Essec II 2011 - Correction

Partie 1 - Description et premiers résultats

1. Par télescopage,
$$\sum_{k=1}^{i} \Delta_k = T_1 + \sum_{k=2}^{i} (T_k - T_{k-1}) = T_1 + T_i - T_1 = T_i.$$
Δ_i représente le temps d'attente pour la i-ème remontée de la carte C_N à partir de la $(i-1)$-ème.

2. Loi de Δ_1.
$(\Delta_1 > n)$ signifie qu'aucune insertion ne s'est faite en $N^{ème}$ position durant les n premières insertions. En notant I_m la position de la $m^{ième}$ insertion on a donc
$$(\Delta_1 > n) = \bigcap_{m=1}^{n} (I_m \leqslant N-1)$$
et par indépendance et équiprobabilité des insertions
$$\mathrm{P}(\Delta_1 > n) = \prod_{m=1}^{n} \mathrm{P}(I_m \leqslant N-1) = \prod_{m=1}^{n} \frac{N-1}{N} = \left(\frac{N-1}{N}\right)^n$$
qui se généralise à $n = 0$ car $\left(\frac{N-1}{N}\right)^0 = 1 = \mathrm{P}(\Delta_1 > 0)$.

Ensuite on a $\Delta_1(\Omega) = [\![1; +\infty[\![$ donc on sait que la variable Δ_1 est discrète donc on peut écrire, avec une réunion incompatible :
$$(\Delta_1 > n-1) = (\Delta_1 = n) \cup (\Delta_1 > n) \quad \text{donc} \quad \mathrm{P}((\Delta_1 > n-1) = \mathrm{P}(\Delta_1 = n) + \mathrm{P}(\Delta_1 > n)$$
et enfin :
$$\mathrm{P}(\Delta_1 = n) = \left(\frac{N-1}{N}\right)^{n-1} - \left(\frac{N-1}{N}\right)^n = \left(\frac{N-1}{N}\right)^{n-1}\left(1 - \frac{N-1}{N}\right) = \frac{1}{N}\left(1 - \frac{1}{N}\right)^{n-1}$$
donc $\Delta_1 \hookrightarrow \mathcal{G}\left(\frac{1}{N}\right)$.

3. Soit $i \in [[2, N-1]]$. Loi de Δ_i

   (a) Pour $n = 0$, le résultat est évident car $\mathrm{P}(\Delta_i > 0) = 1 = \left(\frac{N-i}{N}\right)^0$.

   Pour tout $n \geqslant 1$, $\Delta_i > n$ signifie que les n premières insertions depuis la $(i-1)$-ème remontée (qui a amené C_N à la position $N-i+1$) n'ont pas fait remonter C_N. Cela signifie que ces n insertions se sont faites à des positions inférieures ou égales à la $(N-i)$-ème position, c'est-à-dire :
$$(\Delta_i > n) = \bigcap_{m=1}^{n} (I_m \leqslant N-i)$$
et par indépendance et équiprobabilité des insertions on obtient :
$$\mathrm{P}(\Delta_i > n) = \prod_{m=1}^{n} \frac{N-i}{N} = \left(\frac{N-i}{N}\right)^n.$$

De même que précédemment, on obtient alors $\Delta_i(\Omega) = [\![1;+\infty[\![$ puis :

$$\begin{aligned}\mathrm{P}(\Delta_i = n) &= \mathrm{P}(\Delta_i > n-1) - \mathrm{P}(\Delta_i > n) = \left(\frac{N-i}{N}\right)^{n-1} - \left(\frac{N-i}{N}\right)^{n} = \left(\frac{N-i}{N}\right)^{n-1}\left(1-\frac{N-i}{N}\right)\\ &= \frac{i}{N}\left(1-\frac{i}{N}\right)^{n-1}\end{aligned}$$

donc $\Delta_i \hookrightarrow \mathcal{G}\left(\frac{i}{N}\right)$.

