

Préliminaire

1. $[Y > k] = \cup_{k+1 \leq i \leq N} [Y = i]$ et ces derniers événements sont incompatibles...

$$\sum_{k=0}^{N-1} P(Y > k) = \sum_{k=0}^{N-1} \sum_{i=k+1}^N P(Y = i), \text{ et par inversion de sommes finies, avec } 0 \leq k < i \leq N,$$

on a :

$$\sum_{k=0}^{N-1} P(Y > k) = \sum_{i=1}^N \sum_{k=0}^{i-1} P(Y = i) = \sum_{i=1}^N i P(Y = i) = E(Y).$$

$$\sum_{k=0}^{N-1} (2k+1)P(Y > k) = \sum_{k=0}^{N-1} \sum_{i=k+1}^N (2k+1)P(Y = i) = \sum_{i=1}^N \sum_{k=0}^{i-1} (2k+1)P(Y = i)$$

Or pour i fixé supérieur ou égal à 1, $\sum_{k=0}^{i-1} (2k+1) = 2 \sum_{k=0}^{i-1} k + \sum_{k=0}^{i-1} 1 = \frac{(i-1)i}{2} + i = i^2$ d'où :

$$\sum_{k=0}^{N-1} (2k+1)P(Y > k) = \sum_{i=1}^N i^2 P(Y = i) = E(Y^2)$$

Partie I : Inf et Sup

2. $E(U_1) = \frac{N+1}{2} \quad V(U_1) = \frac{N^2-1}{12}$

3. (a) $P(T_n \leq k) = P(\max(U_1, \dots, U_n) \leq k) = P\left(\bigcap_{i=1}^n [U_i \leq k]\right)$ et les variables aléatoires

$U_1, \dots, U_n$ sont mutuellement indépendantes :

$$P(T_n \leq k) = \prod_{i=1}^n P(U_i \leq k) = (P(U_1 \leq k))^n = \left(\frac{k}{N}\right)^n$$

- (b) Comme $[T_n \leq k] = [T_n \leq k-1] \cup [T_n = k]$ avec incompatibilité :

$$P(T_n = k) = P(T_n \leq k) - P(T_n \leq k-1) = \left(\frac{k}{N}\right)^n - \left(\frac{k-1}{N}\right)^n$$

4. (a) Pour $N \geq 2$ et pour tout k de $\{1, \dots, N-1\}$, $0 < \frac{k}{N} < 1$ et donc $\lim_{n \rightarrow +\infty} \left(\frac{k}{N}\right)^n = 0$.

Par somme finie de termes ayant 0 pour limite : $\forall N \geq 2, \lim_{n \rightarrow +\infty} d_n(N) = 0$.

Cette limite est encore valable pour $N = 1$ d'après la définition de $d_n(1)$.

(b) $E(T_n) = \sum_{k=0}^{N-1} P(T_n > k) = \sum_{k=0}^{N-1} 1 - P(T_n \leq k) = N - \sum_{k=1}^{N-1} P(T_n \leq k) = N - d_n(N).$

On en déduit immédiatement que : $\lim_{n \rightarrow +\infty} E(T_n) = N$.

(c) $E(T_n^2) = \sum_{k=0}^{N-1} (2k+1)P(T_n > k) = 2 \sum_{k=0}^{N-1} kP(T_n > k) + E(T_n) = E(T_n) + 2 \sum_{k=0}^{N-1} k(1 - P(T_n \leq k))$

$$E(T_n^2) = N - d_n(N) + N(N-1) - 2 \sum_{k=0}^{N-1} k \left(\frac{k}{N}\right)^n - d_n(N) + N^2 - 2N \sum_{k=0}^{N-1} \left(\frac{k}{N}\right)^{n+1}$$

$$E(T_n^2) = -d_n(N) + N^2 - 2Nd_{n+1}(N)$$

Par la formule de Huygens, on obtient :

$$V(T_n) = E(T_n^2) - E^2(T_n) = -d_n(N) + N^2 - 2Nd_{n+1}(N) - (N - d_n(N))^2$$

$$V(T_n) = -d_n(N) - 2Nd_{n+1}(N) + 2Nd_n(N) - d_n^2(N)$$

$$V(T_n) = (2N-1)d_n(N) - 2Nd_{n+1}(N) - d_n^2(N)$$

Avec le résultat de 4a), on a immédiatement : $\lim_{n \rightarrow +\infty} V(T_n) = 0$.

(d) Soit $N \geq 2, \forall n \in \mathbb{N}^*, d_n(N) = \frac{1^n + 2^n + \dots + (N-1)^n}{N^n}$

On en déduit que $\frac{d_{n+1}(N)}{d_n(N)} = \frac{1}{N} \cdot \frac{1^{n+1} + 2^{n+1} + \dots + (N-1)^{n+1}}{1^n + 2^n + \dots + (N-1)^n}$.

