



PRÉPASIMPLE

prepasimple.fr

ESSEC

Mathématiques 2 — Option économique

2011

Mathématiques appliquées

Mathématiques appliquées — Prépa ECG

Annales de concours · Corrigé

Partie 1 — Description et premiers résultats

Le nombre N est fixé, sauf dans les questions étudiant sa limite. Les configurations appartiennent à S_N , mais les issues de l'expérience doivent décrire une suite d'insertions. L'espace probabilisé n'est donc pas identifié à S_N muni de l'équiprobabilité. La loi uniforme sur les configurations sera notée π en partie 3.

1) En additionnant les différences successives, les termes intermédiaires s'éliminent :

$$\Delta_1 + \dots + \Delta_i = T_1 + (T_2 - T_1) + \dots + (T_i - T_{i-1}) = T_i.$$

Ainsi, Δ_i est le nombre d'insertions supplémentaires nécessaires pour faire remonter C_N de la position $N - i + 1$ à la position $N - i$. Les deux instants sont inclus selon la convention du nombre d'insertions effectuées entre eux : $\Delta_i \geq 1$.

2) Tant que C_N est au bas du paquet, elle remonte exactement lorsqu'une insertion est faite en dernière place. À chaque essai, la probabilité de cet événement est $1/N$, indépendamment des essais précédents. Pour $n \geq 0$, ne pas avoir encore réussi après n insertions signifie que les n essais ont échoué. Donc

$$P(\Delta_1 > n) = (1 - 1/N)^n.$$

Pour $n = 0$, la probabilité vaut bien 1. Pour tout entier $k \geq 1$,

$$P(\Delta_1 = k) = P(\Delta_1 > k - 1) - P(\Delta_1 > k) = \frac{1}{N}(1 - 1/N)^{k-1}.$$

Par conséquent, Δ_1 suit la loi géométrique sur $\mathbb{N}^*$ de paramètre $1/N$. La probabilité d'attendre indéfiniment est nulle, car $(1 - 1/N)^n \rightarrow 0$.

3)(a) À l'instant T_{i-1} , C_N est en position $N - i + 1$. Lorsqu'on retire la carte du dessus, elle passe provisoirement en position $N - i$. Si la carte retirée est réinsérée avant elle, elle reprend sa position précédente ; si la carte est réinsérée après elle, elle reste à la nouvelle position.

Il y a exactement i places d'insertion après C_N , parmi les N places possibles. Tant qu'aucune de ces places n'est choisie, sa position reste $N - i + 1$. Le choix à chaque insertion est indépendant du passé. La durée jusqu'au premier succès a donc les probabilités

$$P(\Delta_i > n) = (1 - i/N)^n = ((N - i)/N)^n,$$

pour tout $n \geq 0$. Ainsi, Δ_i suit la loi géométrique de paramètre i/N sur $\mathbb{N}^*$. Le paramètre appartient à $]0, 1[$, puisque $1 \leq i \leq N - 1$.

3)(b) Une variable géométrique de paramètre $p \in]0, 1]$ possède un moment d'ordre deux, une espérance $1/p$ et une variance $(1 - p)/p^2$. Ces résultats s'appliquent avec $p = i/N$, d'où

$$E(\Delta_i) = \frac{N}{i}, \quad V(\Delta_i) = \frac{1 - i/N}{(i/N)^2} = \frac{N(N - i)}{i^2}.$$

Les mêmes formules valent pour $i = 1$, d'après 2).

4)(a) Question adaptée au programme actuel : $N \geq 3$ est précisé, car T_2 n'est défini que dans ce cas. Par 1), $T_2 = \Delta_1 + \Delta_2$. Les deux durées sont à valeurs dans $\mathbb{N}^*$, donc, pour $n \geq 2$,

$$[T_2 = n] = \bigcup_{k=1}^{n-1} ([\Delta_1 = k] \cap [\Delta_2 = n - k]).$$

Les événements de cette union sont disjoints. L'additivité et l'indépendance admise des durées donnent

$$P(T_2 = n) = \sum_{k=1}^{n-1} P(\Delta_1 = k)P(\Delta_2 = n - k).$$

Pour $n < 2$, la probabilité est nulle. Aucune formule générale de convolution n'est nécessaire : il s'agit d'une partition finie de l'événement étudié.