(b) On en déduit que pour tout i, Δ_i admet une espérance et une variance, qui vaut :

$$E(\Delta_i) = \frac{N}{i} \quad \text{et} \quad V(\Delta_i) = \frac{1-\frac{i}{N}}{\left(\frac{i}{N}\right)^2} = \frac{N}{i^2}(N-i) = \frac{N(N-i)}{i^2}.$$

4. Loi de T_2. Soit $n \geqslant 2$.-

(a) Avec le système complet d'évènements $(\Delta_1 = k)_{k\in\mathbb{N}^*}$, les probabilités totales donnent :

$$\mathrm{P}(T_2 = n) = \sum_{k=1}^{+\infty} \mathrm{P}\left[(\Delta_1 = k)\cap(T_2 = n)\right] = \sum_{k=1}^{+\infty} \mathrm{P}\left[(\Delta_1 = k)\cap(k+\Delta_2 = n)\right] = \sum_{k=1}^{+\infty} \mathrm{P}\left[(\Delta_1 = k)\cap(\Delta_2 = n-k)\right]$$

Or pour $k \geqslant n$, $n-k \leqslant 0$ donc $(\Delta_2 = n-k)$ est impossible. Par indépendance de Δ_1 et Δ_2 on obtient alors :

$$\mathrm{P}(T_2 = n) = \sum_{k=1}^{n-1} \mathrm{P}(\Delta_1 = k)\,\mathrm{P}(\Delta_2 = n-k) + \sum_{k=n}^{+\infty} 0 = \sum_{k=1}^{n-1} \mathrm{P}(\Delta_1 = k)\,\mathrm{P}(\Delta_2 = n-k).$$

(b) On reconnaît la somme finie des termes d'une suite géométrique de raison $\frac{1-1/N}{1-2/N} \neq 1$, on obtient :

$$\begin{aligned}\sum_{k=1}^{n-1}\left(\frac{1-1/N}{1-2/N}\right)^k &= \left(\frac{1-1/N}{1-2/N}\right)\frac{1-\left(\frac{1-1/N}{1-2/N}\right)^{n-1}}{1-\frac{1-1/N}{1-2/N}} = (1-1/N)\frac{1-\left(\frac{1-1/N}{1-2/N}\right)^{n-1}}{1-\frac{2}{N}-1+\frac{1}{N}}\\ &= N\left(1-\frac{1}{N}\right)\left[\left(\frac{1-1/N}{1-2/N}\right)^{n-1}-1\right]\end{aligned}$$

(le dénominateur valait $\frac{-1}{N}$).

(c) On reprend la formule de 4)a) :

$$\begin{aligned}\mathrm{P}(T_2 = n) &= \sum_{k=1}^{n-1}\frac{1}{N}\left(1-\frac{1}{N}\right)^{k-1}\times\frac{2}{N}\left(1-\frac{2}{N}\right)^{n-k-1} = \frac{2}{N^2}\left(1-\frac{1}{N}\right)^{-1}\left(1-\frac{2}{N}\right)^{n-1}\sum_{k=1}^{n-1}\left(\frac{1-\frac{1}{N}}{1-\frac{2}{N}}\right)^k\\ &= \frac{2}{N^2}\left(1-\frac{1}{N}\right)^{-1}\left(1-\frac{2}{N}\right)^{n-1}\times N\left(1-\frac{1}{N}\right)\left[\left(\frac{1-1/N}{1-2/N}\right)^{n-1}-1\right]\\ &= \frac{2}{N}\left(1-\frac{2}{N}\right)^{n-1}\left[\left(\frac{1-1/N}{1-2/N}\right)^{n-1}-1\right] = \frac{2}{N}\left[\left(1-\frac{1}{N}\right)^{n-1}-\left(1-\frac{2}{N}\right)^{n-1}\right].\end{aligned}$$

5. De manière certaine, la carte C_N se retrouve en position $N-1$ à l'instant T_1 donc l'insertion se fait à la position N.

Par équiprobabilité des positions possible pour chaque insertion, l'insertion de l'instant T_2 se fait de manière équiprobable entre les positions $N-1$ et N, donc avec une probabilité $\frac{1}{2}$ à chaque fois.