Après mise en facteur de $(N-1)^{n+1}$ au numérateur et de $(N-1)^n$ au dénominateur, on peut écrire : $\frac{d_{n+1}(N)}{d_n(N)} = \frac{N-1}{N} \cdot \frac{\alpha_n}{\beta_n}$.

avec $\alpha_n = 1 + \sum_{k=1}^{N-2} \left(\frac{k}{N-1}\right)^{n+1}$ et $\beta_n = 1 + \sum_{k=1}^{N-2} \left(\frac{k}{N-1}\right)^n$. Par limite d'une somme finie, on a : $\lim_{n \rightarrow +\infty} \alpha_n = 1$ et $\lim_{n \rightarrow +\infty} \beta_n = 1$ ce qui entraîne

$$\lim_{n \rightarrow +\infty} \frac{d_{n+1}(N)}{d_n(N)} = \frac{N-1}{N}$$

On aura donc : $\frac{V(T_n)}{d_n(N)} = 2N - 1 - 2N \frac{d_{n+1}(N)}{d_n(N)} - d_n(N)$

et donc par passage à la limite lorsque n tend vers $+\infty$, on obtient directement

$$\lim_{n \rightarrow +\infty} \frac{V(T_n)}{d_n(N)} = 2N - 1 - 2N(1 - 1/N) = 1, \text{ ce qui signifie que } V(T_n) \sim d_n(N).$$

$$5. P(Z_n > k) = P\left(\bigcap_{i=1}^n [U_i > k]\right) = \prod_{i=1}^n P(U_i > k) = (P(U_1 > k))^n = (1 - P(U_1 \leq k))^n = \left(\frac{N-k}{N}\right)^n$$

On en déduit que $P(Z_n = k) = P(Z_n > k-1) - P(Z_n > k)$ et finalement $P(Z_n = k) = \left(\frac{N-(k-1)}{N}\right)^n - \left(\frac{N-k}{N}\right)^n = \left(\frac{N+1-k}{N}\right)^n - \left(\frac{N-k}{N}\right)^n$.

Or si $k \in \{1, \dots, N\}$ alors $(N+1-k) \in \{1, \dots, N\}$ et

$$P(Z_n = k) = P(T_n = N+1-k) = P(N+1-T_n = k)$$

On en déduit que Z_n a même loi que $N+1-T_n$, ce qui permet de donner directement $E(Z_n)$ et $V(Z_n)$: $E(Z_n) = E(N+1-T_n) = N+1 - E(T_n) = 1 + d_n(N)$

$$V(Z_n) = V(T_n) = (2N-1)d_n(N) - 2Nd_{n+1}(N) - d_n^2(N).$$

6. Voici une proposition de fonction en Scilab :

```
function T=simulmax(n);
T=max(grand(1,n,'uin',1,N));
endfunction;
```

Partie II. Couple (Inf,Sup)

7. $\Phi_n(k, l) = P([T_n \leq k] \cap [Z_n \leq l])$.

(a) Soit $(k, l) \in \{1, \dots, N\}^2$.

- Si $k \leq l$ alors $\Phi_n(k, l) = P(T_n \leq k) \cdot P_{[T_n \leq k]}(Z_n \leq l)$.

On sait que si $[T_n \leq k]$ est réalisé alors $[Z_n \leq l]$ aussi et donc $P_{[T_n \leq k]}(Z_n \leq l) = 1$.

On aura donc $\Phi_n(k, l) = P(T_n \leq k) = \left(\frac{k}{N}\right)^n$.

- Si $k > l$ alors en utilisant le système complet d'événements $[[Z_n \leq l], [Z_n > l]]$, on peut écrire : $P(T_n \leq k) = \Phi_n(k, l) + P([T_n \leq k] \cap [Z_n > l])$.

Or $[T_n \leq k] \cap [Z_n > l] = \bigcap_{i=1}^n [l < U_i \leq k]$, ce qui permet d'écrire : $P([T_n \leq k] \cap [Z_n > l]) = (P(l+1 \leq U_1 \leq k))^n = \left(\sum_{i=l+1}^k P(U_1 = i) \right)^n = \left(\frac{k-l}{N} \right)^n$

et finalement $\Phi_n(k, l) = P(T_n \leq k) - P([T_n \leq k] \cap [Z_n > l]) = \left(\frac{k}{N} \right)^n - \left(\frac{k-l}{N} \right)^n$.

On a donc :

$$\Phi_n(k, l) = \begin{cases} \left(\frac{k}{N} \right)^n & \text{si } k \leq l \\ \left(\frac{k}{N} \right)^n - \left(\frac{k-l}{N} \right)^n & \text{si } k > l \end{cases}$$

En fait la dernière égalité est vraie pour $k \geq l$ puisque dans le cas $k = l$ on retrouve $\Phi_n(k, l) = \frac{k^n}{N^n}$.

