



PRÉPASIMPLE

prepasimple.fr

ESSEC

Mathématiques 2

1982

Corrigé de mathématiques approfondies

Mathématiques approfondies — Prépa ECG

Annales de concours · Corrigé

Partie I

L'espace $\mathbb{R}_n[X]$ est de dimension $n + 1$. On utilise uniquement des fonctions polynômes réelles. Un polynôme non nul de degré au plus n possède au plus n racines distinctes.

1.

- a) Pour construire le polynôme demandé, il faut annuler ses valeurs aux n entiers distincts de k et normaliser sa valeur en k .

Le produit

$$\prod_{0 \leq i \leq n, i \neq k} (X - i)$$

est de degré n et s'annule en chaque entier $i \neq k$ de $\{0, \dots, n\}$. Sa valeur en k est :

$$\prod_{\substack{0 \leq i \leq n \\ i \neq k}} (k - i) = \left(\prod_{i=0}^{k-1} (k - i) \right) \left(\prod_{i=k+1}^n (k - i) \right) = k!(-1)^{n-k}(n - k)!.$$

Cette quantité est non nulle. Son inverse est $(-1)^{n-k}/[k!(n - k)!] = (-1)^{n-k} \binom{n}{k}/n!$. Le polynôme donné dans l'énoncé satisfait donc exactement les valeurs imposées.

Pour l'unicité, si deux polynômes de degré au plus n satisfont ces valeurs, leur différence s'annule aux $n + 1$ entiers $0, \dots, n$. Elle est donc identiquement nulle. Ainsi :

$$L_{k,n} = (-1)^{n-k} \frac{\binom{n}{k}}{n!} \prod_{\substack{0 \leq i \leq n \\ i \neq k}} (X - i)$$

est l'unique polynôme demandé. Les produits vides, lorsqu'ils apparaissent aux extrémités dans les calculs, valent 1.

b) Supposons

$$\sum_{k=0}^n c_k L_{k,n} = 0.$$

En évaluant en un entier $i \in \{0, \dots, n\}$, tous les termes s'annulent sauf celui d'indice i , donc $c_i = 0$. La famille est libre. Comme elle possède $n + 1$ éléments dans un espace de dimension $n + 1$, c'est une base.

Si

$$P = \sum_{k=0}^n c_k L_{k,n},$$

l'évaluation en i donne $c_i = P(i)$. Les coordonnées de P sont donc $P(0), P(1), \dots, P(n)$, et :

$$P = \sum_{k=0}^n P(k) L_{k,n}.$$

Pour le polynôme constant $P = 1$, chacune des coordonnées est 1, d'où :

$$\sum_{k=0}^n L_{k,n} = 1.$$

Soit maintenant $x \notin \{0, \dots, n\}$. Le produit

$$D(x) = \prod_{i=0}^n (x - i)$$

est non nul. Pour chaque k , la formule de a) donne :

$$\frac{L_{k,n}(x)}{D(x)} = \frac{(-1)^{n-k} \binom{n}{k}}{n!} \frac{1}{x-k}.$$

En divisant l'identité

$$\sum_k L_{k,n}(x) = 1$$

par $D(x)$, on obtient la décomposition demandée, avec :

$$a_k = (-1)^{n-k} \frac{\binom{n}{k}}{n!} = \frac{(-1)^{n-k}}{k!(n-k)!}, \quad 0 \leq k \leq n.$$

Toutes les divisions sont effectuées hors des racines du dénominateur.

- c) On indexe les lignes et les colonnes de π de 0 à n . La colonne k contient les coefficients de $L_{k,n}$ dans la base $(1, X, \dots, X^n)$.

Le coefficient constant de $L_{k,n}$ est $L_{k,n}(0)$, égal à 1 si $k = 0$ et à 0 sinon. La première ligne de π est donc :

$$(1, 0, \dots, 0).$$

La somme des coefficients d'une ligne d'indice r est le coefficient de X^r dans

$$\sum_{k=0}^n L_{k,n} = 1.$$

Pour toute ligne autre que la première, cette somme est donc 0.

La somme des coefficients de la colonne k est $L_{k,n}(1)$, car évaluer un polynôme en 1 revient à additionner ses coefficients. Ainsi, les sommes des colonnes, dans leur ordre,

sont :

$(0, 1, 0, \dots, 0)$.