Les deux évènements cherchés signifient que l'insertion à T_1 se fait en position N puis celle à T_2 en position N (resp. en position N puis en position $N-1$) de probabilité (par indépendance des insertions) :

$$1\times\frac{1}{2} = \frac{1}{2}.$$

6. (a) a_i peut prendre les valeurs $N-2$, $N-1$ et N pour tout i, mais pas en même temps ! Il y a donc $3! = 6$ résultats possibles : a_1 peut prendre 3 valeurs, puis a_2 les deux restantes et a_3 forcément la dernière restante.

   (b) Quelques exemples. à l'instant T_3 :

      i. $(a_1, a_2, a_3) = (N-2, N-1, N)$ signifie que les trois insertions ont été faites en position N, et sa probabilité est par indépendance des insertions et équiprobabilité entre les positions sous C_N à chaque fois :
$$1 \times \frac{1}{2} \times \frac{1}{3} = \frac{1}{6}.$$

      ii. $(a_1, a_2, a_3) = (N-2, N, N-1)$ signifie que la première insertion se fait à N, la seconde à N et la troisième à $N-1$, et sa probabilité est par indépendance des insertions et équiprobabilité entre les positions sous C_N à chaque fois :
$$1 \times \frac{1}{2} \times \frac{1}{3} = \frac{1}{6}.$$

7. « A partir de l'instant T, toutes les configurations du jeu de cartes sont équiprobables »

   A chaque instant T_i $(i < N)$ la carte insérée en dessous de C_N l'est équiprobablement sur chacune des positions $[\![N-i+1, N]\!]$ et on a alors (par récurrence) équiprobabilité de toutes les permutations pour les positions des cartes qui se trouvent sous C_N.

   A l'instant $T_{N-1}+1$, on insère enfin la carte C_N qui se retrouve équiprobablement en toutes positions.

   Au final, toutes les permutations seront équiprobables à l'instant T.

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

8. Chaque Δ_i admet une espérance et une variance, donc par linéarité de l'espérance et indépendance des Δ_i, T admet une espérance et une variance et on a :
$$E(T) = E\left(\sum_{i=1}^{N-1} \Delta_i + 1\right) = 1 + \sum_{i=1}^{N-1} \frac{N}{i} = 1 + \sum_{i=1}^{N} \frac{N}{i} - \frac{N}{N} = N \sum_{i=1}^{N} \frac{1}{i} = NH_N$$
et (le terme en $i = N$ est nul) :
$$V(T) = V\left(\sum_{i=1}^{N-1} \Delta_i + 1\right) = \sum_{i=1}^{N-1} N\frac{N-i}{i^2} = N\sum_{i=1}^{N} \frac{N-i}{i^2} = N^2 \sum_{i=1}^{N} \frac{1}{i^2} - N\sum_{i=1}^{N} \frac{1}{i} = N^2 \sum_{i=1}^{N} \frac{1}{i^2} - NH_N.$$

9. (a) Pour tout $k \geqslant 1$ et $t \in [k; k+1]$ on a par décroissance de la fonction inverse :
$$\frac{1}{k+1} \leqslant \frac{1}{t} \leqslant \frac{1}{k}$$
   puis en intégrant avec les bornes dans l'ordre croissant :
$$\int_k^{k+1} \frac{1}{k+1}\, dt \leqslant \int_k^{k+1} \frac{1}{t}\, dt \leqslant \int_k^{k+1} \frac{1}{k}\, dt$$
   et en remarquant que les intégrales de gauche et de droite sont des intégrales de constantes :
$$\frac{1}{k+1} \leqslant \int_k^{k+1} \frac{1}{t}\, dt \leqslant \frac{1}{k}.$$

   (b) i. Pour tout $n \geqslant 1$, on a :
$$\begin{aligned} u_{n+1} - u_n &= H_{n+1} - \ln(n+1) - H_n + \ln(n) = \frac{1}{n+1} - \ln(n+1) + \ln(n) \\ &= \frac{1}{n+1} - [\ln(t)]_n^{n+1} = \frac{1}{n+1} - \int_n^{n+1} \frac{1}{t}\, dt \leqslant 0 \end{aligned}$$
   donc la suite $(u_n)_{n\geqslant 1}$ est décroissante.

ii. Ensuite on somme les inégalités de droite de l'encadrement pour k allant de 1 à n, on obtient :

