



PRÉPASIMPLE

prepasimple.fr

HEC — 2009

Corrigé de mathématiques approfondies

Mathématiques approfondies — Prépa ECG

Annales de concours · Corrigé

Préliminaire

Tous les calculs de moments ci-dessous portent sur des variables prenant un nombre fini de valeurs. Leurs espérances, moments d'ordre deux et variances existent donc. Les sommes échangées dans cette correction sont toutes finies.

1. Formules de moments à partir des probabilités de dépassement

Méthode 1 — Décomposition en indicatrices

Pour tout entier $y \in \llbracket 1, N \rrbracket$, exactement y entiers de $\llbracket 0, N - 1 \rrbracket$ vérifient $k < y$. Ainsi :

$$y = \sum_{k=0}^{N-1} \mathbf{1}_{[y > k]}.$$

De plus, la somme des y premiers entiers impairs vaut y^2 , car :

$$\sum_{k=0}^{y-1} (2k+1) = 2 \frac{y(y-1)}{2} + y = y^2.$$

En appliquant ces identités à $Y(\omega)$, pour chaque ω , on obtient :

$$Y = \sum_{k=0}^{N-1} \mathbf{1}_{[Y > k]}, \quad Y^2 = \sum_{k=0}^{N-1} (2k+1) \mathbf{1}_{[Y > k]}.$$

La linéarité de l'espérance pour une somme finie et $E(\mathbf{1}_C) = P(C)$ donnent les deux relations demandées :

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

Méthode 2 — Échange de deux sommes finies

Posons $p_j = P(Y = j)$ pour $1 \leq j \leq N$. On a

$$P(Y > k) = \sum_{j=k+1}^N p_j.$$

Par échange de sommes finies,

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

De même,

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

La première méthode donne une identité entre variables aléatoires ; la seconde travaille directement sur leur loi. Elles sont utiles respectivement pour combiner des indicatrices et pour transformer une somme de probabilités.

Partie I. Inf et Sup

2. Moments d'une uniforme discrète

La variable U_1 suit la loi uniforme sur $\llbracket 1, N \rrbracket$. Les formules du cours, demandées ici sans démonstration, sont :

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

Pour $N = 1$, U_1 est constante égale à 1, et ces expressions donnent bien une espérance égale à 1 et une variance nulle.

3. (a) Répartition du maximum

Fixons $k \in \llbracket 1, N \rrbracket$. Le maximum est inférieur ou égal à k si et seulement si chacun des éléments de l'échantillon l'est :

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

Les U_i sont mutuellement indépendantes, donc les événements de cette intersection sont indépendants. Comme $P(U_i \leq k) = k/N$, on obtient :

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

On utilisera également $P(T_n \leq 0) = 0$, puisque $T_n \geq 1$.

3. (b) Loi du maximum

Le support de T_n est $\llbracket 1, N \rrbracket$. Pour k dans cet ensemble, les événements $[T_n \leq k-1]$ et $[T_n = k]$ forment une partition de $[T_n \leq k]$. Par conséquent :

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

Pour $k \notin \llbracket 1, N \rrbracket$, cette probabilité est nulle. La formule vaut aussi pour $k = 1$, car $n \geq 1$ et $0^n = 0$.

Les probabilités sont strictement positives et leur somme vaut 1 par télescopage. Pour $N = 1$, on retrouve la loi certaine égale à 1.

4. (a) Limite de $d_n(N)$

Méthode 1 — Somme finie de suites géométriques

Si $N = 1$, $d_n(1) = 0$ pour tout n , donc la limite est nulle.

Si $N \geq 2$, pour chacun des indices $1 \leq k \leq N-1$, on a $0 < k/N < 1$. La suite géométrique $(k/N)^n$ tend donc vers 0. Le nombre $N-1$ de termes étant fixe et fini, on

peut additionner leurs limites :

$$d_n(N) = \sum_{k=1}^{N-1} \left(\frac{k}{N}\right)^n \rightarrow 0.$$

Méthode 2 — Encadrement géométrique

Pour $N \geq 2$, posons $q = (N-1)/N \in]0, 1[$. Chaque terme de la somme est majoré par q^n , donc :

$$0 \leq d_n(N) \leq (N-1)q^n.$$

La borne supérieure tend vers 0 ; le théorème d'encadrement donne la même limite.

Dans tous les cas, $d_n(N) \rightarrow 0$. La première approche utilise le caractère fini de la somme ; la seconde fournit une vitesse de décroissance géométrique.

4. (b) Espérance du maximum

La formule de la question 1 s'applique à T_n , à valeurs dans $\llbracket 1, N \rrbracket$. Avec 3(a),

$$E(T_n) = \sum_{k=0}^{N-1} (1 - P(T_n \leq k)) = N - \sum_{k=1}^{N-1} \left(\frac{k}{N}\right)^n.$$

La somme est vide lorsque $N = 1$, conformément à la définition de d_n . Ainsi :

$$E(T_n) = N - d_n(N).$$

La question 4(a) entraîne alors :

$$\lim_{n \rightarrow +\infty} E(T_n) = N.$$

4. (c) Variance du maximum

Appliquons la seconde relation de 1 à T_n :

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

La somme des N premiers entiers impairs vaut N^2 , comme établi en 1. De plus,

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

On obtient donc :

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

En utilisant la formule de Koenig-Huygens et 4(b),

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

Le développement du carré donne :

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

Par 4(a), $d_n(N) \rightarrow 0$ et $d_{n+1}(N) \rightarrow 0$. Tous les termes du membre de droite tendent vers 0, donc :

$$\lim_{n \rightarrow +\infty} V(T_n) = 0.$$

Pour $N = 1$, la formule donne une variance identiquement nulle, comme attendu.

4. (d) Quotient de deux sommes et équivalent de la variance

Supposons $N \geq 2$ et posons $q = (N - 1)/N$. Le terme correspondant à $k = N - 1$

domine la somme. En le factorisant exactement,

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

La somme entre parenthèses après le 1 tend vers 0, puisque ses bases géométriques sont strictement inférieures à 1 et que leur nombre est fixe. Pour $N = 2$, elle est vide, donc déjà nulle.