Pour déterminer l'inverse, soit

$$P(X) = \sum_{r=0}^n b_r X^r.$$

Sa colonne de coordonnées dans la nouvelle base est ${}^t(P(0), \dots, P(n))$, d'après b). Or

$$P(k) = \sum_{r=0}^n k^r b_r.$$

La matrice π^{-1} est donc la matrice d'évaluation :

$$(\pi^{-1})_{k,r} = k^r, \quad 0 \leq k, r \leq n,$$

$$\pi^{-1} = (k^r)_{0 \leq k, r \leq n}.$$

Sa ligne d'indice 0 est $(1, 0, \dots, 0)$ et, pour $1 \leq k \leq n$, sa ligne d'indice k est $(1, k, k^2, \dots, k^n)$. Cette écriture vaut également pour $n = 1$.

Pour la première colonne, $k^0 = 1$ pour chaque k , y compris $k = 0$; il s'agit de l'évaluation du polynôme constant 1. L'inversibilité est déjà justifiée par le fait que les deux familles sont des bases, sans calcul de déterminant.

2.

a) Pour chaque $k \geq 1$, la formule de I.1 a), appliquée avec les deux indices égaux à k , donne :

$$L_{k,k}(X) = \frac{X(X-1) \cdots (X-k+1)}{k!}.$$

Ce polynôme a pour degré k et pour coefficient dominant $1/k! \neq 0$. Avec $L_{0,0} = 1$, les degrés des $n+1$ polynômes sont donc $0, 1, \dots, n$.

Dans une combinaison linéaire nulle de ces polynômes, un coefficient non nul de plus grand indice k donnerait un terme dominant non nul de degré k , que les autres termes

ne pourraient compenser. Tous les coefficients sont donc nuls. La famille est libre et contient $n + 1$ éléments : c'est une base de $\mathbb{R}_n[X]$.

b) Pour $P, R \in \mathbb{R}_n[X]$ et $\alpha, \beta \in \mathbb{R}$, la définition donne :

$$\Delta(\alpha P + \beta R)(X) = \alpha(P(X+1) - P(X)) + \beta(R(X+1) - R(X)).$$

Donc Δ est linéaire. Les deux polynômes $P(X+1)$ et $P(X)$ ont un degré au plus n , donc leur différence appartient à $\mathbb{R}_n[X]$.

Pour calculer sa matrice, on détermine les images des éléments de la base. On a $\Delta L_{0,0} = 0$. Pour $k \geq 1$, on factorise le produit commun, qui vaut 1 si $k = 1$:

$$\Delta L_{k,k}(X) = \frac{\prod_{i=0}^{k-2} (X-i)}{k!} ((X+1) - (X-k+1)),$$

$$\Delta L_{k,k}(X) = \frac{k}{k!} \prod_{i=0}^{k-2} (X-i) = L_{k-1,k-1}(X).$$

Dans la base $(L_{0,0}, \dots, L_{n,n})$, la matrice est donc celle dont les seuls coefficients non nuls sont des 1 sur la première surdiagonale. En indexant de 0 à n , elle est définie par :

$$A_{r,k} = \begin{cases} 1 & \text{si } k = r + 1, \\ 0 & \text{sinon.} \end{cases}$$

En effet, si

$$P = \sum_{k=0}^n c_k L_{k,k},$$

alors

$$\Delta P = \sum_{k=1}^n c_k L_{k-1,k-1}.$$

Ainsi, $\Delta P = 0$ si et seulement si $c_1 = \dots = c_n = 0$. Le noyau est :

$$\text{Ker } \Delta = \text{Vect}(1).$$

Les images des vecteurs de base engendrent $\text{Vect}(L_{0,0}, \dots, L_{n-1,n-1})$, qui est $\mathbb{R}_{n-1}[X]$ par l'argument de a). Donc :

$$\text{Im } \Delta = \mathbb{R}_{n-1}[X].$$

c) Écrivons

$$P = \sum_{k=0}^n c_k L_{k,k}.$$

L'itération de la formule de b) donne, pour $r \geq 0$:

$$\Delta^r L_{k,k} = \begin{cases} L_{k-r,k-r} & \text{si } r \leq k, \\ 0 & \text{si } r > k. \end{cases}$$

Cette identité se vérifie par récurrence : chaque application de Δ abaisse les deux indices d'une unité, jusqu'au polynôme constant $L_{0,0} = 1$, dont la différence est nulle.