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

Enfin on somme les inégalités de gauche de l'encadrement pour k allant de 1 à $n-1$, on obtient :

$$\sum_{k=1}^{n-1} \frac{1}{k+1} \leqslant \int_1^{n} \frac{1}{t}\, dt \Longleftrightarrow \sum_{k=2}^{n} \frac{1}{k} \leqslant \ln(n) \Longleftrightarrow H_n - 1 \leqslant \ln n \Longleftrightarrow H_n \leqslant \ln(n) + 1$$

et enfin on a bien :

$$\ln(n+1) \leqslant H_n \leqslant \ln(n) + 1.$$

(c) On en déduit que pour tout $n \geqslant 1$,

$$\ln(n+1) - \ln n \leqslant u_n \leqslant 1$$

et par croissance du ln, on a :

$$n < n+1 \Longrightarrow \ln(n) < \ln(n+1) \Longrightarrow \ln(n+1) - \ln(n) > 0$$

donc (u_n) est décroissante et minorée par 0 donc converge vers une limite γ. De plus en passant à la limite on obtient :

$$0 \leqslant u_n \leqslant 1 \Longrightarrow 0 \leqslant \gamma \leqslant 1.$$

10. (a) On a tout d'abord :

$$\frac{E(T)}{N \ln N} = \frac{H_N}{\ln N}$$

et en reprenant l'encadrement de H_N :

$$\frac{\ln(N+1)}{\ln N} \leqslant \frac{H_N}{\ln N} \leqslant 1 + \frac{1}{\ln N}$$

Le terme de droite tend clairement vers 1, montrons que c'est aussi le cas de celui de gauche :

$$\frac{\ln(N+1)}{\ln N} = \frac{\ln\left(N\left[1+\frac{1}{N}\right]\right)}{\ln N} = \frac{\ln N + \ln\left(1+\frac{1}{N}\right)}{\ln N} = 1 + \frac{\ln\left(1+\frac{1}{N}\right)}{\ln N} \xrightarrow[N\to+\infty]{} 1$$

par somme, composée et quotient de limites.

On en déduit par encadrement que

$$\frac{E(T)}{N \ln N} = \frac{H_N}{\ln N} \xrightarrow[N\to+\infty]{} 1 \quad \text{et} \quad E(T) \underset{+\infty}{\sim} N \ln N$$

D'autre part montrons que

$$E(T) - N\ln N - N\gamma = o(N) \Longleftrightarrow \frac{E(T) - N\ln N - N\gamma}{N} \xrightarrow[N\to+\infty]{} 0.$$

En effet on a :

$$\frac{E(T) - N\ln N - N\gamma}{N} = \frac{NH_N - N\ln N - N\gamma}{N} = H_N - \ln N - \gamma = u_N - \gamma \xrightarrow[N\to+\infty]{} \gamma - \gamma = 0.$$

ce qui prouve que :

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

(b) On a vu que

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

D'où :

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

Or on remarque que la somme est la somme partielle d'une série de Riemann convergente (car $2 > 1$) donc elle converge en $+\infty$ vers

$$\alpha = \sum_{k=1}^{+\infty} \frac{1}{k^2} > 0$$

tandis que

$$\frac{E(T)}{N^2} \sim \frac{N \ln N}{N^2} = \frac{\ln N}{N} \xrightarrow[N \to +\infty]{} 0$$

par croissances comparées, donc

$$\frac{V(T)}{N^2} \xrightarrow[N \to +\infty]{} \alpha > 0$$

Comme cette limite est finie et non nulle, c'est un équivalent de $\frac{V(T)}{N^2}$, puis on multiplie cet équivalent par N^2 :

$$V(T) \underset{+\infty}{\sim} \alpha N^2$$

Enfin on considère :

$$V(T) - \alpha N^2 = N^2 \left(\sum_{k=1}^{N} \frac{1}{k^2} - \sum_{k=1}^{+\infty} \frac{1}{k^2} \right) - E(T) = - \left(N^2 \sum_{k=N+1}^{+\infty} \frac{1}{k^2} + E(T) \right) \leqslant 0$$

car $E(T)$ est positive et une somme de termes positifs est positive, ce qui donne :

$$V(T) \leqslant \alpha N^2.$$

11. (a) On fait apparaître $E(T)$:

$$\big|T(\omega) - N \ln N\big| = \big|\big(T(\omega) - E(T)\big) + \big(E(T) - N \ln N\big)\big|$$

puis par inégalité triangulaire :

$$\big|T(\omega) - N \ln N\big| \leqslant \big|T(\omega) - E(T)\big| + \big|E(T) - N \ln N\big|$$

