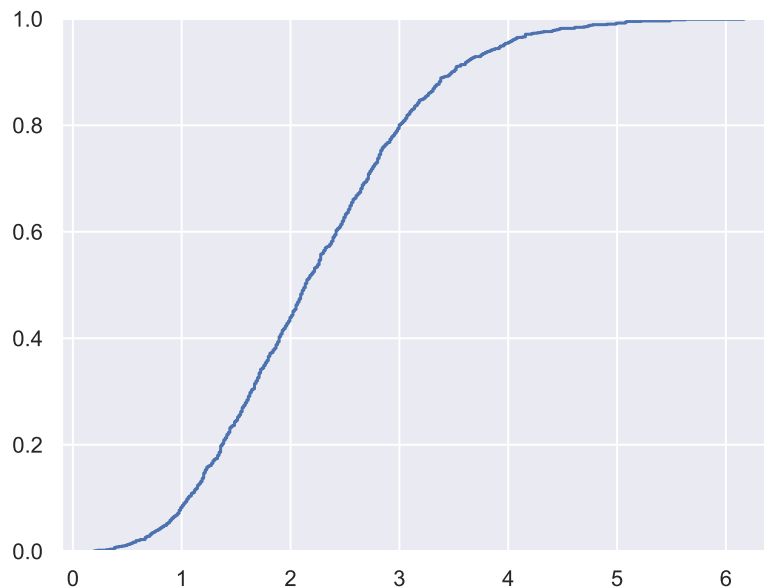
```

Avec $N = 5000$, $(\theta, \alpha) = (1, 0)$ et 1000 tirages, on obtient l'histogramme suivant pour la loi de $\ell^{(5000)}$:



Il n'y a pas à ce stade de concentration, mais on a bien un histogramme centré autour de la valeur $\log 5000 \simeq 8.517 \dots$. La raison de la non-concentration est la renormalisation en $\log n$: la convergence est assez lente (plus précisément, la question suivante montrera qu'on approche bien la loi de $\ell^{(n)}$ dans le cas $\alpha = 0$ par une loi de Poisson de paramètre $\theta \log n$, et dans notre cas ce paramètre n'est pas très grand, donc il n'y a pas de concentration visible).

Si l'on prend maintenant $(\theta, \alpha) = (1, 0.5)$ et qu'on dessine la fonction de répartition empirique de $\frac{\ell^{(n)}}{n^\alpha} = \frac{\ell^{(n)}}{\sqrt{n}}$ avec 1000 tirages, on obtient :



Ceci montre que la renormalisation en $\sqrt{n}$ pour $\alpha = 0.5$ semble être la bonne. Pour illustrer la convergence presque sûre de $\frac{\ell^{(n)}}{n^\alpha}$, on peut modifier le programme `length_partition` pour avoir toute la suite $(\ell^{(n)})_{n \in [1, N]}$, et tracer plusieurs trajectoires de $\frac{\ell^{(n)}}{\sqrt{n}}$:



Les courbes semblent bien tendre vers des limites (aléatoires), avec néanmoins une convergence assez lente (il y a encore des fluctuations importantes à $n = 5000$).

- (5) Si l'on reprend l'argument de la question précédente avec $\alpha = 0$, on obtient la représentation :

$$\ell^{(n+1)} = \ell^{(n)} + 1_{(\xi^{(n)} \geq \frac{n}{\theta+n})} = \ell^{(n)} + B_{n+1},$$

où les variables de Bernoulli B_k sont indépendantes, avec B_k de paramètre $p_k = \frac{\theta}{\theta+k}$. Ceci démontre la première partie de la question. Pour la seconde partie, notons que les variables $B_k^* = 1_{(P_k \geq 1)}$ ont presque pour paramètre $\frac{\theta}{\theta+k}$: en effet,

$$\mathbb{P}[B_k^* = 1] = 1 - e^{-\frac{\theta}{\theta+k}} \simeq_{k \rightarrow \infty} \frac{\theta}{\theta+k}.$$

Comme $\mathbb{P}[B_k^* = 1] \leq \mathbb{P}[B_k = 1]$, on peut réaliser les variables P_k et B_k de sorte que $1_{(P_k \geq 1)} = B_k^* \leq B_k$, et en gardant l'indépendance des P_k entre elles et des B_k entre elles. Alors :

- On a $P_k = B_k^*$ sauf si $P_k \geq 2$, et l'événement $E_k = \{P_k \neq B_k^*\}$ a pour probabilité

$$\mathbb{P}[E_k] = \mathbb{P}[P_k \geq 2] = 1 - e^{-p_k}(1 + p_k) \leq 1 - (1 - p_k)(1 + p_k) = (p_k)^2 \leq \frac{\theta}{k^2}.$$

La série $\sum_{k \geq 1} \mathbb{P}[E_k]$ est donc sommable, et par le lemme de Borel–Cantelli, avec probabilité 1, seulement un nombre fini d'événements E_k se réalise. Autrement dit, $P_k = B_k^*$ pour k assez grand, ce qui implique que la suite $(\sum_{k=1}^n P_k - B_k^*)_{n \geq 1}$ reste bornée.