Pour $0 \leq r \leq n$, la linéarité donne alors :

$$\Delta^r P = \sum_{k=r}^n c_k L_{k-r,k-r}.$$

Or $L_{0,0}(0) = 1$, tandis que $L_{m,m}(0) = 0$ pour tout $m \geq 1$, à cause du facteur X . En

évaluant en 0, on obtient $(\Delta^r P)(0) = c_r$. Ainsi :

$$P = \sum_{k=0}^n (\Delta^k P)(0) L_{k,n}.$$

Cette formule fournit les coordonnées dans la base à degrés croissants au moyen des différences successives.

3.

a) On construit un polynôme à partir des seules valeurs de f aux entiers :

$$P_f(X) = \sum_{k=0}^n f(k) L_{k,n}(X).$$

Ce polynôme appartient à $\mathbb{R}_n[X]$, et les propriétés de I.1 a) donnent $P_f(i) = f(i)$ pour chaque entier $i \in \{0, \dots, n\}$. L'existence est donc établie.

Si deux polynômes vérifient ces conditions, leur différence, de degré au plus n , s'annule en $n+1$ points distincts ; elle est identiquement nulle. Le polynôme P_f est donc unique. Aucune continuité ni dérivabilité de f n'est nécessaire : seules ses valeurs aux entiers interviennent.

b) On démontre par récurrence l'égalité des différences de f et de P_f aux entiers où elles sont définies.

Au rang $i = 0$, l'égalité $f(k) = P_f(k)$ pour $0 \leq k \leq n$ découle de a). Supposons qu'elle soit vérifiée pour les différences d'ordre $i < n$ aux entiers de 0 à $n-i$. Pour $0 \leq k \leq n-i-1$, les deux entiers k et $k+1$ appartiennent à cet intervalle. Alors :

$$(\Delta^{i+1} f)(k) = (\Delta^i f)(k+1) - (\Delta^i f)(k),$$

$$(\Delta^{i+1} f)(k) = (\Delta^i P_f)(k+1) - (\Delta^i P_f)(k) = (\Delta^{i+1} P_f)(k).$$

Cela établit l'égalité demandée pour $1 \leq i \leq n$ et $0 \leq k \leq n - i$.

On applique maintenant I.2 c) au polynôme P_f , puis on l'évalue en n :

$$f(n) = P_f(n) = \sum_{k=0}^n (\Delta^k P_f)(0) L_{k,k}(n).$$

L'égalité des différences en 0, y compris pour $k = 0$, permet de remplacer P_f par f dans chaque coefficient. On obtient :

$$f(n) = \sum_{k=0}^n (\Delta^k f)(0) L_{k,k}(n).$$

On pourra aussi utiliser $L_{k,k}(n) = \binom{n}{k}$, conséquence directe de la formule de I.2 a).

- c) Question adaptée au programme actuel : on précise le domaine $x \in [0, n - k]$ des différences d'ordre k . Rectification de la dernière condition du texte source : $x \geq N$ remplace $x \leq N$, qui autorisait notamment des dénominateurs nuls et ne permettait pas l'utilisation ultérieure en II.5.

Posons $f(x) = 1/(a - x)$, avec $a > n$. Pour $x \in [0, n - k]$, tous les facteurs $a - x - r$, où $0 \leq r \leq k$, sont strictement positifs, car $x + r \leq n < a$.

Au rang $k = 0$, la formule est exactement $f(x) = 0!/(a - x)$. Supposons la formule vraie au rang $k < n$. Pour $x \in [0, n - k - 1]$, elle est applicable en x et en $x + 1$. Ainsi :

$$(\Delta^{k+1} f)(x) = \frac{k!}{\prod_{r=1}^{k+1} (a - x - r)} - \frac{k!}{\prod_{r=0}^k (a - x - r)}.$$

En mettant au même dénominateur :

$$(\Delta^{k+1}f)(x) = \frac{k!((a-x) - (a-x-k-1))}{\prod_{r=0}^{k+1}(a-x-r)} = \frac{(k+1)!}{\prod_{r=0}^{k+1}(a-x-r)}.$$

La récurrence est établie. La relation (1), avec $L_{k,k}(n) = \binom{n}{k}$, devient :

$$\frac{1}{a-n} = \sum_{k=0}^n \frac{\binom{n}{k} k!}{a(a-1) \cdots (a-k)}.$$

Le terme $k=0$ vaut $1/a$. En le soustrayant, on obtient :

$$\sum_{k=1}^n \frac{n(n-1) \cdots (n-k+1)}{a(a-1) \cdots (a-k)} = \frac{1}{a-n} - \frac{1}{a} = \frac{n}{a(a-n)}.$$

Pour la formule finale, on applique ce résultat à $n = N-1$ et au paramètre $a = x \geq N$.