4)(b) Posons $q_1 = 1 - 1/N$, $q_2 = 1 - 2/N$ et $r = q_1/q_2$. Puisque $N \geq 3$, $q_2 > 0$ et $r > 1$,

avec $r - 1 = 1/(N - 2)$. La somme géométrique finie donne

$$\sum_{k=1}^{n-1} r^k = \frac{r(r^{n-1} - 1)}{r - 1} = (N - 1)(r^{n-1} - 1).$$

Comme $N - 1 = N(1 - 1/N)$, c'est l'identité demandée. Le cas $n = 2$ est également inclus.

4)(c) Les lois géométriques obtenues en 2) et 3) fournissent

$$P(T_2 = n) = \frac{2}{N^2} \sum_{k=1}^{n-1} q_1^{k-1} q_2^{n-k-1} = \frac{2}{N^2} \frac{q_2^{n-1}}{q_1} \sum_{k=1}^{n-1} r^k.$$

En remplaçant la somme par le résultat de 4)(b), on obtient

$$P(T_2 = n) = \frac{2}{N} q_2^{n-1} (r^{n-1} - 1) = \frac{2}{N} (q_1^{n-1} - q_2^{n-1}).$$

Cela donne exactement $\frac{2}{N} [(1 - 1/N)^{n-1} - (1 - 2/N)^{n-1}]$ pour $n \geq 2$.

5)(a) Entre T_1 et la première insertion faisant à nouveau remonter C_N , la carte placée sous elle à T_1 reste sous elle. À l'instant de la deuxième remontée, la place choisie est l'une des deux places situées après C_N . Conditionnellement à cette remontée, ces deux places sont équiprobables, car les choix initiaux des N places sont uniformes.

Si la nouvelle carte est placée en dernière position, l'ancienne reste en position $N - 1$. La probabilité demandée est donc $1/2$. Les insertions qui ne font pas remonter C_N ne changent pas l'ordre des cartes déjà situées sous elle.

5)(b) L'autre place possible met la nouvelle carte juste sous C_N et repousse l'ancienne en position N . Elle a elle aussi probabilité $1/2$. Les deux événements de 5)(a) et 5)(b) sont complémentaires.

6)(a) Question adaptée au programme actuel : on précise $N \geq 4$, afin que T_3 existe. Les trois cartes occupent les trois places $N - 2, N - 1, N$. Il y a trois choix pour a_1 , puis deux pour a_2 et une seule place restante pour a_3 . Il y a donc $3 \times 2 \times 1 = 6$ triplets possibles. Ce comptage se fait directement, sans supposer un chapitre général de dénombrement.

6)(b)(i) Pour obtenir l'ordre de haut en bas « carte de T_1 , carte de T_2 , carte de T_3 », les deux premières cartes doivent être dans cet ordre à T_2 , ce qui a probabilité $1/2$. À la troisième remontée, la nouvelle carte doit être insérée après les deux autres, soit dans une des trois places possibles sous C_N .

Cette place a probabilité conditionnelle $1/3$, indépendamment de l'ordre préexistant. Les probabilités composées donnent $(1/2)(1/3) = 1/6$.

6)(b)(ii) Le même ordre des deux premières cartes à T_2 est requis, de probabilité $1/2$. La troisième carte doit cette fois être insérée entre elles, ce qui correspond à une des trois places équiprobables. La probabilité est donc encore $1/6$.

7) Question adaptée au programme actuel : les indications ajoutées permettent de justifier la propriété conditionnelle qui sera nécessaire en 14)(a), et pas seulement l'uniformité à l'instant aléatoire T .