- (b) Soit $(k, l) \in \{1, \dots, N\}^2$. Notons $A_{k,l} = [T_n = k] \cap [Z_n = l]$. On sait que $[T_n \leq k] = [T_n = k] \cup [T_n < k] = [T_n = k] \cup [T_n \leq k-1]$ et il y a incompatibilité des événements des unions précédentes, on en déduit que :

$$\Phi_n(k, l) = P([T_n \leq k] \cap [Z_n \leq l]) = P([T_n = k] \cap [Z_n \leq l]) + P([T_n < k] \cap [Z_n \leq l])$$

$$\text{Or } P([T_n < k] \cap [Z_n \leq l]) = P([T_n \leq k-1] \cap [Z_n \leq l]) = \Phi_n(k-1, l)$$

$$\text{et } P([T_n = k] \cap [Z_n \leq l]) = P(T_n = k, Z_n = l) + P([T_n = k] \cap [Z_n \leq l-1])$$

Or $P([T_n = k, Z_n \leq l-1]) = \Phi_n(k, l-1) - P(T_n \leq k-1, Z_n \leq l-1)$ d'où finalement $\Phi_n(k, l) = P(A_{k,l}) + \Phi_n(k, l-1) - \Phi(k-1, l-1) + \Phi_n(k-1, l)$ Ce qui donne le résultat voulu : $P(A_{k,l}) = \Phi_n(k, l) + \Phi_n(k-1, l-1) - \Phi_n(k, l-1) - \Phi_n(k-1, l)$

- (c) • Si $k < l$ alors il est clair que $[T_n = k] \cap [Z_n = l] = \emptyset$ (le maximum de réels ne peut pas être strictement inférieur au minimum des ces réels). On aura donc $P(T_n = k, Z_n = l) = 0$

- Si $k = l$ alors $[T_n = k, Z_n = l = k] = \bigcap_{i=1}^n [U_i = k]$, par indépendance des variables

$$U_1, \dots, U_n, \text{ on a : } P(T_n = k, Z_n = k) = \prod_{i=1}^n P(U_i = k) = \frac{1}{N^n}$$

- Si $k > l$ alors $k-1 > l-1$, $k-1 \geq l$ et $k > l-1$. On peut remarquer les égalités données en 7a) et 7b) permettent d'écrire : $P(T_n = k, Z_n = l) = \left(\frac{k-l+1}{N} \right)^n + \left(\frac{k-1-l}{N} \right)^n - 2 \left(\frac{k-l}{N} \right)^n$.

8. (a) Par le théorème de transfert, les variables T_n et Z_n étant discrètes finies, on sait que $T_n Z_n$ admet une espérance et

$$E(T_n Z_n) = \sum_{k=1}^N \sum_{l=1}^N kl P(T_n = k, Z_n = l)$$

D'après ce qui a été écrit précédemment, on aura donc :

$$E(T_n Z_n) = \sum_{k=2}^N \sum_{l=1}^{k-1} kl P(T_n = k, Z_n = l) + \sum_{k=1}^N \frac{k^2}{N^n} + \sum_{k=1}^{N-1} \sum_{l=k+1}^N kl P(T_n = k, Z_n = l)$$

$$E(T_n Z_n) = \sum_{k=2}^N \sum_{l=1}^{k-1} \left(kl \left(\frac{k-l+1}{N} \right)^n + \left(\frac{k-1-l}{N} \right)^n - 2 \left(\frac{k-l}{N} \right)^n \right) + \frac{1}{N^n} \sum_{k=1}^N k^2$$

On pose $j = k - l$, on obtient alors

$$E(T_n Z_n) = \frac{1}{N^n} \sum_{k=1}^N k^2 + \frac{1}{N^n} \sum_{k=2}^N \sum_{j=1}^{k-1} k(k-j)[(j+1)^n + (j-1)^n - 2j^n]$$

$$E(T_n Z_n) = \frac{1}{N^n} \sum_{k=1}^N k^2 + \frac{1}{N^n} \sum_{k=2}^N [k^2(k^n - (k-1)^n - 1)] - \frac{1}{N^n} \sum_{k=2}^N [k(k-1)k^n - k^2(k-1)^n]$$

$$E(T_n Z_n) = \frac{1}{N^n} \sum_{k=1}^N k^2 - \frac{1}{N^n} \sum_{k=2}^N k^2 + \frac{1}{N^n} \sum_{k=2}^N k^{n+1} = \frac{1}{N^n} + N \sum_{k=2}^N \left(\frac{k}{N}\right)^{n+1} = N \sum_{k=1}^N \left(\frac{k}{N}\right)^{n+1}$$

Si $N = 1$ alors $E(T_n Z_n) = 1 = N(1 + d_{n+1}(N))$.

Si $N \geq 2$, alors $E(T_n Z_n) = N + N \sum_{k=1}^{N-1} \left(\frac{k}{N}\right)^{n+1}$ et finalement :

$$\forall N \geq 1, \quad E(T_n Z_n) = N(1 + d_{n+1}(N))$$

(b) On déduit de ce qui précède que pour $N \geq 2$: $Cov(T_n, Z_n) = E(T_n Z_n) - E(T_n)E(Z_n) = N + Nd_{n+1}(N) - (N - d_n(N))(1 + d_n(N))$

$$\rho_n = \frac{Cov(T_n, Z_n)}{\sqrt{V(T_n)V(Z_n)}} = \frac{Cov(T_n, Z_n)}{V(T_n)} = \frac{Nd_{n+1}(N) + d_n(N) - Nd_n(N) + d_n^2(N)}{(2N-1)d_n(N) - 2Nd_{n+1}(N) - d_n^2(N)}$$

On a vu que $V(T_n) \sim d_n(N)$, on aura donc $\lim_{n \rightarrow +\infty} \rho_n = \lim_{n \rightarrow +\infty} \left(N \frac{d_{n+1}(N)}{d_n(N)} + 1 - N + d_n(N) \right)$.