Ainsi, $d_n(N) \sim q^n$ et :

$$\frac{d_{n+1}(N)}{d_n(N)} = q \frac{1 + \sum_{k=1}^{N-2} (k/(N-1))^{n+1}}{1 + \sum_{k=1}^{N-2} (k/(N-1))^n} \rightarrow q = 1 - \frac{1}{N}.$$

Comme $d_n(N) > 0$ pour $N \geq 2$, on peut diviser la formule de 4(c) par $d_n(N)$:

$$\frac{V(T_n)}{d_n(N)} = (2N-1) - 2N \frac{d_{n+1}(N)}{d_n(N)} - d_n(N).$$

Sa limite vaut $(2N-1) - 2N(1-1/N) = 1$. Par conséquent :

$$V(T_n) \sim d_n(N) \sim \left(1 - \frac{1}{N} \right)^n.$$

L'hypothèse $N \geq 2$ est indispensable pour cet équivalent : si $N = 1$, la variance et $d_n(N)$ sont toutes deux nulles, et leur quotient n'est pas défini.

5. Loi et moments du minimum

Méthode 1 — Probabilités de dépassement

Pour $k \in \llbracket 0, N \rrbracket$,

$$[Z_n > k] = \bigcap_{i=1}^n [U_i > k].$$

Par indépendance et uniformité,

$$P(Z_n > k) = \left(\frac{N-k}{N}\right)^n.$$

En prenant la différence entre les probabilités $Z_n > \ell - 1$ et $Z_n > \ell$, on obtient, pour $1 \leq \ell \leq N$:

$$P(Z_n = \ell) = \left(\frac{N-\ell+1}{N}\right)^n - \left(\frac{N-\ell}{N}\right)^n.$$

En dehors de $\llbracket 1, N \rrbracket$, la probabilité est nulle. La question 1 donne :

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

Pour le second moment, effectuons le changement d'indice $j = N - k$ dans la somme finie :

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

Le terme $j = N$ vaut 1, et les autres termes donnent :

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

En soustrayant $(1 + d_n(N))^2$, on obtient :

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

Méthode 2 — Réflexion de l'échantillon

Posons $V_i = N + 1 - U_i$. Chaque V_i suit encore la loi uniforme sur $\llbracket 1, N \rrbracket$; les V_i sont

mutuellement indépendantes par transformation individuelle des variables indépendantes.

De plus, pour chaque réalisation,

$$\max(V_1, \dots, V_n) = N + 1 - Z_n.$$

Le maximum des V_i a la même loi que T_n . Ainsi, Z_n a la même loi que $N + 1 - T_n$. La formule de 3(b), avec l'indice $N + 1 - \ell$, fournit donc la loi obtenue ci-dessus.

Pour les moments, la linéarité de l'espérance et la formule de variance d'une transformation affine donnent directement :

$$E(Z_n) = N + 1 - E(T_n) = 1 + d_n(N),$$

$$V(Z_n) = V(T_n).$$

La première méthode décrit directement le minimum ; la seconde réutilise les résultats sur le maximum par symétrie. Pour $N = 1$, $Z_n = T_n = 1$ et les deux méthodes donnent une espérance égale à 1 et une variance nulle.

6. Simulation du maximum

Code adapté en Python. Question adaptée au programme actuel : la fonction reçoit explicitement les deux entiers $n \geq 1$ et $N \geq 1$.

Python

```
import numpy.random as rd

def simulmax(n, N):
    T = 1
    for i in range(n):
        U = rd.randint(1, N + 1)
        if U > T:
            T = U
    return T
```

La borne supérieure de `rd.randint` est exclue ; $N + 1$ assure donc un tirage uniforme sur $\llbracket 1, N \rrbracket$. La boucle effectue exactement n tirages indépendants. L'initialisation $T = 1$ est valide, puisque chaque tirage est au moins égal à 1 ; après chaque itération, T contient le maximum des valeurs déjà simulées.

En sortie, la fonction renvoie donc une réalisation de T_n . Pour $N = 1$, elle renvoie toujours 1.

Le code a été exécuté. Un contrôle sur toutes les suites possibles de petits échantillons a vérifié la valeur du maximum renvoyé ; une simulation répétée a également été comparée à la loi exacte de 3(b). Ces vérifications portent sur l'implémentation et ne remplacent pas la preuve de la loi.

Partie II. Couple (Inf, Sup)

7. (a) Fonction de répartition jointe

Si $k \leq \ell$, $T_n \leq k$ implique $Z_n \leq T_n \leq \ell$. Ainsi,

$$\phi_n(k, \ell) = P(T_n \leq k) = \left(\frac{k}{N}\right)^n.$$

Si $k > \ell$, on retire de $[T_n \leq k]$ le sous-événement $[T_n \leq k] \cap [Z_n > \ell]$. Ce dernier signifie que tous les U_i appartiennent à $\llbracket \ell + 1, k \rrbracket$, qui contient $k - \ell$ entiers. Par indépendance,

$$P(T_n \leq k, Z_n > \ell) = \left(\frac{k - \ell}{N}\right)^n.$$

D'où :

$$\phi_n(k, \ell) = \left(\frac{k}{N}\right)^n - \left(\frac{k - \ell}{N}\right)^n \quad (k > \ell).$$

Pour les calculs aux bords, on observe aussi que $\phi_n(0, \ell) = \phi_n(k, 0) = 0$, puisque T_n et Z_n sont au moins égaux à 1.

7. (b) Passage de la répartition jointe à la loi jointe

Les variables étant entières,

$$\mathbf{1}_{[T_n = k]} = \mathbf{1}_{[T_n \leq k]} - \mathbf{1}_{[T_n \leq k-1]},$$

$$\mathbf{1}_{[Z_n=\ell]} = \mathbf{1}_{[Z_n \leq \ell]} - \mathbf{1}_{[Z_n \leq \ell-1]}.$$

Le produit des deux indicatrices de gauche est l'indicatrice de l'événement $[T_n = k] \cap [Z_n = \ell]$. En développant le produit de droite, puis en prenant l'espérance, on obtient :

$$P(T_n = k, Z_n = \ell) = \phi_n(k, \ell) + \phi_n(k-1, \ell-1) - \phi_n(k-1, \ell) - \phi_n(k, \ell-1).$$

Cette formule inclut $k = 1$ et $\ell = 1$ grâce aux valeurs nulles précisées à la fin de 7(a).

7. (c) Loi jointe du maximum et du minimum

Méthode 1 — Différences de la fonction de répartition

Si $k < \ell$, le minimum ne peut pas dépasser le maximum. L'événement $[T_n = k, Z_n = \ell]$ est donc vide et sa probabilité vaut 0.

