

## Corrigé.

Les programmes des questions 1., 4., 5., 6. et 8. sont dans la feuille de calcul Jupyter.

2. À chaque temps  $n$ , on rajoute un nouveau groupe à  $c_{n-1}$  pour obtenir  $c_n$  avec probabilité  $\frac{1}{n}$ . On a donc l'identité en loi :

$$\ell_n = \ell_{n-1} + \text{Ber}\left(\frac{1}{n}\right),$$

avec une variable de Bernoulli indépendante des choix précédents. Ceci implique l'identité annoncée.

3. En prenant les espérances, on obtient :

$$\mathbb{E}[\ell_n] = \sum_{k=1}^n \frac{1}{k},$$

et il est bien connu que la somme harmonique à droite est équivalente à  $\log n$ . Ceci prouve que  $\frac{\mathbb{E}[\ell_n]}{\log n}$  tend vers 1. Pour les variances, comme les variables de Bernoulli sont indépendantes, on a :

$$\text{var}(\ell_n) = \sum_{k=1}^n \text{var}\left(\text{Ber}\left(\frac{1}{k}\right)\right) = \sum_{k=1}^n \frac{1}{k} \left(1 - \frac{1}{k}\right) = \mathbb{E}[\ell_n] - \sum_{k=1}^n \frac{1}{k^2}.$$

La somme des inverses des carrés est bornée par  $\sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6}$ , donc on a de nouveau  $\frac{\text{var}(\ell_n)}{\log n} \rightarrow 1$ .

5. La courbe  $k \in \llbracket 1, n \rrbracket \mapsto \frac{\ell_k}{\log(k+1)}$  semble toujours converger vers la valeur 1 : on peut donc conjecturer la convergence presque sûre de  $\frac{\ell_n}{\log n}$  vers la constante 1.
7. Les deux histogrammes sont très proches, et on peut donc conjecturer la convergence en loi de  $P_n$  vers un variable de Poisson de paramètre 1.
9. Les fonctions de répartition empiriques sont toutes proches de la fonction  $F(x) = x$ , qui est la fonction de répartition empirique de la loi uniforme sur  $[0, 1]$ . On peut donc conjecturer que la variable  $U_n$  converge en loi vers une variable  $U$  de loi uniforme sur  $[0, 1]$ .
10. La fonction  $n \mapsto c_{n,1}$  est croissante par pas 0 ou +1 ; si elle atteint le niveau  $k$  au temps  $n$ , alors elle est entièrement déterminée par les entiers  $a_2, \dots, a_k$  auxquels  $c_{m,1}$  augmente d'une unité. Si ces entiers sont donnés, alors la probabilité d'avoir cette suite est :

$$\begin{aligned} & \frac{2-1}{2} \frac{3-1}{3} \dots \frac{a_2-1-1}{a_2-1} \times \frac{1}{a_2} \\ & \times \frac{a_2+1-2}{a_2+1} \frac{a_2+2-2}{a_2+2} \dots \frac{a_3-1-2}{a_3-1} \times \frac{2}{a_3} \\ & \quad \vdots \quad \quad \quad \vdots \\ & \times \frac{a_{k-1}+1-(k-1)}{a_{k-1}+1} \dots \frac{a_k-1-(k-1)}{a_k-1} \times \frac{k-1}{a_k} \\ & \times \frac{a_k+1-k}{a_k+1} \dots \frac{n-k}{n} \end{aligned}$$

en utilisant les formules donnant les probabilités de transition de la suite  $c_n$ . On remarque que le produit des numérateurs est toujours  $(n-k)!(k-1)!$ , et le produit des dénominateurs est  $n!$ . Ainsi, la probabilité d'obtenir  $c_{n,1} = k$  et une suite donnée  $a_2 < a_3 < \dots < a_k$  est toujours  $\frac{(n-k)!(k-1)!}{n!}$ .