Or $\lim_{n \rightarrow +\infty} \frac{d_{n+1}(N)}{d_n(N)} = \frac{N-1}{N}$ et $\lim_{n \rightarrow +\infty} d_n(N) = 0$, donc finalement $\lim_{n \rightarrow +\infty} \rho_n = 0$

$$9. (a) P_{[T_n=k]}(Z_n = l) = \frac{P(T_n = k, Z_n = l)}{P(T_n = k)}$$

Par les résultats de la question 7 et le résultat de la question 3 on a immédiatement :

$$P_{[T_n=k]}(Z_n = l) = \begin{cases} 0 & \text{si } k < l \\ \frac{1}{k^n - (k-1)^n} & \text{si } k = l \\ \frac{(k-l-1)^n + (k-l+1)^n - 2(k-l)^n}{k^n - (k-1)^n} & \text{si } k > l \end{cases}$$

(b) De ce qui précède, on tire : $E(Z_n/[T_n = k]) = \sum_{l=1}^N l P_{[T_n=k]}(Z_n = l)$ On aura donc

$$E(Z_n/[T_n = k]) = \frac{k}{k^n - (k-1)^n} + \frac{1}{k^n - (k-1)^n} \sum_{l=1}^{k-1} l((k-l-1)^n + (k-l+1)^n - 2(k-l)^n)$$

On pose $j = k - l$ et on a :

$$E(Z_n/[T_n = k]) = \frac{k}{k^n - (k-1)^n} + \frac{1}{k^n - (k-1)^n} \sum_{j=1}^{k-1} (k-j)((j+1)^n - 2j^n + (j-1)^n)$$

Et en utilisant les résultats donnés à la question 8, on obtient :

$$E(Z_n/[T_n = k]) = \frac{k}{k^n - (k-1)^n} + \frac{1}{k^n - (k-1)^n} (k(k^n - (k-1)^n - 1) - ((k-1)k^n - k(k-1)^n))$$

Après simplification, on a : $E(Z_n/[T_n = k]) = \frac{k^n}{k^n - (k-1)^n}$

Partie III. Prévion

Pour $t = (t_1, \dots, t_N) \in \mathbf{R}^N$, on pose $W_t(T_n) = \sum_{k=1}^N t_k \mathbf{1}_{[T_n=k]}$.

10. $([T_n = i])_{1 \leq i \leq N}$ est un système complet d'évènements, on en déduit par la formule des proba-

bilités totales que : $P(W_t(T_n) = t_k) = \sum_{i=1}^N P([W_t(T_n) = t_k] \cap [T_n = i]) = \sum_{i=1}^N P(T_n = i) P_{[T_n=i]}(W_t(T_n) = t_k)$

Or pour i et j dans $\{1, \dots, N\}$ avec $i \neq j$, $[T_n = k] \cap [T_n = i] = \emptyset$ et donc $\forall \omega \in [T_n = i]$,

$W_t(T_n)(\omega) = \sum_{j=1}^N t_j \mathbf{1}_{[T_n=j]}(\omega) = t_i$, ce qui signifie que si $[T_n = i]$ est réalisé avec $i \neq k$ alors

$[W_t(T_n) = t_k]$ est réalisé et donc : $P_{[T_n=i]}(W_t(T_n) = t_k) = \begin{cases} 0 & \text{si } t_i \neq t_k \\ 1 & \text{si } t_i = t_k \end{cases}$

On va ici supposer que si $i \neq k$ alors $t_i \neq t_k$ pour obtenir l'égalité demandée sinon cela ne fonctionne pas.

On aura donc $P(W_t(T_n) = t_k) = P(T_n = k).1 + \sum_{i \neq k, i=1}^N P(T_n = i).0 = P(T_n = k)$

11. Par la formule de l'espérance totale, puisque $([T_n = i])_{1 \leq i \leq N}$ est un système complet d'évènements et que $X_{n+1} = T_{n+1} \mathbf{1}_{[T_n=k]}$ est une variable aléatoire discrète finie avec $X_{n+1}(\Omega) \subset$

$\{0, 1, \dots, N\}$, on a : $E(X_{n+1}) = E(T_{n+1} \mathbf{1}_{[T_n=k]}) = \sum_{i=1}^N E(T_{n+1} \mathbf{1}_{[T_n=k]}/[T_n = i]).P(T_n = i)$

- Pour i dans $\{1, \dots, N\}$ si $i \neq k$ alors

$$E(X_{n+1}/[T_n = i]) = \sum_{x \in X_{n+1}(\Omega)} x P_{[T_n=i]}(X_{n+1} = x) = \sum_{x=0}^N x P_{[T_n=i]}(T_{n+1} \mathbf{1}_{[T_n=k]} = x)$$

D'après le système complet d'évènements, $\forall \omega \in [T_n = i]$, $\mathbf{1}_{[T_n=k]}(\omega) = 0$ alors

$$P_{[T_n=i]}(T_{n+1} \mathbf{1}_{[T_n=k]} = x) = \begin{cases} 0 & \text{si } x \neq 0 \\ 1 & \text{si } x = 0 \end{cases}$$

ce qui entraîne $\forall i \neq k$, $E(X_{n+1}/[T_n = i]) = 0$

- Si $i = k$ alors $E(X_{n+1}/[T_n = k]) = \sum_{x=0}^N x P_{[T_n=k]}(X_{n+1} = x) = \sum_{x=1}^N x P_{[T_n=k]}(X_{n+1} = x)$ et