Si $k = \ell$, les quatre termes de 7(b), après multiplication par N^n , valent respectivement k^n , $(k-1)^n$, $(k-1)^n$ et $k^n - 1$. Leur combinaison vaut 1. Ainsi,

$$P(T_n = k, Z_n = k) = \frac{1}{N^n}.$$

Si $k > \ell$, posons $j = k - \ell \geq 1$. La substitution des formules de 7(a) dans 7(b) donne, après multiplication par N^n :

$$(k^n - j^n) + ((k-1)^n - j^n) - ((k-1)^n - (j-1)^n) - (k^n - (j+1)^n).$$

Les termes en k^n et en $(k-1)^n$ s'annulent. Il reste $(j+1)^n - 2j^n + (j-1)^n$.

La loi jointe est donc :

$$P(T_n = k, Z_n = \ell) = \begin{cases} 0 & k < \ell, \\ \frac{1}{N^n} & k = \ell, \\ \frac{(k - \ell + 1)^n - 2(k - \ell)^n + (k - \ell - 1)^n}{N^n} & k > \ell. \end{cases}$$

Méthode 2 — Dénombrement et inclusion-exclusion

Chaque suite de n tirages dans $\llbracket 1, N \rrbracket$ a la probabilité N^{-n} , par indépendance et uniformité.

Pour $k > \ell$, l'événement demandé impose que les valeurs appartiennent à $\llbracket \ell, k \rrbracket$, et que chacune des deux extrémités ℓ et k apparaisse au moins une fois.

Il y a $(k - \ell + 1)^n$ suites dans cet intervalle. Parmi elles, $(k - \ell)^n$ n'utilisent pas ℓ , et autant n'utilisent pas k . Les suites n'utilisant ni l'une ni l'autre ont été retirées deux fois ; il faut réajouter leur nombre $(k - \ell - 1)^n$. La formule d'inclusion-exclusion redonne donc le numérateur ci-dessus.

Pour $k = \ell$, une seule suite convient : tous les tirages valent k . Pour $k < \ell$, aucune suite ne convient. Cela établit de nouveau toute la loi jointe.

La première méthode systématise l'utilisation d'une répartition jointe ; la seconde explique la présence de la différence seconde de puissances.

8. (a) Espérance du produit

Méthode 1 — Utilisation des deux identités fournies

Posons $b_j = (j+1)^n - 2j^n + (j-1)^n$ pour $j \geq 1$. La loi jointe de 7(c) permet d'écrire :

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

Le terme k dans la parenthèse provient du cas diagonal $Z_n = T_n = k$.

Pour $k \geq 2$, le changement d'indice $j = k - \ell$ donne :

$$\sum_{\ell=1}^{k-1} \ell b_{k-\ell} = k \sum_{j=1}^{k-1} b_j - \sum_{j=1}^{k-1} j b_j.$$

Appliquons les deux identités de l'énoncé avec $m = k - 1$:

$$\sum_{j=1}^{k-1} b_j = k^n - (k-1)^n - 1,$$

$$\sum_{j=1}^{k-1} j b_j = (k-1)k^n - k(k-1)^n.$$

La combinaison précédente devient :

$$k[k^n - (k-1)^n - 1] - [(k-1)k^n - k(k-1)^n] = k^n - k.$$

Pour $k = 1$, la somme est vide et vaut $0 = 1^n - 1$. Ainsi, la parenthèse de la formule de transfert vaut k^n pour tout k , d'où :

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

Le terme $k = N$ vaut 1, et les autres forment $d_{n+1}(N)$. Donc :

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

Méthode 2 — Produit de deux décompositions en indicatrices

La question 1 donne les identités

$$T_n = \sum_{a=0}^{N-1} \mathbf{1}_{[T_n > a]}$$

et

$$Z_n = \sum_{b=0}^{N-1} \mathbf{1}_{[Z_n > b]}.$$

Leur produit, suivi de la prise d'espérance, fournit :

$$E(T_n Z_n) = \sum_{a=0}^{N-1} \sum_{b=0}^{N-1} P(T_n > a, Z_n > b).$$

Si $a \leq b$, $Z_n > b$ implique $T_n > a$, et cette probabilité vaut $((N - b)/N)^n$.

Si $a > b$, on retire du même événement les trajectoires dont toutes les valeurs sont dans $\llbracket b + 1, a \rrbracket$. La probabilité vaut alors $((N - b)/N)^n - ((a - b)/N)^n$.

En regroupant les termes, on obtient :

$$E(T_n Z_n) = N \sum_{b=0}^{N-1} \left(\frac{N - b}{N} \right)^n - \sum_{0 \leq b < a \leq N-1} \left(\frac{a - b}{N} \right)^n.$$

Pour chaque différence $j = a - b \in \llbracket 1, N - 1 \rrbracket$, il y a $N - j$ couples possibles. Par conséquent,

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

La dernière somme vaut $Nd_n(N) - Nd_{n+1}(N)$. On retrouve $N(1 + d_{n+1}(N))$.

La première méthode suit les indications de l'énoncé ; la seconde montre comment les probabilités de dépassement permettent aussi de calculer un moment mixte. Les deux calculs incluent $N = 1$, pour lequel le produit vaut constamment 1.

8. (b) Limite du coefficient de corrélation

Supposons $N \geq 2$. Pour chaque $n \geq 1$, le maximum prend au moins deux valeurs avec des probabilités strictement positives, d'après 3(b), donc sa variance est strictement positive. Il en va de même pour le minimum, par 5. Le coefficient de corrélation est donc bien défini.

Pour alléger les calculs, posons $a = d_n(N)$ et $b = d_{n+1}(N)$. Les questions 4(b), 5 et 8(a) donnent :

$$\text{Cov}(T_n, Z_n) = N(1 + b) - (N - a)(1 + a) = Nb - (N - 1)a + a^2.$$

Comme les deux variances sont égales,

$$\rho_n = \frac{\text{Cov}(T_n, Z_n)}{V(T_n)}.$$

La division du numérateur par $a > 0$ donne :

$$\frac{\text{Cov}(T_n, Z_n)}{a} = N\frac{b}{a} - (N - 1) + a \longrightarrow N\left(1 - \frac{1}{N}\right) - (N - 1) = 0.$$

Par 4(d), $V(T_n)/a \rightarrow 1$. Le quotient de ces deux expressions donne donc :

$$\lim_{n \rightarrow +\infty} \rho_n = 0.$$

On ne conclut pas à une indépendance pour un n fixé. Pour $N = 1$, les variances sont nulles et le coefficient de corrélation n'est pas défini, ce qui explique l'hypothèse de la question.

9. (a) Loi conditionnelle du minimum sachant le maximum

Pour $1 \leq k \leq N$, posons $A_k = k^n - (k-1)^n > 0$. Par 3(b), $P(T_n = k) = A_k/N^n > 0$; la probabilité conditionnelle est donc définie.