11. On en déduit, puisqu'il y a  $\binom{n-1}{k-1}$  possibilités pour la suite  $a_2 < \dots < a_k$ , que :

$$\mathbb{P}[c_{n,1} = k] = \frac{(n-k)!(k-1)!}{n!} \frac{(n-1)!}{(n-k)!(k-1)!} = \frac{1}{n}.$$

Autrement dit,  $c_{n,1}$  est une variable aléatoire entière uniformément répartie sur  $\llbracket 1, n \rrbracket$ . On a donc pour tout  $x \in [0, 1]$ :

$$\mathbb{P}\left[\frac{c_{n,1}}{n} \leq x\right] = \frac{\lfloor nx \rfloor}{n} \rightarrow_{n \rightarrow \infty} x.$$

La convergence des fonctions de répartition vers la fonction continue  $F(x) = x$  implique la convergence en loi  $U_n \xrightarrow{n \rightarrow \infty} U$ , comme conjecturé dans la question 9.

12. Pour montrer l'identité des lois, on utilise une construction par récurrence des permutations aléatoires uniformes. Si  $\sigma \in \mathfrak{S}(n)$ , considérons tous les produits

$$\tau_{\sigma,k} = \sigma \circ (k, n+1)$$

avec  $k \in \llbracket 1, n+1 \rrbracket$ , et par convention  $(n+1, n+1) = \text{id}_{\llbracket 1, n+1 \rrbracket}$ .

- Toutes ces permutations sont différentes, car  $\tau_{\sigma,k}(n+1) = \sigma(k)$  pour  $k \neq n+1$ , et  $\tau_{\sigma,n+1}(n+1) = n+1$ . Or,  $n+1$  et les  $n$  valeurs  $\sigma(1), \dots, \sigma(n)$  sont toutes différentes.
- On peut reconstituer  $\sigma$  à partir de n'importe quel  $\tau_{\sigma,k}$ : en effet, le  $k$  de  $\tau_{\sigma,k}$  est  $(\tau_{\sigma,k})^{-1}(n+1)$ , et on récupère ensuite  $\sigma$  en considérant  $\tau_{\sigma,k} \circ (k, n+1)$ .

Ces deux propriétés impliquent que les parties  $A_\sigma = \{\tau_{\sigma,k}, k \in \llbracket 1, n+1 \rrbracket\}$  de  $\mathfrak{S}(n+1)$  sont toutes disjointes, et toutes de cardinal  $n+1$ . Par cardinalité, on a donc :

$$\mathfrak{S}(n+1) = \bigsqcup_{\sigma \in \mathfrak{S}(n)} A_\sigma.$$

Alors, si  $\sigma$  est choisie uniformément dans  $\mathfrak{S}(n)$ , et  $k$  est choisi uniformément dans  $\llbracket 1, n+1 \rrbracket$ , alors  $\tau = \tau_{\sigma,k}$  est de loi uniforme dans  $\mathfrak{S}(n+1)$  : en effet, l'identité ensembliste ci-dessus montre que toute permutation  $\tau \in \mathfrak{S}(n+1)$  est obtenue, et toutes ces permutations ont la même probabilité  $\frac{1}{(n+1)!}$ .

Voyons alors comment les cycles évoluent par multiplication par  $(k, n+1)$  avec  $k \sim \text{Unif}(\llbracket 1, n+1 \rrbracket)$ . Si  $\sigma = C_1 \circ C_2 \circ \dots \circ C_\ell$ , alors :

- avec probabilité  $\frac{1}{n+1}$ ,  $k = n+1$  et  $\tau_{\sigma,k}$  a un nouveau cycle  $C_{\ell+1} = (n+1)$ .
- pour tout  $j \in \llbracket 1, \ell \rrbracket$ , avec probabilité  $\frac{|C_j|}{n+1}$ ,  $k \in C_j$ , et

$$C_j \circ (k, n+1) = (a, \dots, z, k) \circ (k, n+1) = (a, \dots, z, k, n+1).$$

Autrement dit,  $\tau_{\sigma,k}$  a le même nombre de cycles que  $\sigma$ , et la taille de  $C_j$  augmente d'une unité.

Conclusion : on peut construire par récurrence une permutation  $\sigma_n$  de loi uniforme sur  $\mathfrak{S}(n)$ , de sorte que les tailles des cycles évoluent exactement comme défini au début du problème. Ceci implique l'identité en loi souhaitée.