- De même, on a $\mathbb{P}[B_k^* \neq B_k] = p_k - (1 - e^{-p_k}) = \sum_{n \geq 2} \frac{(-p_k)^n}{n!} \leq \frac{(p_k)^2}{2} \leq \frac{\theta}{2k^2}$, donc par le même argument que ci-dessus, $B_k = B_k^*$ pour k assez grand, et la suite $(\sum_{k=1}^n B_k - B_k^*)_{n \geq 1}$ reste bornée.

Le comportement asymptotique de la suite $(\ell^{(n)})_{n \geq 1}$ est donc le même que celui de $(\sum_{k=1}^n P_k)_{n \geq 1}$. Or, on peut réaliser la suite $(\sum_{k=1}^n P_k)_{n \geq 1}$ comme la suite des valeurs $(\mathcal{P}_{t_n})_{n \geq 1}$, où $(\mathcal{P}_t)_{t \geq 0}$ est un processus de Poisson standard, et $t_n = \sum_{k=1}^n p_k$. Par la loi des grands nombres pour les processus de Poisson,

$$\frac{\mathcal{P}_{t_n}}{t_n} \xrightarrow{n \rightarrow \infty, \text{ p.s. }} 1,$$

et il reste à voir que t_n est équivalent à $\theta \sum_{k=1}^n \frac{1}{k} \simeq \theta \log n$.

- (6) Si π et π' sont conjuguées par une permutation σ , écrivons les parts de π par ordre décroissant de taille :

$$\pi = (\{a_1, \dots, a_{\lambda_1}\}, \{a_{\lambda_1+1}, \dots, a_{\lambda_1+\lambda_2}\}, \dots, \{a_{\lambda_1+\dots+\lambda_{\ell-1}+1}, \dots, a_n\})$$

avec $\lambda(\pi) = (\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_\ell)$ signature de la partition π . Par hypothèse,

$$\pi' = (\{\sigma(a_1), \dots, \sigma(a_{\lambda_1})\}, \{\sigma(a_{\lambda_1+1}), \dots, \sigma(a_{\lambda_1+\lambda_2})\}, \dots, \{\sigma(a_{\lambda_1+\dots+\lambda_{\ell-1}+1}), \dots, \sigma(a_n)\}),$$

ce qui prouve que π' a la même signature : $\lambda(\pi) = \lambda(\pi')$. Réciproquement, si $\lambda(\pi) = \lambda(\pi')$, alors on peut écrire :

$$\pi' = (\{b_1, \dots, b_{\lambda_1}\}, \{b_{\lambda_1+1}, \dots, b_{\lambda_1+\lambda_2}\}, \dots, \{\sigma(b_{\lambda_1+\dots+\lambda_{\ell-1}+1}), \dots, \sigma(b_n)\})$$

avec la même suite décroissante $(\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_\ell)$. Alors, π et π' sont conjuguées par la permutation σ définie par $\sigma(a_i) = b_i$ pour tout $i \in [1, n]$.