En divisant les probabilités conjointes de 7(c) par cette probabilité, on obtient, pour $1 \leq \ell \leq N$:

$$P_{[T_n=k]}(Z_n = \ell) = \begin{cases} 0 & \ell > k, \\ \frac{1}{A_k} & \ell = k, \\ \frac{(k-\ell+1)^n - 2(k-\ell)^n + (k-\ell-1)^n}{A_k} & 1 \leq \ell < k. \end{cases}$$

Cette loi ne dépend pas de N , une fois k et n fixés. Pour $k = 1$, elle est certaine égale à 1.

9. (b) Espérance conditionnelle du minimum

Méthode 1 — Somme de la loi conditionnelle

En utilisant 9(a) et la notation b_j de 8(a),

$$E(Z_n \mid T_n = k) = \frac{k + \sum_{\ell=1}^{k-1} \ell b_{k-\ell}}{A_k}.$$

La simplification effectuée en 8(a), valable aussi pour $k = 1$, montre que le numérateur

est k^n . Ainsi :

$$E(Z_n \mid T_n = k) = \frac{k^n}{k^n - (k-1)^n}.$$

Méthode 2 — Probabilités conditionnelles de dépassement

Pour $j \in \llbracket 0, k-1 \rrbracket$, l'événement $[Z_n > j, T_n = k]$ impose que tous les tirages soient dans $\llbracket j+1, k \rrbracket$, avec au moins un tirage égal à k . Le nombre de suites correspondantes est $(k-j)^n - (k-j-1)^n$.

En divisant sa probabilité par $P(T_n = k)$, on obtient :

$$P_{[T_n=k]}(Z_n > j) = \frac{(k-j)^n - (k-j-1)^n}{A_k}.$$

La formule de la question 1 s'applique à la loi conditionnelle, qui est une loi finie sur $\llbracket 1, k \rrbracket$. En sommant sur $j = 0, \dots, k-1$, le numérateur est télescopique et vaut k^n . On retrouve donc la même espérance.

La première méthode réutilise les identités fournies ; la seconde transforme le calcul de moyenne en une simple somme télescopique de probabilités de dépassement.

Partie III. Préviation

Pour toute cette partie, fixons $n \geq 1$ et posons, pour $1 \leq k \leq N$:

$$p_k = P(T_n = k) = \frac{k^n - (k-1)^n}{N^n} > 0.$$

Le symbole p_k désigne ici les probabilités de la loi du maximum, pour les paramètres n et N fixés.

10. Loi de la variable de préviation

Question adaptée au programme actuel : rectification d'une erreur de l'énoncé original. L'égalité $P(W_t(T_n) = t_k) = P(T_n = k)$ n'est vraie en général que si la valeur t_k n'est partagée par aucun autre coefficient.

Par exemple, pour $N \geq 2$ et $t_1 = \dots = t_N = 0$, on a $W_t(T_n) = 0$ avec certitude, alors que $P(T_n = k) < 1$ pour chaque k .

Pour chaque réalisation, un unique événement $[T_n = \ell]$ est réalisé. Sur cet événement, $W_t(T_n) = t_\ell$. Par conséquent :

$$[W_t(T_n) = t_k] = \bigcup_{\substack{1 \leq \ell \leq N \\ t_\ell = t_k}} [T_n = \ell].$$

Les événements de cette réunion sont disjoints. Donc :

$$P(W_t(T_n) = t_k) = \sum_{\substack{1 \leq \ell \leq N \\ t_\ell = t_k}} p_\ell.$$

Si les coefficients sont deux à deux distincts, la somme ne contient que l'indice $\ell = k$, et l'on retrouve l'égalité annoncée dans ce cas particulier.

Sans aucune hypothèse de distinction des coefficients, la linéarité de l'espérance donne :

$$E(W_t(T_n)) = E\left(\sum_{k=1}^N t_k \mathbf{1}_{[T_n=k]}\right) = \sum_{k=1}^N t_k p_k.$$

Cette formule reste valide lorsque plusieurs coefficients sont égaux et suffit pour les calculs ultérieurs.

11. Espérance d'un produit avec une indicatrice

Posons $B_k = [T_n = k]$. Sa probabilité p_k est strictement positive, donc l'espérance conditionnelle à cet événement est définie.

Comme T_{n+1} prend ses valeurs dans $\llbracket 1, N \rrbracket$, la définition de l'espérance conditionnelle donne :

$$E(T_{n+1} \mid B_k) = \sum_{j=1}^N j \frac{P(T_{n+1} = j, B_k)}{p_k}.$$

En multipliant par p_k , le membre de droite devient :

$$\sum_{j=1}^N j P(T_{n+1} = j, B_k) = E(T_{n+1} \mathbf{1}_{B_k}).$$

On a donc établi :

$$E(T_{n+1} \mathbf{1}_{[T_n=k]}) = E(T_{n+1} \mid T_n = k) p_k.$$

12. (a) Loi jointe de deux maxima successifs

La variable T_n dépend seulement de $U_1, \dots, U_n$. Elle est indépendante de U_{n+1} , par indépendance de l'échantillon et lemme des coalitions. De plus, $T_{n+1} = \max(T_n, U_{n+1})$.

Si $j < k$, l'événement $[T_n = k, T_{n+1} = j]$ est impossible, puisque le maximum ne diminue pas.

Si $j = k$, cet événement est $[T_n = k] \cap [U_{n+1} \leq k]$. Son indépendance donne une probabilité $p_k k / N$.

Si $j > k$, l'événement est $[T_n = k] \cap [U_{n+1} = j]$, de probabilité p_k / N .

Ainsi :

$$P(T_n = k, T_{n+1} = j) = \begin{cases} 0 & j < k, \\ \frac{k}{N} p_k & j = k, \\ \frac{1}{N} p_k & j > k. \end{cases}$$

12. (b) Loi conditionnelle du maximum suivant

Comme $p_k > 0$, on divise la formule de 12(a) par p_k :

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

Ces probabilités totalisent $k/N + (N - k)/N = 1$. Pour $k = N$, la loi est certaine égale à N ; pour $N = 1$, elle est certaine égale à 1.