Or on remarque que :

$$E(T) - N \ln N = N H_N - N \ln N = N u_N \geqslant 0$$

car $u_N \geqslant 0$, donc

$$|E(T) - N \ln N| = E(T) - N \ln N = N u_N \leqslant N$$

car $u_N \leqslant 1$, et on obtient enfin :

$$\big|T(\omega) - N \ln N\big| \leqslant \big|T(\omega) - E(T)\big| + N.$$

On en déduit finalement que :

$$(|T - N \ln(N) \geqslant cN|) \Longrightarrow (|T - E(T)| + N \geqslant |T - N \ln(N) \geqslant cN) \Longrightarrow (|T - E(T)| \geqslant cN - N = (c-1)N)$$

donc l'évènement de gauche est inclus dans celui de droite :

$$(|T - N \ln(N) \geqslant cN|) \subset (|T - E(T)| \geqslant (c-1)N).$$

(b) Par croissance de la probabilité puis par l'inégalité de Bienaymé-Chebychev on en déduit :

$$\mathrm{P}(|T - N \ln(N) \geqslant cN|) \leqslant \mathrm{P}(|T - E(T)| \geqslant (c-1)N) \leqslant \frac{V(T)}{(c-1)^2 N^2} \leqslant \frac{\alpha N^2}{(c-1)^2 N^2} = \frac{\alpha}{(c-1)^2}.$$

Par encadrement (une probabilité est positive) on en déduit qu'à N fixé,

$$\lim_{c \to +\infty} \mathrm{P}(|T - N \ln(N) \geqslant cN|) = 0.$$

12. Pour tout $\varepsilon > 0$,
$$\varepsilon \ln N \xrightarrow[N\to+\infty]{} +\infty$$
donc cette quantité dépasse 1 à partir d'un certain rang.

A partir de ce rang, on obtient donc en remplaçant c par $\varepsilon \ln N$:
$$0 \leqslant \mathrm{P}(|T - N\ln(N) \geqslant \varepsilon N \ln(N)|) \leqslant \frac{\alpha}{(\varepsilon \ln N - 1)^2}$$
et par encadrement,
$$\lim_{N\to+\infty} \mathrm{P}(|T - N\ln(N) \geqslant \varepsilon N \ln(N)|) = 0.$$

13. (a)
```
function Jeu=Init()
    Jeu=1:32
endfunction
```
(b)
```
function Jeu=Insertion(Jeu)
    k= floor(rand()*32)+1
    cartedessus=Jeu(1)
    if k >1 then
         for i=1:k-1 do
               Jeu(i)=Jeu(i+1)
        end
        Jeu(k)=cartedessus
    end
endfunction
```
(c) La fonction T effectue les insertions jusqu'à ce que la carte 32 (celle du dessous) se retrouve au dessus. La variable n compte le nombre d'insertions nécessaires pour y parvenir.

Donc `T` simule T_{N-1} (et non pas T, il faudrait encore rajouter 1).

C'est donc un de moins que le nombre d'insertions pour avoir un paquet bien mélangé.

(d)
```
S=0
for k=1:100 do
    S= S+T()
end
disp(S/100)
```

Partie 3 - Distance variationnelle à la loi uniforme

14. (a) A partir de l'instant T, le paquet a été mélangé et toutes les configurations sont équiprobables.