- (7) La formule est vraie si $n = 1$: il n'y a qu'une partition $\{1\}$ de taille 1, et on a bien

$$\frac{\alpha^1}{\theta^{\uparrow 1}} \left(\frac{\theta}{\alpha} \right)^{\uparrow 1} (1 - \alpha)^{\uparrow 0} = \frac{\alpha}{\theta} \frac{\theta}{\alpha} = 1.$$

Supposons la formule établie au rang n , et considérons une partition II de taille $n + 1$. Puisque la chaîne de Markov de paramètres (θ, α) donne une suite de partitions cohérente, si π est la partition de taille n obtenue en supprimant $n + 1$ de la part correspondante de II , alors on a :

$$\begin{aligned}\mathbb{P}_{n+1}[II] &= \mathbb{P}[\pi^{(n+1)} = II] = \mathbb{P}[\pi^{(n)} = \pi \text{ et } \pi^{(n+1)} = II] \\ &= \mathbb{P}_n[\pi] P(\pi, II).\end{aligned}$$

Distinguons alors deux cas :

(a) Si $\{n + 1\}$ est une part de II , alors $II = \pi[\ell(\pi) + 1]$, et on a :

$$\begin{aligned}\mathbb{P}_{n+1}[II] &= \mathbb{P}_n[\pi] P(\pi, II) = \mathbb{P}_n[\pi] \frac{\theta + \ell(\pi) \alpha}{\theta + n} \\ &= \frac{\alpha^{\ell(\pi)}}{\theta^{\uparrow n}} \left(\frac{\theta}{\alpha}\right)^{\uparrow \ell(\pi)} \left(\prod_{k=1}^{\ell(\pi)} (1 - \alpha)^{\uparrow c_k(\pi) - 1}\right) \frac{\theta + \ell(\pi) \alpha}{\theta + n} \\ &= \frac{\alpha^{\ell(II)}}{\theta^{\uparrow n+1}} \left(\frac{\theta}{\alpha}\right)^{\uparrow \ell(II)} \left(\prod_{k=1}^{\ell(II)} (1 - \alpha)^{\uparrow c_k(II) - 1}\right) \frac{\theta + \ell(\pi) \alpha}{\alpha} \\ &= \frac{\alpha^{\ell(II)}}{\theta^{\uparrow n+1}} \left(\frac{\theta}{\alpha}\right)^{\uparrow \ell(II)} \left(\prod_{k=1}^{\ell(II)} (1 - \alpha)^{\uparrow c_k(II) - 1}\right)\end{aligned}$$

car $\ell(II) = \ell(\pi) + 1$, et $c_k(II) = c_k(\pi)$ pour tout $k \in [1, \ell(\pi)]$.

(b) Sinon, $n + 1$ appartient à une part II_k de taille plus grande que 2, et $II = \pi[k]$ avec $k \in [1, \ell(\pi)]$. On a dans ce cas :

$$\begin{aligned}\mathbb{P}_{n+1}[II] &= \mathbb{P}_n[\pi] P(\pi, II) = \mathbb{P}_n[\pi] \frac{c_k(\pi) - \alpha}{\theta + n} \\ &= \frac{\alpha^{\ell(\pi)}}{\theta^{\uparrow n}} \left(\frac{\theta}{\alpha}\right)^{\uparrow \ell(\pi)} \left(\prod_{j=1}^{\ell(\pi)} (1 - \alpha)^{\uparrow c_j(\pi) - 1}\right) \frac{c_k(\pi) - \alpha}{\theta + n} \\ &= \frac{\alpha^{\ell(II)}}{\theta^{\uparrow n+1}} \left(\frac{\theta}{\alpha}\right)^{\uparrow \ell(II)} \left(\prod_{j \neq k} (1 - \alpha)^{\uparrow c_j(II) - 1}\right) (1 - \alpha)^{\uparrow c_k(\pi) - 1} (c_k(\pi) - \alpha) \\ &= \frac{\alpha^{\ell(II)}}{\theta^{\uparrow n+1}} \left(\frac{\theta}{\alpha}\right)^{\uparrow \ell(II)} \left(\prod_{j=1}^{\ell(II)} (1 - \alpha)^{\uparrow c_j(II) - 1}\right)\end{aligned}$$

car $\ell(II) = \ell(\pi)$, $c_j(II) = c_j(\pi)$ pour tout $j \neq k$, et $c_k(II) = c_k(\pi) + 1$.

La formule est donc vraie au rang $n + 1$ dans les deux cas.

(8) On souhaite montrer que la loi conditionnelle de $\pi^{(n)}$ sachant la tribu $\mathcal{F}_{-n}$ est la loi uniforme sur l'ensemble des partitions de signature $\lambda^{(n)}$. Par définition de l'espérance conditionnelle, il s'agit donc de montrer que pour tout événement $A = A(\mu^{(n)}, \dots, \mu^{(n+k)})$,

$$\mathbb{E}[1_{(\pi^{(n)}=\pi)} 1_A] = \frac{1}{B_{n,A}} \mathbb{E}[1_{(\lambda^{(n)}=A)} 1_A],$$

avec $A = \lambda(\pi)$. Autrement dit, la quantité $\mathbb{P}[\pi^{(n)} = \pi \text{ et } A]$ ne dépend que de la signature $\lambda(\pi)$.