12. (c) Moyenne conditionnelle du maximum suivant

En utilisant la loi de 12(b),

$$E(T_{n+1} \mid T_n = k) = \frac{k^2}{N} + \frac{1}{N} \sum_{j=k+1}^N j.$$

La formule de somme arithmétique donne :

$$\sum_{j=k+1}^N j = \frac{N(N+1) - k(k+1)}{2}.$$

Après regroupement,

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

Notons cette quantité m_k . Pour $k = N$, elle vaut N , ce qui confirme le cas certain décrit en 12(b).

12. (d) Relation entre les moments des maxima

Les événements $[T_n = k]$, pour $1 \leq k \leq N$, forment un système complet d'événements de probabilités non nulles. Toutes les variables considérées sont bornées, donc la

formule de l'espérance totale s'applique :

$$E(T_{n+1}) = \sum_{k=1}^N m_k p_k.$$

En remplaçant m_k par l'expression de 12(c) et en utilisant

$$\sum p_k = 1,$$

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

Les deux sommes sont respectivement $E(T_n^2)$ et $E(T_n)$. Ainsi :

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

13. Carré d'une combinaison d'indicateurs disjointes

Pour chaque réalisation, exactement une des indicateurs $\mathbf{1}_{[T_n=k]}$ vaut 1 ; toutes les autres valent 0. Si $T_n(\omega) = k$, alors $W_t(T_n)(\omega) = t_k$ et son carré vaut t_k^2 .

On en déduit l'identité entre variables aléatoires :

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

On peut également vérifier cette identité en développant le carré : les produits croisés d'indicateurs sont nuls, car les événements sont disjoints, et chaque indicatrice est égale à son carré.

14. (a) Expression explicite de la fonction à minimiser

Toutes les variables sont à valeurs finies pour t fixé, donc l'espérance du carré existe. Développons :

$$g(t) = E(T_{n+1}^2) - 2E(T_{n+1}W_t(T_n)) + E(W_t(T_n)^2).$$

Par 11, la linéarité et 13,

$$E(T_{n+1}W_t(T_n)) = \sum_{k=1}^N t_k p_k m_k, \quad E(W_t(T_n)^2) = \sum_{k=1}^N t_k^2 p_k.$$

La fonction est donc le polynôme quadratique :

$$g(t_1, \dots, t_N) = E(T_{n+1}^2) + \sum_{k=1}^N p_k (t_k^2 - 2m_k t_k),$$

avec :

$$p_k = \frac{k^n - (k-1)^n}{N^n}, \quad m_k = \frac{N+1}{2} + \frac{k^2 - k}{2N}.$$

Le terme constant peut lui aussi être explicité à l'aide de 4(c), au rang $n+1$:

$$E(T_{n+1}^2) = N^2 - 2Nd_{n+2}(N) - d_{n+1}(N).$$

On obtient ainsi une expression de g entièrement donnée par N, n et les variables $t_1, \dots, t_N$.

14. (b) Minimum global de g

Méthode 1 — Gradient et hessienne

La fonction g est un polynôme sur l'ouvert convexe $\mathbb{R}^N$. Elle est donc de classe C^2 sur cet ouvert. Pour tout k ,

$$\frac{\partial g}{\partial t_k}(t) = 2p_k(t_k - m_k).$$

Comme chaque p_k est strictement positif, le gradient s'annule si et seulement si $t_k = m_k$ pour tous les indices. Il existe donc un unique point critique :

$$\theta = (m_1, \dots, m_N), \quad \theta_k = E(T_{n+1} \mid T_n = k).$$

La hessienne est constante et diagonale :

$$H_g(t) = \text{diag}(2p_1, \dots, 2p_N).$$

Pour chaque vecteur e_k de la base canonique, $H_g(t)e_k = 2p_k e_k$. Les valeurs propres sont donc $2p_1, \dots, 2p_N$, toutes strictement positives. Le critère hessien sur un ouvert convexe assure que le point critique est un minimum global ; il est aussi un minimum local strict.

Tout minimum global d'une fonction dérivable sur l'ouvert $\mathbb{R}^N$ est un point critique. Comme le point critique θ est unique, le minimum global obtenu est lui aussi unique.

Méthode 2 — Complétion des carrés

Pour chaque k , $t_k^2 - 2m_k t_k = (t_k - m_k)^2 - m_k^2$. Ainsi :

$$g(t) = E(T_{n+1}^2) - \sum_{k=1}^N p_k m_k^2 + \sum_{k=1}^N p_k (t_k - m_k)^2.$$

Les deux premiers termes sont indépendants de t . Le dernier est positif ou nul, et, puisque $p_k > 0$, il est nul si et seulement si $t_k = m_k$ pour tous les k .

Par conséquent,

$$g(t) - g(\theta) = \sum_{k=1}^N p_k (t_k - m_k)^2 \geq 0,$$

avec égalité exactement lorsque $t = \theta$. Le minimum global est unique et atteint en :

$$\theta_k = \frac{N+1}{2} + \frac{k^2 - k}{2N} \quad (1 \leq k \leq N).$$

La première méthode utilise le critère du programme en dimension N ; la seconde est plus directe pour une somme de termes quadratiques séparés. Aucune contrainte supplémentaire n'est nécessaire pour obtenir ce minimum.

15. Moyenne et variance de la prévision optimale

Posons $W = W_\theta(T_n)$. Par 10 et $\theta_k = m_k$, puis par la formule de l'espérance totale,

$$E(W) = \sum_{k=1}^N p_k m_k = E(T_{n+1}).$$

La prévision qui minimise le carré moyen vérifie donc automatiquement la condition d'égalité des moyennes.

Pour la variance, 11 donne :

$$E(T_{n+1}W) = \sum_{k=1}^N m_k E(T_{n+1} \mathbf{1}_{[T_n=k]}) = \sum_{k=1}^N p_k m_k^2.$$

D'après 13, cette dernière somme vaut $E(W^2)$. Le développement du carré de l'erreur donne alors :

$$E((T_{n+1} - W)^2) = E(T_{n+1}^2) - E(W^2) \geq 0.$$

Les moyennes étant égales, la soustraction de leur carré commun entraîne :

$$V(T_{n+1}) - V(W) = E((T_{n+1} - W)^2) \geq 0.$$

Ainsi,

$$E(W_\theta(T_n)) = E(T_{n+1}), \quad V(W_\theta(T_n)) \leq V(T_{n+1}).$$

La preuve repose uniquement sur les espérances conditionnelles à des événements et sur la positivité d'un carré.

16. (a) Puissances du maximum et indicatrices

Fixons $i \in \mathbb{N}^*$. Pour toute réalisation telle que $T_n(\omega) = j$, l'unique terme non nul de la somme demandée est celui d'indice $k = j$, et il vaut j^i .