Donc sachant l'instant $n \geqslant T$, la probabilité est l'équiprobable : c'est π. On en déduit par probabilités composées que :
$$\mathrm{P}(E_n \cap [T \leqslant n]) = \mathrm{P}(T \leqslant n)\,\mathrm{P}_{(T\leqslant n)}(E_n) = \pi(A)\,\mathrm{P}(T \leqslant n).$$

(b) Par définition d'une intersection, puis par croissance de la probabilité :
$$(E_n \cap [T > n]) \subset (T > n) \quad \text{donc} \quad \mathrm{P}(E_n \cap [T > n]) \leqslant \mathrm{P}(T > n).$$

(c) Par définition $\mu_n(A) = \mathrm{P}(E_n)$. Avec le système complet d'évènements $([T \leqslant n], [T > n])$ les probabilités totales donnent :
$$\mu_n(A) = \mathrm{P}(E_n) = \mathrm{P}(E_n \cap [T \leqslant n]) + \mathrm{P}(E_n \cap [T > n]) \leqslant \pi(A)\,\mathrm{P}(T \leqslant n) + \mathrm{P}(T > n).$$

15. (a) μ_n et π sont des probabilités donc :
$$\mu_n(\overline{A}) = 1 - \mu_n(A) \quad \text{et} \quad \pi(\overline{A}) = 1 - \pi(A).$$
On en déduit que :
$$\mu_n(\overline{A}) - \pi(\overline{A}) = 1 - \mu_n(A) - 1 + \pi(A) = \pi(A) - \mu_n(A) = -\left[\mu_n(A) - \pi(A)\right].$$

(b) La question 14)c) donne immédiatement :

$$\mu_n(A) - \pi(A) \leqslant \mathrm{P}(T > n)$$

D'autre part les résultats obtenus jusqu'ici sont vrais pour une partie quelconque A, donc ils sont aussi vrais pour $\overline{A}$: donc

$$\mu_n(\overline{A}) - \pi(\overline{A}) \leqslant \mathrm{P}(T > n) \Longrightarrow -\left[\mu_n(A) - \pi(A)\right] \leqslant \mathrm{P}(T > n) \Longrightarrow \mu_n(A) - \pi(A) \geqslant -\mathrm{P}(T > n)$$

On obtient l'encadrement :

$$-\mathrm{P}(T > n) \leqslant \mu_n(A) - \pi(A) \leqslant \mathrm{P}(T > n) \Longleftrightarrow |\mu_n(A) - \pi(A)| \leqslant \mathrm{P}(T > n).$$

16. Pour toute partie A, la quantité $|\mu_n(A) - \pi(A)|$ est comprise entre 0 et $\mathrm{P}(T > n)$.

On en déduit que le maximum sur toutes les parties A de cette quantité est aussi compris entre ces deux valeurs, c'est-à-dire :

$$0 \leqslant d(\mu_n, \pi) \leqslant \mathrm{P}(T > n).$$

Enfin T est une variable aléatoire donc sa fonction de répartition vérifie :

$$\mathrm{P}(T \leqslant n) \xrightarrow[n \to +\infty]{} 1 \Longrightarrow \mathrm{P}(T > n) = 1 - \mathrm{P}(T \leqslant n) \xrightarrow[n \to +\infty]{} 0.$$

Par encadrement on en déduit que :

$$d(\mu_n, \pi) \xrightarrow[n \to +\infty]{} 0.$$

Partie 4 Une majoration de $\mathrm{P}(T > n)$

17. S_1 est le temps d'attente du premier timbre qui arrive forcément le premier jour donc S_1 est certaine égale à 1.

18. Pour tout $k \in [\![2; N]\!]$, S_k est le temps d'attente du premier succès dans une succession d'épreuves indépendantes et identiques (car tant qu'on ne reçoit pas de nouveau timbre, la probabilité d'en obtenir un nouveau est toujours la même), dont la probabilité de succès est :

$$\mathrm{P}(S) = \frac{N - (k-1)}{N}$$

car le collectionneur a déjà $k-1$ timbres, il lui en faut un des $[N - (k-1)]$ autres, et les timbres sont équiprobables.

On en déduit que $S_k \hookrightarrow \mathcal{G}\left(1 - \frac{k-1}{N}\right)$.