• L'identité ci-dessus est évidente si $A \neq \mu^{(n)}$ (dans ce cas, les deux parties valent 0).

- Supposons maintenant $\Lambda = \mu^{(n)}$, et fixons deux partitions π et π' avec la même signature, et une permutation $\sigma \in \mathfrak{S}(n)$ telle que $\pi' = \sigma(\pi)$. Remarquons que σ agit aussi sur $[1, n+k]$: $\sigma(x) = x$ si $x > n$. Soit L l'ensemble des partitions Π de taille $n+k$ telle que, si $(\Pi^{(m)})_{m \in [1, n+k]}$ est l'unique suite cohérente avec $\Pi^{(n+k)} = \Pi$, alors $\lambda(\Pi^{(m)}) = \mu^{(m)}$ pour tout $m \in [n, n+k]$. L'événement A est simplement $\{\pi^{(n+k)} \in L\}$. Or, si $\Pi \in L$, alors on a aussi $\sigma^{-1}(\Pi) \in L$: la permutation σ change dans quelles parts tombent les éléments $1, 2, \dots, n$, mais elle ne change pas la taille de ces parts, y compris après intersection avec $[1, n+l]$ pour $l \in [0, k]$. Autrement dit, $\sigma(L) = L$. Par échangeabilité de la suite de partitions, on a donc :

$$\begin{aligned} \mathbb{P}[\pi^{(n)} = \pi \text{ et } A] &= \mathbb{P}[\pi^{(n)} = \pi \text{ et } \pi^{(n+k)} \in L] = \mathbb{P}[\pi^{(n)} = \sigma(\pi) \text{ et } \pi^{(n+k)} \in \sigma(L)] \\ &= \mathbb{P}[\pi^{(n)} = \pi' \text{ et } \pi^{(n+k)} \in L] = \mathbb{P}[\pi^{(n)} = \pi' \text{ et } A]. \end{aligned}$$

C'est ce que l'on souhaitait démontrer.

Calculons maintenant $B_{n,\Lambda}$ si $\Lambda = (s, \dots, s, s-1, \dots, s-1, \dots, 2, \dots, 2, 1, \dots, 1)$ avec m_i entrées égales à i pour chaque $i \in [1, s]$. Pour toute permutation $\sigma \in \mathfrak{S}(n)$, on peut construire une partition de signature Λ comme suit :

$$\begin{aligned} &\{\sigma(1), \dots, \sigma(s)\}, \{\sigma(s+1), \dots, \sigma(2s)\}, \dots, \{\sigma((m_s-1)s+1), \dots, \sigma(m_s s)\}, \\ &\{\sigma(m_s s+1), \dots, \sigma(m_s s+s-1)\}, \dots \end{aligned}$$

Autrement dit, on remplit avec les valeurs de la permutation les parts de taille s , puis celles de taille $s-1$, etc. jusqu'aux parts de taille 1. Toute partition de signature Λ est obtenue de cette façon. De plus, il y a $\prod_{i=1}^s (i!)^{m_i} (m_i!)$ permutations qui donnent la même partition, car on peut sans changer la partition :

- permuter dans chaque part les valeurs de la permutation (d'où le facteur $\prod_{i=1}^s (i!)^{m_i}$);
- permuter les parts de même taille dans l'écriture ci-dessus (d'où le facteur $\prod_{i=1}^s (m_i!)$).

La formule pour $B_{n,\Lambda}$ s'en déduit par le lemme des bergers.