Ainsi, pour chaque réalisation,

$$\sum_{k=1}^N k^i \mathbf{1}_{[T_n=k]} = T_n^i.$$

Cette identité n'exige ni indépendance supplémentaire ni calcul de loi.

16. (b) Formule de la prévision optimale

Par définition de W_θ et la valeur de θ_k trouvée en 14(b),

$$W_\theta(T_n) = \sum_{k=1}^N \left(\frac{N+1}{2} + \frac{k^2 - k}{2N} \right) \mathbf{1}_{[T_n=k]}.$$

Les indicatrices somment à 1. En appliquant 16(a) aux puissances $i = 2$ et $i = 1$, on obtient :

$$W_\theta(T_n) = \frac{N+1}{2} + \frac{1}{2N}(T_n^2 - T_n).$$

Pour $N = 1$, $T_n = 1$ et cette expression vaut 1, ce qui correspond à la prévision certaine.

Partie IV. Estimation

Dans cette partie, N est inconnu. Un estimateur doit être calculable à partir de l'échantillon et de n , sans utiliser la valeur inconnue de N dans sa définition.

Pour tout N fixé, l'échantillon ne prend qu'un nombre fini de valeurs. Toute fonction réelle de cet échantillon a donc des moments de tous ordres ; les moments nécessaires ci-dessous existent. Lorsqu'on affirme qu'un estimateur est sans biais, l'égalité d'espérance doit être valable pour toutes les valeurs admissibles de N .

17. (a) Une correction faisant intervenir le paramètre inconnu

D'après 4(b), la variable $T_n + d_n(N)$ vérifie bien :

$$E(T_n + d_n(N)) = N.$$

Cependant, pour N inconnu, le terme $d_n(N)$ n'est pas calculable à partir du seul échantillon : son expression contient le paramètre que l'on veut estimer.

Ainsi, $T_n + d_n(N)$ a l'espérance recherchée mais ne constitue pas un estimateur de N dans ce problème. On ne peut donc pas la présenter comme un estimateur sans biais.

Le cas particulier $N = 1$ donne $d_n(1) = 0$, mais on ne connaît pas à l'avance que ce cas est réalisé ; il ne rend pas la formule utilisable pour l'ensemble du modèle.

17. (b) Biais asymptotique du maximum

Le maximum T_n est une fonction de l'échantillon ne faisant pas intervenir N dans sa définition. C'est donc un estimateur du paramètre N .

Son biais vaut, d'après 4(b),

$$E(T_n) - N = -d_n(N).$$

Par 4(a), ce biais tend vers 0 pour chaque N fixé. La suite (T_n) est donc asymptotiquement sans biais.

17. (c) Inclusions d'événements

Fixons $\varepsilon > 0$. La relation $E(T_n) = N - d_n(N)$ permet d'écrire :

$$T_n - N = T_n - E(T_n) - d_n(N).$$

L'inégalité triangulaire entraîne :

$$|T_n - N| \leq |T_n - E(T_n)| + |d_n(N)|.$$

Si le membre de gauche est au moins égal à ε , le membre de droite l'est aussi. Donc

$$A_n(\varepsilon) \subset B_n(\varepsilon).$$

Par 4(a), il existe n_0 tel que $n > n_0$ entraîne $|d_n(N)| < \varepsilon/2$. Si $\omega \in B_n(\varepsilon)$ pour un tel n , alors :

$$|T_n(\omega) - E(T_n)| \geq \varepsilon - |d_n(N)| > \frac{\varepsilon}{2}.$$

On en déduit bien :

$$B_n(\varepsilon) \subset [|T_n - E(T_n)| \geq \varepsilon/2] \quad (n > n_0).$$

17. (d) Convergence en probabilité du maximum

Méthode 1 — Inclusions et inégalité de Bienaymé-Tchebychev

Pour tout $\varepsilon > 0$ et tout n suffisamment grand, 17(c) donne :

$$P(|T_n - N| \geq \varepsilon) \leq P(|T_n - E(T_n)| \geq \varepsilon/2).$$

La variable T_n admet une variance finie. L'inégalité de Bienaymé-Tchebychev s'applique :

$$P(|T_n - N| \geq \varepsilon) \leq \frac{4V(T_n)}{\varepsilon^2} \rightarrow 0,$$

puisque la variance tend vers 0 par 4(c). C'est la définition de $T_n \xrightarrow{P} N$.

Méthode 2 — Probabilité de ne pas avoir observé N

Pour tout $\varepsilon > 0$, si $|T_n - N| \geq \varepsilon$, alors $T_n \neq N$. Comme $T_n \leq N$ et prend des valeurs entières, cela impose $T_n \leq N - 1$.

Pour $N \geq 2$, la question 3(a) donne donc :

$$0 \leq P(|T_n - N| \geq \varepsilon) \leq P(T_n \leq N - 1) = \left(\frac{N-1}{N}\right)^n \rightarrow 0.$$

Pour $N = 1$, $T_n = N$ avec certitude, et la probabilité d'écart est toujours nulle. La convergence est donc établie dans tous les cas.

La première méthode suit la chaîne de questions et se réutilise pour des estimateurs dont on connaît les moments. La seconde exploite la loi exacte du maximum et donne une borne géométrique.

18. (a) Probabilité d'un échantillon fixé

Fixons $(u_1, \dots, u_n) \in \llbracket 1, N \rrbracket^n$. Les U_i sont mutuellement indépendantes et chacune prend chaque valeur de son support avec probabilité $1/N$. Ainsi :

$$P\left(\bigcap_{i=1}^n [U_i = u_i]\right) = \prod_{i=1}^n \frac{1}{N} = \frac{1}{N^n}.$$

Toutes les suites admissibles ont donc la même probabilité.

18. (b) Loi conditionnelle de l'échantillon sachant son maximum

Pour $1 \leq k \leq N$, la probabilité $P(T_n = k) = (k^n - (k-1)^n)/N^n$ est strictement positive.

Un échantillon fixé $(u_1, \dots, u_n)$ est compatible avec $T_n = k$ si et seulement si tous ses termes appartiennent à $\llbracket 1, k \rrbracket$ et si son maximum vaut k . Dans ce cas, la probabilité de l'échantillon et de $T_n = k$ vaut N^{-n} ; sinon elle est nulle.