Appelons A_i la carte passée sous C_N à l'instant T_i . Ces cartes sont distinctes : une carte déjà sous C_N ne peut revenir au sommet avant que C_N ait elle-même atteint le sommet. Avant l'instant T_i , il y a $i - 1$ cartes sous C_N . Au premier choix faisant remonter C_N , les i places d'insertion sous elle sont équiprobables.

Montrons que, conditionnellement aux durées d'attente et à l'évolution du paquet au-dessus de C_N , les i cartes sous elle sont uniformément ordonnées. Pour $i = 1$, il n'y a qu'un ordre. Supposons la propriété vraie pour $i - 1$. Les insertions précédant la prochaine remontée ont lieu au-dessus de C_N : elles ne modifient pas l'ordre des cartes sous elle. Elles déterminent l'identité A_i de la carte à ajouter, sans utiliser l'ordre des cartes sous C_N .

À l'instant de la remontée, chacune des i places sous C_N a probabilité conditionnelle $1/i$, quel que soit cet ordre. Pour obtenir un ordre final prescrit des i cartes, l'ordre des $i - 1$ anciennes cartes est imposé, ainsi que la place de la nouvelle. Sa probabilité est donc

$$\frac{1}{(i-1)!} \times \frac{1}{i} = \frac{1}{i!}.$$

Les durées fixées n'altèrent pas cet argument : elles précisent combien de choix au-dessus de C_N ont précédé le choix réussi ; conditionnellement à un succès, chacune des places sous C_N reste équiprobable. L'induction est établie.

À T_{N-1} , C_N est au sommet et les $N - 1$ autres cartes sont uniformément ordonnées sous elle, même conditionnellement à la valeur de T_{N-1} . L'insertion suivante, à l'instant T , place C_N dans l'une des N positions équiprobables, indépendamment du passé. Tout ordre des N cartes est obtenu avec probabilité $1/((N - 1)!N) = 1/N!$. En particulier, pour chaque t tel que $P(T = t) > 0$, la configuration conditionnellement à $T = t$ est uniforme.

Vérifions enfin que l'uniformité est préservée par une insertion indépendante. Fixons une configuration finale σ et une place d'insertion k . La configuration précédente est entièrement déterminée : il faut retirer de σ la carte située en position k et la replacer au sommet. Sous une loi uniforme préalable et un choix uniforme indépendant de k , la probabilité du couple «cette configuration précédente, cette place» vaut $1/(N!N)$. En sommant sur les N places, la probabilité finale de σ vaut $1/N!$.

Ainsi, pour tout $t \leq n$ de probabilité positive, la configuration à l'instant n , conditionnellement à $T = t$, est uniforme. Arrêter exactement à T fournit un mélange uniforme ; poursuivre ensuite avec des insertions indépendantes conserve cette uniformité dans le sens conditionnel précisé.

Partie 2 — Estimation du nombre d'insertions pour bien mélanger les cartes

- 8) Les variables Δ_i ont toutes un moment d'ordre deux, et elles sont en nombre fini. Par linéarité de l'espérance,

$$E(T) = 1 + \sum_{i=1}^{N-1} \frac{N}{i} = 1 + NH_{N-1} = NH_N.$$

L'indépendance des durées permet d'additionner les variances ; ajouter la constante 1 ne change pas la variance. Donc

$$V(T) = \sum_{i=1}^{N-1} \frac{N(N-i)}{i^2} = N^2 \sum_{i=1}^{N-1} \frac{1}{i^2} - N \sum_{i=1}^{N-1} \frac{1}{i}.$$

En ajoutant dans chaque somme le terme d'indice N , la différence ajoutée est $N^2/N^2 - N/N = 0$. On obtient ainsi

$$V(T) = N^2 \sum_{i=1}^N \frac{1}{i^2} - NH_N.$$

9)(a) Sur $[k, k+1]$, la fonction $t \mapsto 1/t$ est continue et décroissante. On a donc $1/(k+1) \leq 1/t \leq 1/k$. En intégrant sur un intervalle de longueur 1, il vient

$$\frac{1}{k+1} \leq \int_k^{k+1} \frac{dt}{t} \leq \frac{1}{k}.$$

Les inégalités intégrées sont même strictes, car les écarts sont continus, positifs à l'intérieur du segment, et non identiquement nuls.