19. On remarque que S_k suit la même loi que Δ_{N-k+1} et on écrit :

$$S = 1 + \sum_{k=2}^{N} S_k$$

et comme les S_k sont indépendants les uns des autres comme les Δ_k, S suit la même loi que (avec un changement d'indice $l = N - k + 1$) :

$$1 + \sum_{k=2}^{N} \Delta_{N-k+1} = 1 + \sum_{l=1}^{N-1} \Delta_l = \sum_{k=1}^{N-1} \Delta_k + 1 = T$$

20. Soit $m \in \mathbb{N}^*$

(a) $(S > m)$ signifie que le $m^{ième}$ jour, le collectionneur n'a toujours pas les N timbres. Donc que l'un au moins des timbre n'a pas été reçu au jour m :

$$(S > m) = \bigcup_{j=1}^{N} B_j^m.$$

(b) Pour tout entier $j \in [\![1, N]\!]$, B_j^m signifie que le timbre j n'a pas été reçu pendant les m jours (intersection d'événements indépendants, chacun de probabilité $\frac{N-1}{N}$). D'où :

$$\mathrm{P}(B_j^m) = \left(1 - \frac{1}{N}\right)^m.$$

(c) On a vu que

$$(S > m) = \bigcup_{j=1}^{N} B_j^m \quad \text{donc} \quad \mathrm{P}(S > m) \leqslant \sum_{j=1}^{N} \mathrm{P}(B_j^m) = \sum_{k=1}^{N} \left(1 - \frac{1}{N}\right)^m = N\left(1 - \frac{1}{N}\right)^m.$$

21. (a) La fonction $x \to \ln(1+x)$ étant concave, sa courbe est en dessous de sa tangente en 0 d'équation $y = \ln'(1) \times (x - 0) + \ln 1 = x$. D'où

$$\forall x \in]-1; +\infty[\quad , \quad \ln(1 + x) \leqslant x.$$

(b) Comme T et S suivent la même loi,

$$\mathrm{P}(T > m) = \mathrm{P}(S > m) \leqslant N\left(1 - \frac{1}{N}\right)^m = Ne^{m\ln\left(1-\frac{1}{N}\right)}.$$

Or $\frac{1}{N} < 1$ donc $-\frac{1}{N} > -1$ et on applique l'inégalité ci-dessus :

$$\ln\left(1 - \frac{1}{N}\right) \leqslant -\frac{1}{N} \Longrightarrow m\ln\left(1 - \frac{1}{N}\right) \leqslant -\frac{m}{N} \Longrightarrow e^{m\ln\left(1-\frac{1}{N}\right)} \leqslant e^{-\frac{m}{N}} \Longrightarrow Ne^{m\ln\left(1-\frac{1}{N}\right)} \leqslant Ne^{-\frac{m}{N}}.$$

Enfin on obtient bien :

$$\mathrm{P}(T > m) \leqslant Ne^{-\frac{m}{N}}.$$

22. On reprend les notations introduites dans la partie précédente.

(a) On a vu à la fin de la partie III (puis on utilise le résultat ci-dessus) que :

$$0 \leqslant d(\mu_n, \pi) \leqslant \mathrm{P}(T > n) \leqslant Ne^{-\frac{n}{N}}.$$

Or on suppose que :

$$n \geqslant N\ln(N) + cN \Longrightarrow \frac{n}{N} \leqslant \ln(N) + c \Longrightarrow -\frac{n}{N} \leqslant -c - \ln N$$

ce qui donne :

$$d(\mu_n, \pi) \leqslant Ne^{-c-\ln N} = e^{-c}\frac{N}{e^{\ln N}} = e^{-c}.$$

(b) On veut $d(\mu_n, \pi) \leqslant 0,2$. Il suffit pour cela que

$$e^{-c} \leqslant 0,2 \Longleftrightarrow -c \leqslant \ln(0,2) = \ln(1/5) = -\ln 5 \Longleftrightarrow c \geqslant \ln 5$$

Il faut alors réaliser un nombre n d'insertions supérieur à :

$$32\ln(32) + 32\ln 5 = 32(\ln 32 + \ln 5) = 32(5\ln 2 + \ln 5)$$

donc à partir de

$$n = \lfloor 32(5\ln 2 + \ln 5)\rfloor + 1$$

le paquet est mélangé de façon acceptable.