La division par la probabilité de conditionnement donne :

$$P_{[T_n=k]}\left(\bigcap_{i=1}^n [U_i = u_i]\right) = \begin{cases} \frac{1}{k^n - (k-1)^n} & \text{si } (u_1, \dots, u_n) \in \llbracket 1, k \rrbracket^n \text{ et } \max_i u_i = k, \\ 0 & \text{sinon.} \end{cases}$$

Il y a exactement $k^n - (k-1)^n$ suites dans le premier cas : on retire des k^n suites à valeurs au plus égales à k les $(k-1)^n$ suites à valeurs au plus égales à $k-1$. La loi conditionnelle est donc bien normalisée.

La valeur de chaque probabilité et l'ensemble des suites possibles ne dépendent que de k et de n , et non de N , dès lors que $N \geq k$. Ce fait sera essentiel en 20(c).

19. (a) Un estimateur sans biais construit à partir des deux extrêmes

La variable $S_n = T_n + Z_n - 1$ est calculable à partir de l'échantillon sans connaître N . Elle est donc un estimateur admissible.

Par linéarité de l'espérance et les questions 4(b) et 5,

$$E(S_n) = (N - d_n(N)) + (1 + d_n(N)) - 1 = N.$$

Ainsi, S_n est sans biais pour toute valeur $N \geq 1$.

19. (b) Moyenne conditionnelle de S_n

Sur $[T_n = k]$, la variable T_n est constante égale à k . Par linéarité de l'espérance sous la loi conditionnelle et 9(b),

$$E(S_n | T_n = k) = k - 1 + \frac{k^n}{k^n - (k - 1)^n}.$$

Le dénominateur est strictement positif. En réunissant les termes,

$$E(S_n | T_n = k) = \frac{(k - 1)(k^n - (k - 1)^n) + k^n}{k^n - (k - 1)^n}.$$

Le numérateur vaut $k^{n+1} - (k - 1)^{n+1}$. Donc :

$$E(S_n | T_n = k) = \frac{k^{n+1} - (k - 1)^{n+1}}{k^n - (k - 1)^n} = \psi_n(k).$$

Pour $k = 1$, cette expression vaut 1, conformément au fait que tout l'échantillon vaut alors 1.

19. (c) Absence de biais après remplacement par la moyenne conditionnelle

La fonction ψ_n est définie sur tous les entiers strictement positifs par une expression qui ne dépend que de k et de n . Son dénominateur est positif pour chaque k . Par conséquent, $\psi_n(T_n)$ est un estimateur calculable sans connaître N .

Méthode 1 — Espérance totale

Les événements $[T_n = k]$ forment un système complet d'événements de probabilités non nulles. Avec 19(b),

$$E(\psi_n(T_n)) = \sum_{k=1}^N \psi_n(k)P(T_n = k) = \sum_{k=1}^N E(S_n \mid T_n = k)P(T_n = k) = E(S_n) = N.$$

L'estimateur est donc sans biais.

Méthode 2 — Télescopage direct

En utilisant la loi de T_n obtenue en 3(b), les dénominateurs se simplifient :

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

La somme est télescopique et vaut N^{n+1} . Ainsi, l'espérance vaut N .

La première méthode explique le rôle de la moyenne conditionnelle ; la seconde vérifie directement la formule explicite et prépare l'argument d'unicité de 20(b).

19. (d) Comparaison des variances

Méthode 1 — Polynôme obtenu à partir d'un carré

Fixons $k \in \llbracket 1, N \rrbracket$. Sous la loi conditionnelle sachant $T_n = k$, tous les moments considérés existent. Pour tout $\lambda \in \mathbb{R}$, développons :

$$E((S_n - \lambda)^2 \mid T_n = k) = \varphi_n(k) - 2\lambda\psi_n(k) + \lambda^2,$$

par 19(b). Le membre de gauche est positif ou nul. En choisissant $\lambda = \psi_n(k)$, on obtient :

$$\varphi_n(k) - \psi_n(k)^2 \geq 0.$$

Cela établit l'inégalité demandée pour chaque k .

En multipliant par $P(T_n = k)$ puis en sommant, la formule de l'espérance totale donne :

$$E(\psi_n(T_n)^2) \leq \sum_{k=1}^N \varphi_n(k)P(T_n = k) = E(S_n^2).$$

Les deux variables ont la même espérance N , par 19(a) et 19(c). En soustrayant N^2 et en appliquant Koenig-Huygens,

$$V(\psi_n(T_n)) \leq V(S_n).$$

Méthode 2 — Cauchy-Schwarz pour une loi finie conditionnelle

Notons $q_s = P_{[T_n=k]}(S_n = s)$, sur le support fini conditionnel de S_n . Alors

$$\sum_s q_s = 1$$

et :

$$\psi_n(k) = \sum_s s q_s, \quad \varphi_n(k) = \sum_s s^2 q_s.$$

Appliquons Cauchy-Schwarz aux deux vecteurs réels de coordonnées $s\sqrt{q_s}$ et $\sqrt{q_s}$:

$$\left(\sum_s s q_s \right)^2 \leq \left(\sum_s s^2 q_s \right) \left(\sum_s q_s \right) = \sum_s s^2 q_s.$$

On retrouve $\psi_n(k)^2 \leq \varphi_n(k)$. La multiplication par les probabilités de $T_n = k$, la sommation et la soustraction du carré de la moyenne commune donnent ensuite la même comparaison des variances.

La première preuve suit l'indication de l'énoncé et exprime une variance conditionnelle comme un carré moyen. La seconde est une application du produit scalaire à des sommes finies.

19. (e) Variance de S_n et convergence de l'estimateur amélioré

La constante -1 ne modifie pas la variance. Les moments d'ordre deux existent, donc la formule de variance d'une somme s'applique :

$$V(S_n) = V(T_n) + V(Z_n) + 2 \operatorname{Cov}(T_n, Z_n).$$

Posons $a = d_n(N)$ et $b = d_{n+1}(N)$. Par 4(c), 5 et le calcul de covariance de 8(b), cette expression est :

$$2[(2N-1)a - 2Nb - a^2] + 2[Nb - (N-1)a + a^2].$$

Les termes quadratiques se compensent ; le coefficient de a vaut $2N$ et celui de b vaut $-2N$. Ainsi :

$$V(S_n) = 2N(d_n(N) - d_{n+1}(N)).$$

Cette expression est positive ou nulle, puisque chaque base k/N est dans $[0, 1[$, et elle tend vers 0 par 4(a). Pour $N = 1$, elle est identiquement nulle.