9)(b)(i) Pour $n \geq 1$,

$$u_{n+1} - u_n = \frac{1}{n+1} - \ln(n+1) + \ln n = \frac{1}{n+1} - \int_n^{n+1} \frac{dt}{t} \leq 0.$$

La suite est donc décroissante ; l'observation finale de 9)(a) montre même qu'elle est strictement décroissante.

9)(b)(ii) Question adaptée au programme actuel : la borne supérieure stricte du PDF est rectifiée à $n = 1$. Pour tout $n \geq 1$, en sommant les bornes strictes $\int_k^{k+1} dt/t < 1/k$ pour $k = 1, \dots, n$, on obtient

$$\ln(n+1) = \int_1^{n+1} \frac{dt}{t} < \sum_{k=1}^n \frac{1}{k} = H_n.$$

Pour $n \geq 2$, les inégalités $1/k < \int_{k-1}^k dt/t$ pour $k = 2, \dots, n$ donnent

$$H_n - 1 < \int_1^n \frac{dt}{t} = \ln n.$$

Ainsi, $H_n < \ln n + 1$ pour $n \geq 2$. Pour $n = 1$, $H_1 = 1 = \ln 1 + 1$. Le bon encadrement pour tout $n \geq 1$ est donc $\ln(n+1) < H_n \leq \ln n + 1$.

9)(c) Puisque $\ln(n+1) > \ln n$, l'encadrement précédent entraîne $0 < u_n \leq 1$. La suite est décroissante et minorée; elle converge donc vers un réel $\gamma \in [0, 1]$.

10)(a) Le résultat $u_N \rightarrow \gamma$ s'écrit $H_N = \ln N + \gamma + \varepsilon_N$, avec $\varepsilon_N \rightarrow 0$. En multipliant par N ,

$$E(T) = N \ln N + N\gamma + N\varepsilon_N = N \ln N + N\gamma + o(N).$$

De plus, pour $N \geq 2$,

$$\frac{E(T)}{N \ln N} = 1 + \frac{\gamma + \varepsilon_N}{\ln N} \rightarrow 1.$$

On a donc aussi $E(T) \sim N \ln N$.

10)(b) Posons

$$a_N = V(T)/N^2 = \sum_{k=1}^N 1/k^2 - H_N/N.$$

Une soustraction explicite, en utilisant $H_{N+1} = H_N + 1/(N+1)$, donne

$$a_{N+1} - a_N = \frac{1}{(N+1)^2} - \frac{H_N + 1/(N+1)}{N+1} + \frac{H_N}{N} = \frac{H_N}{N(N+1)} > 0.$$

La suite est strictement croissante. Pour justifier sa convergence sans invoquer un résultat supplémentaire sur les séries, on peut noter que, pour $k \geq 2$,

$$\frac{1}{k^2} \leq \frac{1}{k(k-1)} = \frac{1}{k-1} - \frac{1}{k}.$$

Les sommes partielles positives de

$$\sum 1/k^2$$

sont donc croissantes et majorées par 2 : elles convergent vers un réel

$$\alpha = \sum_{k=1}^{+\infty} 1/k^2 \geq 1 > 0.$$

Par 9)(b), $0 \leq H_N/N \leq (\ln N + 1)/N \rightarrow 0$. On en déduit $a_N \rightarrow \alpha$.

Puisque a_N est croissante vers α , $a_N \leq \alpha$. Ainsi,

$$V(T) \sim \alpha N^2, \quad V(T) \leq \alpha N^2.$$

La nature de la suite demandée est donc : strictement croissante et convergente. Il n'est pas nécessaire de calculer exactement α .