$\forall \omega \in [T_n = k]$, $X_{n+1}(\omega) = T_{n+1}(\omega)$, donc

$$E(X_{n+1}/[T_n = k]) = \sum_{x=1}^N x P_{[T_n=k]}(T_{n+1} = x) = E(T_{n+1}/[T_n = k])$$

et donc finalement

$$E(T_{n+1} \mathbf{1}_{[T_n=k]}) = P(T_n = k) E(X_{n+1}/[T_n = k]) + 0 = P(T_n = k) E(T_{n+1}/[T_n = k])$$

12. (a) Soit $(k, j) \in \{1, \dots, N\}^2$, $T_{n+1} = \text{Sup}(T_n, U_{n+1}) = \text{Max}(T_n, U_{n+1})$, on a donc :

- Si $j < k$ alors $P([T_n = k] \cap [T_{n+1} = j]) = 0$,
- Si $j > k$ alors par indépendance des variables $U_1, \dots, U_n, U_{n+1}$:
 $P([T_n = k] \cap [T_{n+1} = j]) = P([T_n = k] \cap [U_{n+1} = j]) = P(T_n = k) P(U_{n+1} = j)$
 $P([T_n = k] \cap [T_{n+1} = j]) = \frac{1}{N} P(T_n = k) = \frac{k^n - (k-1)^n}{N^{n+1}}$
- Si $j = k$ alors
 $P([T_n = k] \cap [T_{n+1} = j]) = P([T_n = k] \cap [U_{n+1} \leq k]) P(T_n = k) P(U_{n+1} \leq k)$
 $P([T_n = k] \cap [T_{n+1} = j]) = P(T_n = k) \cdot \frac{k}{N} = \frac{k}{N^{n+1}} (k^n - (k-1)^n)$

(b) On en déduit que

$$P_{[T_n=k]}(T_{n+1} = j) = \frac{P([T_n = k] \cap [T_{n+1} = j])}{P(T_n = k)} = \begin{cases} 0 & \text{si } j < k \\ \frac{k}{N} & \text{si } j = k \\ \frac{1}{N} & \text{si } j > k \end{cases}$$

(c) On en déduit alors :

$$E(T_{n+1}/[T_n = k]) = \sum_{j=1}^N j P_{[T_n=k]}(T_{n+1} = j) = \frac{k^2}{N} + \frac{1}{N} \sum_{j=k+1}^N j = \frac{k^2}{N} + \frac{(N-k)(k+1+N)}{2N}$$

$$E(T_{n+1}/[T_n = k]) = \frac{N+1}{2} + \frac{k^2}{2N} - \frac{k}{2N}$$

(d) Par la formule de l'espérance totale, on obtient :

$$E(T_{n+1}) = \sum_{k=1}^N P(T_n = k) \cdot E(T_{n+1}/[T_n = k])$$

En utilisant le résultat de la question précédente, on peut décomposer en :

$$E(T_{n+1}) = \frac{N+1}{2} \sum_{k=1}^N P(T_n = k) + \frac{1}{2N} \sum_{k=1}^N k^2 P(T_n = k) - \frac{1}{2N} \sum_{k=1}^N k P(T_n = k)$$

$$\text{Ce qui donne : } E(T_{n+1}) = \frac{N+1}{2} + \frac{1}{2N} (E(T_n^2) - E(T_n))$$

13. Par développement d'une somme de termes élevé au carré, on a :

$$(W_t(T_n))^2 = \left(\sum_{k=1}^N t_k \mathbf{1}_{[T_n=k]} \right)^2 = \sum_{k=1}^N t_k^2 (\mathbf{1}_{[T_n=k]})^2 + 2 \sum_{1 \leq i < j \leq N} t_i t_j \mathbf{1}_{[T_n=i]} \mathbf{1}_{[T_n=j]}$$

Or $(\mathbf{1}_{[T_n=k]})^2 = \mathbf{1}_{[T_n=k]} \cdot \mathbf{1}_{[T_n=k]} = \mathbf{1}_{[T_n=k] \cap [T_n=k]} = \mathbf{1}_{[T_n=k]}$ et pour $i < j$ $\mathbf{1}_{[T_n=i]} \cdot \mathbf{1}_{[T_n=j]} = \mathbf{1}_{[T_n=i] \cap [T_n=j]} = 0$ on rappelle que $([T_n = k])_{1 \leq k \leq N}$ est un système complet d'événements, les événements sont donc deux à deux incompatibles.

$$\text{Finalement } (W_t(T_n))^2 = \sum_{k=1}^N t_k^2 \mathbf{1}_{[T_n=k]}$$

14. Soit g telle que $\forall (t_1, \dots, t_N) \in \mathbf{R}^N$ $g(t_1, \dots, t_N) = E((T_{n+1} - W_t(T_n))^2)$

(a) En appliquant la formule de l'espérance totale avec le système complet $([T_n = k])_{1 \leq k \leq N}$,

$$g(t) = \sum_{k=1}^N E((T_{n+1} - W_t(T_n))^2 | [T_n = k]) P(T_n = k) = \sum_{k=1}^N E((T_{n+1}^2 - 2\theta_k t_k + t_k^2) | [T_n = k]) P(T_n = k)$$