Calculons maintenant le noyau $Q(\Lambda, \mu) = \mathbb{P}[\lambda(E(\Pi)) = \mu]$. Si Π est de signature $\Lambda = (s^{m_s}, (s-1)^{m_{s-1}}, \dots, 2^{m_2}, 1^{m_1})$, alors l'effacement d'un élément de $[1, n]$ dans une part de taille r donne une partition $\pi = E(\Pi)$ dont la signature μ s'écrit :

$$\mu = (s^{m_s}, \dots, r^{m_r-1}, (r-1)^{m_{r-1}+1}, \dots, 1^{m_1}).$$

Si μ a cette forme, alors le noyau de transition s'écrit $Q(\Lambda, \mu) = \frac{r m_r}{n}$, et on vérifie aisément que ce ratio est aussi égal à

$$\frac{(m_{r-1} + 1) B_{n-1, \mu}}{B_{n, \Lambda}}.$$

Voyons finalement comment calculer la loi conditionnelle de $\lambda^{(n-1)}$ sachant $\mathcal{F}_{-n}$. Par définition, $\pi^{(n-1)}$ est obtenue à partir de $\pi^{(n)}$ en effaçant n , donc :

$$\mathbb{P}[\pi^{(n-1)} = \pi \mid \mathcal{F}_{-n}] = \sum_{k \in [1, \ell(\pi)+1]} \mathbb{P}[\pi^{(n)} = \pi[k] \mid \mathcal{F}_{-n}] = \sum_{k \in [1, \ell(\pi)+1]} \frac{1_{(\lambda(\pi[k])=\lambda^{(n)})}}{B_{n, \lambda^{(n)}}}.$$

Si π et π' partitions de taille $n-1$ ont même signature, alors il est facile de voir que les signatures de leurs successeurs forment le même multi-ensemble. Donc, si μ est une signature de taille $n-1$ et π est une partition arbitraire de signature μ , alors

$$\mathbb{P}[\lambda^{(n-1)} = \mu \mid \mathcal{F}_{-n}] = \frac{B_{n-1, \mu}}{B_{n, \lambda^{(n)}}} \sum_{k \in [1, \ell(\pi)+1]} 1_{(\lambda(\pi[k])=\lambda^{(n)})}.$$

Si μ n'est pas obtenue à partir de $\Lambda = \lambda^{(n)}$ en diminuant une part d'une unité, alors la somme à droite vaut 0. Sinon, μ est obtenue à partir de Λ en changeant une part de taille r en une part de taille $r - 1$, et il y a $m_{r-1} + 1$ indices k tels que $\lambda(\pi[k]) = \Lambda$ (le nombre de parts de μ égales à $r - 1$). On a donc montré que :

$$\mathbb{P}[\lambda^{(n-1)} = \mu \mid \mathcal{F}_{-n}] = Q(\Lambda, \mu) = Q(\lambda^{(n)}, \mu).$$

- (9) Pour la première partie de la question, remarquons que si A est une partie aléatoire de $[1, n + 1]$ qui est $\mathcal{G}$ -mesurable et X_{n+1} est une variable uniforme sur $[1, n + 1]$ qui est indépendante de $\mathcal{G}$, alors $\mathbb{P}[X_{n+1} \in A \mid \mathcal{G}] = \frac{|A|}{n+1}$. En effet,

$$\begin{aligned} \mathbb{P}[X_{n+1} \in A \mid \mathcal{G}] &= \sum_{k=1}^{n+1} \mathbb{E}[1_{(k \in A)} 1_{(X_{n+1}=k)} \mid \mathcal{G}] \\ &= \sum_{k=1}^{n+1} 1_{(k \in A)} \mathbb{E}[1_{(X_{n+1}=k)}] = \frac{1}{n+1} \sum_{k=1}^{n+1} 1_{(k \in A)} = \frac{|A|}{n+1}. \end{aligned}$$

Par conséquent, avec $\mathcal{G}$ comme dans le texte, puisque $A = \pi_{\rho(1)}^{(n+1)} \sqcup \dots \sqcup \pi_{\rho(k)}^{(n+1)}$ est $\mathcal{G}$ -mesurable et X_{n+1} est indépendante de $\mathcal{G}$,

$$\mathbb{P}[X_{n+1} \in A \mid \mathcal{G}] = \frac{|\pi_{\rho(1)}^{(n+1)} \sqcup \dots \sqcup \pi_{\rho(k)}^{(n+1)}|}{n+1} = V_{-(n+1),k}.$$

En prenant l'espérance conditionnelle sachant $\mathcal{F}_{-(n+1)} \subset \mathcal{G}$, on obtient :

$$\mathbb{E}[1_{(X_{n+1} \in A)} \mid \mathcal{F}_{-(n+1)}] = V_{-(n+1),k},$$

et finalement $\mathbb{E}[|A| - 1_{(X_{n+1} \in A)} \mid \mathcal{F}_{-(n+1)}] = (n+1) V_{-(n+1),k} - V_{-(n+1),k} = n V_{-(n+1),k}$.