11)(a) L'inégalité triangulaire et 9)(b) donnent, pour toute issue,

$$|T - N \ln N| \leq |T - E(T)| + |E(T) - N \ln N| = |T - E(T)| + Nu_N \leq |T - E(T)| + N.$$

Par conséquent, si $|T - N \ln N| \geq cN$, alors $|T - E(T)| \geq (c - 1)N$. On obtient l'inclusion

$$[|T - N \ln N| \geq cN] \subset [|T - E(T)| \geq (c - 1)N].$$

11)(b) Le moment d'ordre deux de T a été établi en 8), et $(c - 1)N > 0$. L'inclusion précédente puis Bienaymé-Tchebychev donnent

$$P(|T - N \ln N| \geq cN) \leq \frac{V(T)}{N^2(c - 1)^2} \leq \frac{\alpha}{(c - 1)^2}.$$

La probabilité est positive et son majorant tend vers 0 lorsque $c \rightarrow +\infty$. Pour N fixé, sa limite est donc 0.

12) Fixons $\varepsilon > 0$. Puisque $\ln N \rightarrow +\infty$, on a $\varepsilon \ln N > 1$ pour tout N assez grand. La question 11)(b), appliquée avec $c = \varepsilon \ln N$, fournit alors

$$0 \leq P(|T - N \ln N| \geq \varepsilon N \ln N) \leq \frac{\alpha}{(\varepsilon \ln N - 1)^2}.$$

Le dernier terme tend vers 0. Le théorème d'encadrement prouve la limite demandée. Cet argument est une majoration explicite de probabilités, sans recours à des règles générales de calcul sur les convergences de variables aléatoires.

13)(a) Question adaptée au programme actuel. Code adapté en Python. Une compréhension de liste produit les numéros 1 à 32 dans l'ordre :

Python

```
def init():
    return [i for i in range(1, 33)]
```

La borne supérieure 33 de range est exclue, ce qui donne exactement 32 cartes.

13)(b) Code adapté en Python. Le programme complet de l'insertion est :

Python

```
import numpy as np

def insertion(jeu):
    k = np.random.randint(0, 32)
    carte_dessus = jeu[0]
    for i in range(k):
        jeu[i] = jeu[i + 1]
    jeu[k] = carte_dessus
```

La variable k est l'indice final de la carte du dessus, donc la position mathématique vaut $k + 1$. Les cartes d'indices $1, \dots, k$ sont décalées d'un cran vers le sommet, puis la carte retirée est remplacée en indice k . Les affectations se font dans l'ordre croissant des indices, de sorte qu'une valeur encore nécessaire n'a pas été écrasée. Si $k == 0$, la boucle est vide et le paquet est inchangé; si $k == 31$, la carte du dessus est placée en dernière position.

13)(c) La fonction réinitialise le paquet, puis effectue et compte les insertions jusqu'à la première arrivée de C_{32} au sommet. La carte initialement au bas du paquet remonte d'une place lors de chaque succès; ce premier instant est T_{31} .

La valeur renvoyée est donc $T_{31} = T - 1$, et non T . La dernière insertion rendant l'ordre complètement uniforme, étudiée en 7), n'a pas encore été effectuée. Le comportement

de l'extrait original est conservé par la conversion ; la distinction entre sa fonction informatique et la variable mathématique est essentielle.

Le temps est fini presque sûrement, car il est une somme finie de variables géométriques de paramètres strictement positifs. Dans l'exemple à quatre cartes, la même règle s'arrête au temps 6, tandis que $T = 7$.