On obtient la formule voulue en remarquant que :

$$T_{n+1}^2 - 2\theta_k t_k + t_k^2 = (t_k - \theta_k)^2 + E(T_{n+1}^2) - \theta_k^2$$

$$\text{Alors : } g(t) = \sum_{k=1}^N (t_k - \theta_k)^2 P(T_n = k) + \sum_{k=1}^N (E(T_{n+1}^2) - \theta_k^2) P(T_n = k)$$

(b) Comme la deuxième somme est constante (indépendante de t), le minimum de g est atteint quand la première somme est minimum. Or cette somme est positive, et nulle uniquement pour $\forall k, t_k = \theta_k$, donc pour $t = \theta$.

Le minimum global de g sur $\mathbf{R}^N$ est atteint au point θ (uniquement).

15. Par linéarité de l'espérance, on a :

$$E(W_\theta(T_n)) = \sum_{k=1}^N \theta_k E(\mathbf{1}_{[T_n=k]}) = \sum_{k=1}^N E(T_{n+1}/[T_n = k])P(T_n = k) = E(T_{n+1})$$

Par la formule de Huygens, on a :

$$V(W_\theta(T_n)) = E(W_\theta^2(T_n)) - (E(W_\theta(T_n)))^2 = E(W_\theta^2(T_n)) - E(T_{n+1}^2)$$

$$V(W_\theta(T_n)) = E(W_\theta^2(T_n)) - E(T_{n+1}^2) + V(T_{n+1})$$

$$\text{Or } g(\theta) = E(T_{n+1}^2) - 2E(T_{n+1}W_\theta(T_n)) + E(W_\theta^2(T_n))$$

et par le résultat de la question 11, on a :

$$E(T_{n+1}W_\theta(T_n)) = \sum_{k=1}^N \theta_k E(T_{n+1}\mathbf{1}_{[T_n=k]}) = \sum_{k=1}^N \theta_k^2 P(T_n = k) = E(W_\theta^2(T_n)) \quad (\text{résultat 13})$$

$$\text{donc } g(\theta) = E(T_{n+1}^2) - E(W_\theta^2(T_n)) \text{ et } V(W_\theta(T_n)) = V(T_{n+1}) - g(\theta),$$

$$\text{mais par définition } g(\theta) = E((T_{n+1} - W_\theta(T_n))^2) \geq 0, \text{ donc } V(W_\theta(T_n)) \leq V(T_{n+1})$$

16. (a) Soit $\omega \in \Omega$, puisque $([T_n = j])_{1 \leq j \leq N}$ est un système complet d'évènements, il existe un unique $j \in \{1, \dots, N\}$ tel que $T_n(\omega) = j$. On en déduit que si $k \neq j$ alors $\mathbf{1}_{[T_n=k]}(\omega) = 0$,

$$\text{et donc } \sum_{k=1}^N k^i \mathbf{1}_{[T_n=k]}(\omega) = j^i \mathbf{1}_{[T_n=j]}(\omega) = j^i = T_n^i(\omega).$$

$$\text{Finalement : } \forall \omega \in \Omega, \quad T_n^i(\omega) = \sum_{k=1}^N k^i \mathbf{1}_{[T_n=k]}(\omega), \text{ donc } \sum_{k=1}^N k^i \mathbf{1}_{[T_n=k]} = T_n^i.$$

$$(b) \text{ Pour } k \in \{1, \dots, N\}, \theta_k = E(T_{n+1}/[T_n = k]) = \frac{N+1}{2} + \frac{k^2}{2N} - \frac{k}{2N}.$$

$$\text{On en déduit } W_\theta(T_n) = \sum_{k=1}^N \theta_k \mathbf{1}_{[T_n=k]} = \frac{N+1}{2} \sum_{k=1}^N \mathbf{1}_{[T_n=k]} + \frac{1}{2N} \sum_{k=1}^N k^2 \mathbf{1}_{[T_n=k]} - \frac{1}{2N} \sum_{k=1}^N k \mathbf{1}_{[T_n=k]}.$$

$$\text{Ce qui donne avec le résultat précédent : } W_\theta(T_n) = \frac{N+1}{2} + \frac{1}{2N}(T_n^2 - T_n).$$

Partie IV Estimation

17. (a) D'après 4(b), $E(T_n)$ tend vers N lorsque n tend vers $+\infty$.
 (b) D'après 4(c), $V(T_n)$ tend vers 0 lorsque n tend vers $+\infty$.
 (c) Lorsque n devient grand, T_n se comporte quasiment comme une variable d'espérance N et de variance nulle, soit une variable constante égale à N : lorsque n devient grand, T_n prend une valeur proche de N , donc fournit une estimation de N .
18. (a) Par indépendance des variables $U_1, \dots, U_n$, et d'après la loi U , on a :

$$\forall (u_1, \dots, u_n) \in \{1, \dots, N\}^N \quad P\left(\bigcap_{i=1}^n [U_i = u_i]\right) = \prod_{i=1}^n P(U_i = u_i) = \frac{1}{N^n}.$$