Par 19(d),

$$0 \leq V(\psi_n(T_n)) \leq 2N(d_n(N) - d_{n+1}(N)) \rightarrow 0.$$

L'estimateur est sans biais, donc, pour chaque $\varepsilon > 0$, Bienaymé-Tchebychev donne :

$$P(|\psi_n(T_n) - N| \geq \varepsilon) \leq \frac{V(\psi_n(T_n))}{\varepsilon^2} \rightarrow 0.$$

Par conséquent, $\psi_n(T_n) \xrightarrow{P} N$. C'est un estimateur convergent de N .

20. (a) Réduction de variance pour un estimateur sans biais quelconque

Fixons n et soit R_n un estimateur sans biais de N . Comme il est une fonction réelle de l'échantillon, il a un support fini pour chaque N fixé, donc un moment d'ordre deux fini.

Pour $k \in \llbracket 1, N \rrbracket$, posons $f_n(k) = E(R_n \mid T_n = k)$ et $h_n(k) = E(R_n^2 \mid T_n = k)$. Pour tout réel λ , la positivité d'un carré et sa décomposition donnent :

$$0 \leq E((R_n - \lambda)^2 \mid T_n = k) = h_n(k) - 2\lambda f_n(k) + \lambda^2.$$

En prenant $\lambda = f_n(k)$, on obtient $f_n(k)^2 \leq h_n(k)$. La multiplication par $P(T_n = k)$ et la sommation entraînent :

$$E(f_n(T_n)^2) \leq E(R_n^2).$$

Par espérance totale,

$$E(f_n(T_n)) = E(R_n) = N.$$

En soustrayant le carré de cette moyenne commune, on conclut :

$$V(f_n(T_n)) \leq V(R_n).$$

Il reste à vérifier que la fonction f_n est calculable indépendamment de N et à l'identifier ; ce sera fait en 20(c) à l'aide de 18(b).

20. (b) Unicité d'une fonction du maximum sans biais pour tout N

Question adaptée au programme actuel : le quantificateur de la conclusion est explicité sur tous les $k \in \mathbb{N}^*$, puisque la condition d'espérance porte sur toutes les valeurs possibles de N . La fonction ψ_n est définie sur ce même ensemble.

Fixons $n \geq 1$. Supposons d'abord que $F : \mathbb{N}^* \rightarrow \mathbb{R}$ vérifie $E(F(T_n)) = N$ pour chaque

$N \geq 1$. Avec la loi de 3(b), cette propriété équivaut à :

$$\sum_{k=1}^N F(k)[k^n - (k-1)^n] = N^{n+1} \quad \text{pour tout } N \geq 1.$$

Pour $N = 1$, cette identité donne $F(1) = 1 = \psi_n(1)$.

Pour $N \geq 2$, soustrayons l'identité au paramètre $N - 1$ de celle au paramètre N . Tous les termes sauf le dernier s'annulent :

$$F(N)[N^n - (N-1)^n] = N^{n+1} - (N-1)^{n+1}.$$

Le facteur $N^n - (N-1)^n$ étant strictement positif, on peut diviser et obtenir :

$$F(N) = \frac{N^{n+1} - (N-1)^{n+1}}{N^n - (N-1)^n} = \psi_n(N).$$

Cela prouve $F(k) = \psi_n(k)$ pour tous les entiers $k \geq 1$.

Réciproquement, supposons $F(k) = \psi_n(k)$ pour tout $k \geq 1$. Pour chaque N ,

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

Les deux implications sont établies. La fonction ψ_n est donc l'unique fonction du maximum dont l'espérance vaut N pour toutes les valeurs du paramètre.

La condition « pour tout N » est essentielle : une seule contrainte d'espérance pour un N fixé ne déterminerait pas tous les coefficients d'une fonction.

20. (c) Optimalité parmi tous les estimateurs sans biais

Fixons n et un estimateur sans biais R_n . Écrivons :

$$R_n = r_n(U_1, \dots, U_n),$$

où $r_n : (\mathbb{N}^*)^n \rightarrow \mathbb{R}$ est une fonction qui ne dépend pas de N . Cette indépendance du paramètre fait partie de la définition d'un estimateur.

D'après la loi conditionnelle de l'échantillon obtenue en 18(b), pour tout $k \geq 1$ et tout paramètre $N \geq k$,

$$E(R_n | T_n = k) = \frac{1}{k^n - (k-1)^n} \sum_{\substack{(u_1, \dots, u_n) \in \llbracket 1, k \rrbracket^n \\ \max_i u_i = k}} r_n(u_1, \dots, u_n).$$

Cette somme est finie et son expression ne dépend pas de N . Elle définit donc une fonction f_n sur tous les entiers strictement positifs, utilisable comme estimateur $f_n(T_n)$ sans connaître N .

Pour chaque N , la formule de l'espérance totale donne :

$$E(f_n(T_n)) = E(R_n) = N.$$

La fonction f_n vérifie ainsi l'hypothèse de 20(b) pour toutes les valeurs du paramètre. L'unicité démontrée dans cette question implique :

$$f_n(k) = \psi_n(k) \quad \text{pour tout } k \geq 1.$$

Enfin, la comparaison de 20(a) entraîne, pour chaque $N \geq 1$:

$$V(\psi_n(T_n)) = V(f_n(T_n)) \leq V(R_n).$$

Comme R_n était un estimateur sans biais quelconque, $\psi_n(T_n)$ a la variance minimale parmi tous les estimateurs sans biais fondés sur l'échantillon. La preuve est entièrement obtenue par des lois finies, une comparaison de carrés et un argument de télescopage ; aucun théorème général d'optimalité n'est invoqué.

Relevé des adaptations

- 6 : code adapté en Python ; la fonction reçoit explicitement n et N , et la borne exclue de `rd.randint` est précisée.
- 10 : rectification de l'égalité de probabilités, fausse pour des coefficients répétés. La probabilité correcte regroupe les indices partageant une même valeur ; l'égalité

originale est conservée dans le cas de coefficients distincts. La formule générale de l'espérance est ajoutée pour préserver l'objectif de la question.

- 17 à 20 : la définition d'un estimateur comme fonction de l'échantillon ne dépendant pas du paramètre inconnu est explicitée dans le préambule de la partie IV. Elle est nécessaire à 17(a) et à la preuve de 20(c).
- 18(b) : le support conditionnel est écrit directement comme l'ensemble des suites à valeurs dans $\llbracket 1, k \rrbracket$ dont le maximum est k , ce qui explicite son indépendance vis-à-vis de N .
- 19 et 20(b) : la définition de ψ_n et le quantificateur de l'égalité $F(k) = \psi_n(k)$ sont explicités sur $\mathbb{N}^*$, en cohérence avec la condition portant sur toutes les valeurs du paramètre.