13)(d) Code adapté en Python. Le programme suivant contient toutes les fonctions et le programme principal :

Python

```
import numpy as np

def init():
    return [i for i in range(1, 33)]

def insertion(jeu):
    k = np.random.randint(0, 32)
    carte_dessus = jeu[0]
    for i in range(k):
        jeu[i] = jeu[i + 1]
    jeu[k] = carte_dessus

def temps(jeu):
    jeu[:] = init()
    n = 0
    while jeu[0] != 32:
        insertion(jeu)
        n = n + 1
    return n

jeu = init()
total = 0
for experience in range(100):
    total = total + temps(jeu)
```

```
moyenne = total / 100
print(moyenne)
```

La variable `total` additionne les 100 temps simulés, et `moyenne` est leur moyenne observée. Chaque appel réinitialise le paquet. La quantité théorique approchée est $E(T - 1) = 32H_{32} - 1 \simeq 128,872$, pas $32H_{32}$. Pour observer une moyenne de T , on additionnerait `temps(jeu) + 1`.

Les blocs complétés ont été exécutés : les 32 positions d'insertion préservent toutes les cartes et leur ordre attendu, et une simulation de 100 expériences se termine avec des temps entiers au moins égaux à 31. Une simulation illustre les résultats ; elle ne démontre ni l'uniformité ni les formules d'espérance.

Partie 3 — Distance variationnelle à la loi uniforme

14)(a) Question adaptée au programme actuel : une probabilité conditionnelle n'est écrite que si son événement de conditionnement a une probabilité strictement positive. Pour tout $t \leq n$ tel que $P(T = t) > 0$, le résultat renforcé de 7) donne $P(E_n \mid T = t) = \pi(A)$. Les événements $[T = t]$ étant disjoints,

$$P(E_n \cap [T \leq n]) = \sum_{t=1}^n P(E_n \cap [T = t]) = \sum_{t=1}^n \pi(A)P(T = t) = \pi(A)P(T \leq n).$$

Les termes de probabilité nulle valent 0 sans qu'il soit nécessaire de les conditionner. Si $P(T \leq n) > 0$, la division par cette probabilité donne $P_{[T \leq n]}(E_n) = \pi(A)$.

Si $P(T \leq n) = 0$, l'événement $E_n \cap [T \leq n]$ a lui aussi probabilité nulle ; l'identité non conditionnelle reste donc vraie. En particulier, $T \geq N$, car les $N - 1$ durées valent au moins 1 ; pour $n < N$, le conditionnement du PDF ne serait pas défini.

14)(b) L'inclusion $E_n \cap [T > n] \subset [T > n]$ donne directement $P(E_n \cap [T > n]) \leq P(T > n)$.

14)(c) Les événements $[T \leq n]$ et $[T > n]$ forment une partition. Par 14)(a) et 14)(b),

$$\mu_n(A) = P(E_n) = \pi(A)P(T \leq n) + P(E_n \cap [T > n]) \leq \pi(A) + P(T > n),$$

car $0 \leq \pi(A)$ et $P(T \leq n) \leq 1$.

15)(a) Comme μ_n et π sont des probabilités sur S_N , $\mu_n(A) = 1 - \mu_n(\bar{A})$ et $\pi(A) = 1 - \pi(\bar{A})$. Ainsi,

$$\mu_n(A) - \pi(A) = -[\mu_n(\bar{A}) - \pi(\bar{A})].$$

15)(b) La question 14)(c) appliquée à A donne $\mu_n(A) - \pi(A) \leq P(T > n)$. Appliquée à $\bar{A}$, elle donne, par 15)(a), $-\mu_n(A) + \pi(A) \leq P(T > n)$. Ces deux bornes équivalent à

$$|\mu_n(A) - \pi(A)| \leq P(T > n).$$

16) Il existe un nombre fini de parties de S_N , et chaque écart absolu est positif ou nul.

Le maximum définissant la distance existe donc et, par 15)(b),

$$0 \leq d(\mu_n, \pi) \leq P(T > n).$$

Méthode 1 — Contrôle par la variance

Pour $n > E(T)$, l'événement $[T > n]$ est inclus dans $[|T - E(T)| \geq n - E(T)]$. Le moment d'ordre deux de T existe et N est maintenant fixé. Bienaymé-Tchebychev fournit donc

$$P(T > n) \leq \frac{V(T)}{(n - E(T))^2} \rightarrow 0.$$

L'encadrement précédent entraîne $d(\mu_n, \pi) \rightarrow 0$.