- (b) Par définition d'une probabilité conditionnelle, on a :

$$P_{[T_n=k]} \left(\bigcap_{i=1}^n [U_i = u_i] \right) = \frac{P\left(\left(\bigcap_{i=1}^n [U_i = u_i]\right) \cap [T_n = k]\right)}{P(T_n = k)}.$$

Or si $u_i \notin \{1, \dots, N\}$ alors $[U_i = u_i] = \emptyset$ et par la définition de T_n si $\max(u_i) \neq k$ alors

$$\bigcap_{i=1}^n [U_i = u_i] \cap [T_n = k] = \emptyset.$$

$$\text{Si } \forall i \in \{1, \dots, n\} \text{ et } \max(u_i) = k \text{ alors } \bigcap_{i=1}^n [U_i = u_i] \cap [T_n = k] = \bigcap_{i=1}^n [U_i = u_i] \text{ d'où}$$

$$P_{[T_n=k]} \left(\bigcap_{i=1}^n [U_i = u_i] \right) = \begin{cases} \frac{P \left(\bigcap_{i=1}^n [U_i = u_i] \right)}{P(T_n = k)} = \frac{1}{k^n - (k-1)^n} & \text{si } 1 \leq u_i \leq N \text{ et } \max(u_i) = k \\ 0 & \text{sinon} \end{cases}$$

19. On pose $S_n = T_n + Z_n - 1$ et $\psi_n(k) = \frac{k^{n+1} - (k-1)^{n+1}}{k^n - (k-1)^n}$.

(a) S_n est une variable aléatoire fonction, indépendante de N , du n -échantillon iid $(U_1, \dots, U_n)$ de la loi U . De plus $E(S_n) = E(T_n) + E(Z_n) - 1 = N - d_n(N) + 1 + d_n(N) - 1 = N$.
On en déduit que S_n est un estimateur sans biais de N .

(b) S_n est une variable aléatoire discrète finie alors pour $k \in \{1, \dots, N\}$

$$E(S_n/[T_n = k]) = E(T_n/[T_n = k]) + E(Z_n/[T_n = k]) - 1.$$

On a calculé $E(Z_n/[T_n = k])$ dans la question 9b) et $E(T_n/[T_n = k]) = \sum_{i=1}^N iP_{[T_n=k]}(T_n = i) = k$,

d'où

$$E(S_n/[T_n = k]) = k + \frac{k^n}{k^n - (k-1)^n} - 1 = \frac{k^{n+1} - k(k-1)^n + k^n - k^n + (k-1)^n}{k^n - (k-1)^n}$$

$$E(S_n/[T_n = k]) = \frac{k^{n+1} - (k-1)^n(k-1)}{k^n - (k-1)^n} = \psi_n(k)$$

(c) $\psi_n(T_n)$ est une variable aléatoire ne dépendant pas de N , qui, par le théorème de transfert ($T_n(\Omega)$ est inclus dans le domaine de définition de ψ_n), admet une espérance et

$$E(\psi_n(T_n)) = \sum_{k=1}^N \psi_n(k)P(T_n = k) = \sum_{k=1}^N \left(\frac{k^{n+1} - (k-1)^{n+1}}{N^n} \right)$$

Par télescopage, il reste : $E(\psi_n(T_n)) = \frac{N^{n+1}}{N^n} = N$

(d) On pose $\varphi_n(k) = E(S_n^2/[T_n = k])$ pour k entier compris entre 1 et N .

Soit la fonction $f : \lambda \mapsto E((S_n - \lambda)^2/[T_n = k])$, il est clair que f est à valeurs positives et par linéarité de l'espérance on a : $f(\lambda) = \varphi_n(k) - 2\lambda\psi_n(k) + \lambda^2$

donc en particulier $0 \leq f(\psi_n(k))$ et $f(\psi_n(k)) = \varphi_n(k) - \psi_n^2(k)$ donc $\psi_n^2(k) \leq \varphi_n(k)$.

On en déduit, par produit par un terme positif et par somme, que :

$$\sum_{k=1}^N \psi_n^2(k)P(T_n = k) \leq \sum_{k=1}^N \varphi_n(k)P(T_n = k) \text{ Ce qui donne, par le théorème de transfert}$$

et la formule de l'espérance totale : $E(\psi_n^2(T_n)) \leq E(S_n^2)$.

Or $E(\psi_n(T_n)) = N = E(S_n)$ alors par la formule de Huygens, on a : $V(\psi_n(T_n)) \leq V(S_n)$.

(e) Par formule sur la variance d'une somme de variables aléatoires, on a :

$$V(S_n) = V(T_n) + V(Z_n) + 2Cov(Z_n, T_n)$$

On a vu $V(Z_n) = V(T_n)$ et on a calculé $Cov(T_n, Z_n)$ dans la question 8b), on a donc

$$V(S_n) = 2((2N-1)d_n(N) - 2Nd_{n+1}(N) - d_n^2(N) + Nd_{n+1}(N) + d_n(N) - Nd_n(N + d_n^2(N)))$$

$$V(S_n) = Nd_n(N) - Nd_{n+1}(N) = N(d_n(N) - d_{n+1}(N))$$

On en déduit que $\lim_{n \rightarrow +\infty} V(S_n) = 0$ et puisque $0 \leq V(\psi_n(T_n)) \leq V(S_n)$, par théorème d'encadrement, on a aussi $\lim_{n \rightarrow +\infty} V(\psi_n(T_n)) = 0$.