Pour la seconde partie, il s'agit de simplifier la quantité

$$\mathbb{E}[\lambda_1(E(\pi^{(n+1)})) + \dots + \lambda_k(E(\pi^{(n+1)})) \mid \mathcal{F}_{-(n+1)}].$$

Or, par le lemme établi dans la question précédente,

$$\mathbb{P}[\lambda^{(n)} = \mu \mid \mathcal{F}_{-(n+1)}] = Q(\lambda^{(n+1)}, \mu) = \mathbb{P}[\lambda(E(\pi^{(n+1)})) = \mu \mid \mathcal{F}_{-(n+1)}].$$

L'identité s'obtient donc en décomposant en fonction des valeurs possibles pour la signature $\lambda(E(\pi^{(n+1)}))$:

$$\begin{aligned} &\mathbb{E}[\lambda_1(E(\pi^{(n+1)})) + \dots + \lambda_k(E(\pi^{(n+1)})) \mid \mathcal{F}_{-(n+1)}] \\ &= \sum_{\mu \in \mathfrak{Y}(n)} \mathbb{P}[\lambda(E(\pi^{(n+1)})) = \mu \mid \mathcal{F}_{-(n+1)}] (\mu_1 + \dots + \mu_k) \\ &= \sum_{\mu \in \mathfrak{Y}(n)} \mathbb{P}[\lambda^{(n)} = \mu \mid \mathcal{F}_{-(n+1)}] (\mu_1 + \dots + \mu_k) \\ &= \mathbb{E}[\lambda_1^{(n)} + \dots + \lambda_k^{(n)} \mid \mathcal{F}_{-(n+1)}] = \mathbb{E}[n V_{-n,k} \mid \mathcal{F}_{-(n+1)}]. \end{aligned}$$

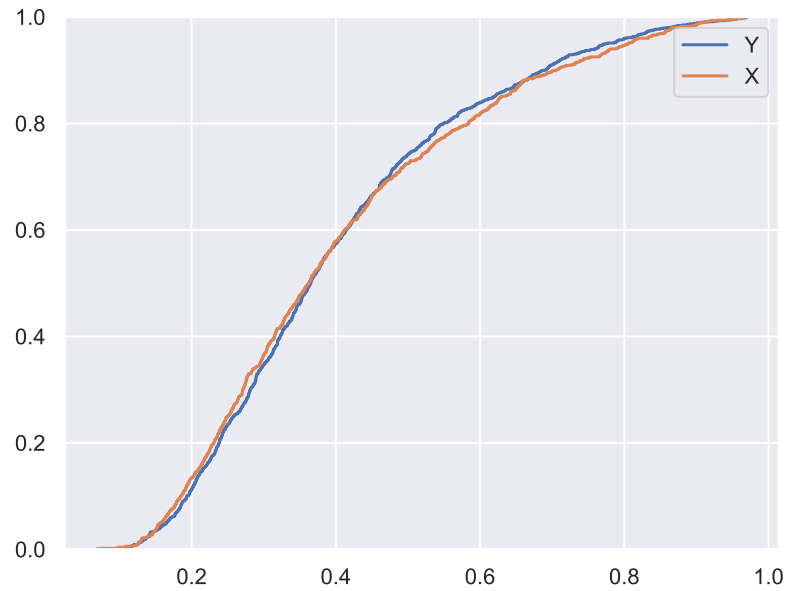
- (10) On obtient un échantillon de la variable Y_1 en utilisant le programme :

```
def Y1(N, theta, alpha):
    return max([len(part) for part in markov_partition(N, theta, alpha)]) / N
```

avec N grand, par exemple $N = 5000$. Par ailleurs, on obtient un échantillon de $\max(X_{k \geq 1})$ avec le programme :

```
def maximum_X(theta, alpha, K=100):  
    W = np.array([scs.beta(1 - alpha, theta + (k+1) * alpha).rvs() for k in range(K)])  
    X = [np.prod(1 - W[:k]) * W[k] for k in range(K)]  
    return max(X)
```

Avec $\theta = 1$, $\alpha = 0.5$ et des échantillons de taille 1000, on obtient les fonctions de répartition empiriques suivantes :



La correspondance des lois est donc confirmée par ces programmes.