Méthode 2 — Contrôle par la seule espérance

Comme $T \geq 0$ et $E(T) < +\infty$, pour tout entier $n \geq 1$ la définition de l'espérance discrète donne

$$E(T) = \sum_{k \geq 0} kP(T = k) \geq \sum_{k > n} nP(T = k) = nP(T > n).$$

Ainsi $P(T > n) \leq E(T)/n \rightarrow 0$, puis $d(\mu_n, \pi) \rightarrow 0$. Cette preuve particulière de la majoration nécessaire n'exige que le premier moment ; la première méthode exploite le moment d'ordre deux déjà calculé.

Partie 4 — Une majoration de $P(T > n)$

- 17) Le premier timbre reçu est nécessairement nouveau. Donc $S_1 = 1$ presque sûrement : S_1 suit la loi certaine de valeur 1.
- 18) Lorsque $k - 1$ timbres différents ont été obtenus, il reste $N - k + 1$ timbres nouveaux possibles. Chaque jour, la probabilité d'obtenir un nouveau timbre vaut donc $(N - k + 1)/N$, quel que soit l'ensemble des timbres déjà possédés.

Les choix journaliers étant indépendants, pour tout $r \geq 0$,

$$P(S_k > r) = ((k - 1)/N)^r.$$

Ainsi, S_k suit la loi géométrique sur $\mathbb{N}^*$ de paramètre $(N - k + 1)/N$. Le paramètre est strictement positif pour $2 \leq k \leq N$.

- 19) Pour $i \in \llbracket 1, N - 1 \rrbracket$, la variable S_{N-i+1} suit la loi géométrique de paramètre i/N , donc la même loi que Δ_i . Les deux familles sont indépendantes d'après les hypothèses de leurs modèles respectifs. Les vecteurs $(S_N, S_{N-1}, \dots, S_2)$ et $(\Delta_1, \dots, \Delta_{N-1})$ ont donc les mêmes probabilités conjointes, produits des mêmes probabilités marginales.

En additionnant leurs coordonnées, on obtient la même loi de somme. Puisque $S_1 = 1$ et $T = 1 + \Delta_1 + \dots + \Delta_{N-1}$, il vient

$S = 1 + S_N + S_{N-1} + \dots + S_2$ a la même loi que T .

Il ne suffirait pas de comparer uniquement les espérances et variances : c'est bien l'indépendance et l'égalité des lois de chaque terme qui permettent cette conclusion.

20)(a) La collection n'est pas complète après m jours si et seulement si au moins un timbre n'a pas encore été reçu. Par conséquent,

$$[S > m] = \bigcup_{j=1}^N B_j^m.$$

20)(b) Pour un timbre fixé j , la probabilité de ne pas le recevoir un jour donné vaut $1 - 1/N$. L'indépendance des m choix quotidiens donne

$$P(B_j^m) = (1 - 1/N)^m.$$

20)(c) L'inégalité rappelée dans l'énoncé, appliquée à l'union de 20)(a), fournit

$$P(S > m) \leq \sum_{j=1}^N P(B_j^m) = N(1 - 1/N)^m.$$

Les événements B_j^m ne sont pas supposés indépendants ; une telle hypothèse n'est pas nécessaire pour cette majoration.

21)(a) Méthode 1 — Étude d'une fonction

La fonction $g(x) = x - \ln(1 + x)$ est définie et dérivable sur $] -1, +\infty[$, avec $g'(x) = x/(1 + x)$. Elle est décroissante sur $] -1, 0]$ puis croissante sur $[0, +\infty[$. Elle admet donc un minimum en 0, de valeur $g(0) = 0$. Ainsi $\ln(1 + x) \leq x$ pour tout $x > -1$.