$(\psi_n(T_n))$ est une suite d'estimateurs de N sans biais dont la variance tend vers 0 lorsque n tend vers $+\infty$, on sait alors (par inégalité de Bienaymé-Tchebychev :

$$\forall \varepsilon > 0, \quad P(|\psi_n(T_n) - N| \geq \varepsilon) \leq \frac{V(\psi_n(T_n))}{\varepsilon^2}, \text{ que la suite d'estimateurs de } N (\psi_n(T_n))$$

est convergente.

20. R_n est un estimateur sans biais du paramètre N . On pose pour k entier compris entre 1 et N : $f_n(k) = E(R_n/[T_n = k])$.

(a) On va supposer ici que R_n admet une variance, alors en utilisant la fonction positive $\lambda \mapsto E((R_n - \lambda)^2/[T_n = k])$ comme fait en question précédente, on trouve :

$$\forall \lambda \in \mathbf{R}, \quad 0 \leq E((R_n - \lambda)^2/[T_n = k]) = \lambda^2 - 2\lambda f_n(k) + E(R_n^2/[T_n = k])$$

et pour $\lambda = f_n(k)$, on a :

$$\forall k \in \{1, \dots, N\}, \quad 0 \leq -f_n(k)^2 + E(R_n^2/[T_n = k]).$$

Par produit par $P(T_n = k) \geq 0$ puis par somme pour k allant de 1 à N , on obtient avec

$$\text{la formule de l'espérance totale : } 0 \leq - \sum_{k=1}^N f_n(k)^2 P(T_n = k) + E(R_n^2),$$

$$\text{et par le théorème de transfert : } 0 \leq E(R_n^2) - E(f_n^2(T_n)) \quad (*).$$

Toujours par le théorème de transfert et la formule de l'espérance totale, on a

$$E(f_n(T_n)) = \sum_{k=1}^N f_n(k) P(T_n = k) = E(R_n) = N, \text{ d'où finalement par la formule de Huygens et } (*), \text{ on a : } 0 \leq V(f_n(T_n)) - V(R_n).$$

(b) On suppose que pour tout $N \in \mathbb{N}^*$, $E(F(T_n)) = N = \sum_{k=1}^N F(k) P(T_n = k)$. Montrons par

récurrence sur $k \in \mathbb{N}^*$, que $F(k) = \psi_n(k)$.

On rappelle que $E(\psi_n(T_n)) = N$.

$$\text{En prenant } N = 1, \quad N = E(F(T_n)) = \sum_{k=1}^N F(k) P(T_n = k) = F(1) \text{ et}$$

$$E(\psi_n(T_n)) = N = \sum_{k=1}^N \psi_n(k) P(T_n = k) = \psi_n(1), \text{ donc } F(1) = \psi_n(1).$$

Supposons que pour k fixé dans $\mathbb{N}^*$, on ait $\forall i \in \{1, \dots, k\} \quad F(i) = \psi_n(i)$.

En prenant $N = k + 1$, on aura

$$N = E(F(T_n)) = \sum_{i=1}^N F(i) P(T_n = i) = \sum_{i=1}^k \psi_n(i) P(T_n = i) + F(k+1) P(T_n = k+1),$$

et par application de l'hypothèse de récurrence faite ci-dessus, on a :

$$N = E(F(T_n)) = E(\psi_n(T_n)) - \psi_n(k+1) P(T_n = k+1) + F(k+1) P(T_n = k+1),$$

$$N = N + P(T_n = N) (F(k+1) - \psi_n(k+1)).$$

Or $P(T_n = N) \neq 0$, donc $F(k+1) = \psi_n(k+1)$. ce qui termine la récurrence :

On a ainsi : $\forall k \in \mathbb{N}^*, \quad F(k) = \psi_n(k)$.

Ce qui donne bien sûr : $\forall N \in \mathbb{N}^*, \quad \forall k \in \{1, \dots, N\}, \quad F(k) = \psi_n(k)$.

(c) Soit R_n une variable d'espérance N , par le résultat 20a) et 20b), on a $f_n(T_n)$ a pour espérance N et donc $\forall k \in \{1, \dots, N\}, \quad f_n(k) = \psi_n(k)$. ce qui entraîne :

$\forall \omega \in \Omega, \quad f_n(T_n(\omega)) = \psi_n(T_n(\omega))$ ce qui signifie que $f_n(T_n) = \psi_n(T_n)$ et on a :

$$V(f_n(T_n)) \leq V(R_n) \quad \text{et} \quad V(f_n(T_n)) = V(\psi_n(T_n)) \text{ donc } V(\psi_n(T_n)) \leq V(R_n).$$

$\psi_n(T_n)$ est donc une variable d'espérance N optimale, on dit c'est un « estimateur sans biais de N optimal » (de variance minimale).