Méthode 2 — Comparaison d'intégrales

Si $x \geq 0$, alors $1/(1 + t) \leq 1$ sur $[0, x]$, donc $\ln(1 + x) = \int_0^x dt/(1 + t) \leq x$. Si

$-1 < x < 0$, alors $1/(1+t) \geq 1$ sur $[x, 0]$, donc

$$\ln(1+x) = - \int_x^0 dt/(1+t) \leq - \int_x^0 dt = x.$$

La première méthode identifie le minimum d'un écart; la seconde montre directement le rôle du sens des bornes d'intégration de part et d'autre de 0. Les deux donnent l'égalité pour $x = 0$.

21)(b) Pour $N \geq 2$, le réel $-1/N$ appartient à $] -1, 0[$. D'après 21)(a), $\ln(1 - 1/N) \leq -1/N$. Pour $m \geq 1$, la multiplication par m puis la croissance de l'exponentielle donnent

$$(1 - 1/N)^m = \exp(m \ln(1 - 1/N)) \leq e^{-m/N}.$$

Par 19) et 20)(c), $P(T > m) = P(S > m) \leq Ne^{-m/N}$. Pour $m = 0$, on a $P(T > 0) = 1 \leq N = Ne^0$. La majoration est donc valable pour tout entier naturel m .

22)(a) Soient $c > 0$ et un entier $n \geq N \ln N + cN$. Par 16) et 21)(b),

$$d(\mu_n, \pi) \leq P(T > n) \leq Ne^{-n/N} \leq Ne^{-\ln N - c} = e^{-c}.$$

La condition sur n est une condition suffisante explicite pour la précision de mélange demandée.

22)(b) On veut garantir $d(\mu_n, \pi) \leq 0,2 = 1/5$. On peut choisir $c = \ln 5 > 0$, puis appliquer 22)(a). Pour $N = 32$, il suffit que

$$n \geq 32 \ln 32 + 32 \ln 5 = 32 \ln 160 \simeq 162,406.$$

Le plus petit entier satisfaisant cette condition est 163. Il suffit donc de réaliser 163 insertions pour garantir la distance acceptable à l'aide de cette majoration. Ce nombre est une garantie issue de la borne démontrée; celle-ci ne permet pas d'affirmer que 162 insertions seraient effectivement insuffisantes.

Relevé des adaptations

Préambule : remplacement de l'identification de l'univers probabilisé à S_N muni de l'équiprobabilité. Cette identification est incompatible avec une configuration initiale

déterministe et des durées géométriques non bornées. Le processus est défini sur un espace adapté, tandis que π désigne seulement la loi uniforme cible sur S_N .

4), 5), 6) et 10)(b) : explicitation des tailles de paquet nécessaires aux temps mentionnés ; $N \geq 3$ pour T_2 , $N \geq 4$ pour T_3 , et $N \geq 2$ dans la famille étudiée. Cela évite notamment une division par zéro en 4)(b).

7) : ajout d'indications élémentaires sur les insertions successives dans l'ordre des cartes et sur le maintien de l'uniformité. La propriété conditionnelle au temps T , nécessaire à 14)(a), est démontrée sans supposer une théorie générale des temps d'arrêt.

9)(b)(ii) : rectification de la borne supérieure stricte erronée à $n = 1$. La borne large vaut dès $n = 1$; elle est stricte pour $n \geq 2$.

13)(a)–(d) : code adapté en Python, indexation de 0 à 31 et initialisations remplaçant les déclarations de l'ancien langage. Le comportement original est conservé : la fonction s'arrête à $T - 1$, et cette différence est explicitée en 13)(c). Les fonctions complétées ont été exécutées et contrôlées sur les positions d'insertion et une simulation de 100 expériences.

14)(a) : l'égalité conditionnelle est réservée à $P(T \leq n) > 0$; l'identité non conditionnelle est conservée pour tout $n \geq 1$, notamment pour $n < N$ où le conditionnement original est impossible.

18) et 20)(b) : explicitation de l'indépendance des choix quotidiens de timbres dans le modèle du collectionneur, nécessaire aux probabilités géométriques et aux produits de probabilités calculés.