L'hypothèse $a > n$ est alors satisfaite. On multiplie l'égalité par N :

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

Sur le domaine $x \geq N$, tous les dénominateurs sont strictement positifs.

Partie II

Préliminaire

1. Posons $p_m = P(X = m)$. Pour $u \in [0, 1]$ et chaque entier $m \geq 0$, on a $0 \leq p_m u^m \leq p_m$, avec $u^0 = 1$, y compris à $u = 0$.

La variable X étant à valeurs dans $\mathbb{N}$,

$$\sum_{m=0}^{\infty} p_m = 1.$$

Par comparaison de séries à termes positifs, la série définissant la fonction génératrice converge. Sa somme est comprise entre 0 et 1. En particulier, elle vaut p_0 à $u = 0$ et 1 à $u = 1$.

2. Méthode 1 — Espérance de variables bornées indépendantes

Pour $u \in [0, 1]$, la variable u^X prend ses valeurs dans $[0, 1]$ et possède donc une espérance. La formule de transfert donne :

$$E(u^X) = \sum_{m=0}^{\infty} P(X = m)u^m = g_X(u).$$

De même, $E(u^Y) = g_Y(u)$ et $E(u^{X+Y}) = g_{X+Y}(u)$. Comme X et Y sont indépendantes, leurs transformations u^X et u^Y sont indépendantes. Elles sont bornées, donc le résultat sur l'espérance du produit s'applique :

$$g_{X+Y}(u) = E(u^{X+Y}) = E(u^X u^Y) = E(u^X)E(u^Y) = g_X(u)g_Y(u).$$

L'égalité $u^{X+Y} = u^X u^Y$ reste valable à $u = 0$ avec la convention donnée : les deux membres valent 1 exactement lorsque $X = Y = 0$.

Méthode 2 — Sommes finies et encadrement

Pour $r, s \geq 0$, posons $a_r = P(X = r)u^r$ et $b_s = P(Y = s)u^s$. L'indépendance et la formule des probabilités totales donnent, pour chaque m :

$$P(X + Y = m)u^m = \sum_{r=0}^m a_r b_{m-r}.$$

Posons

$$A_M = \sum_{r=0}^M a_r,$$

$$B_M = \sum_{s=0}^M b_s$$

et

$$T_M = \sum_{m=0}^M P(X + Y = m)u^m.$$

Les sommes étant finies :

$$T_M = \sum_{\substack{r,s \geq 0 \\ r+s \leq M}} a_r b_s.$$

Les termes sont positifs. Le carré d'indices $0 \leq r, s \leq \lfloor M/2 \rfloor$ est inclus dans le triangle $r + s \leq M$, lui-même inclus dans le carré $0 \leq r, s \leq M$. Ainsi :

$$A_{\lfloor M/2 \rfloor} B_{\lfloor M/2 \rfloor} \leq T_M \leq A_M B_M.$$

Les deux bornes tendent vers $g_X(u)g_Y(u)$ d'après le préliminaire 1. Le théorème d'encadrement donne donc $g_{X+Y}(u) = g_X(u)g_Y(u)$.

La première méthode utilise la propriété d'espérance des variables indépendantes; la seconde établit le résultat directement avec des sommes finies, sans invoquer de théorie générale des séries entières ni de réarrangement d'une série double.

Étude des tirages

Les tirages sont indépendants et uniformes sur les N numéros. Pour justifier que les temps X_n sont finis presque sûrement, fixons un numéro : la probabilité qu'il ne soit pas

apparu en m tirages est $(1 - 1/N)^m$. La probabilité qu'au moins un numéro manque est donc au plus $N(1 - 1/N)^m$, par la majoration d'une réunion finie.

Cette borne tend vers 0. Les événements « tous les numéros sont apparus en m tirages » sont croissants ; la continuité croissante des probabilités montre que tous les numéros apparaissent en un temps fini avec probabilité 1. En particulier, chaque X_n est fini presque sûrement. On peut fixer les variables arbitrairement sur l'éventuel événement exceptionnel de probabilité nulle, sans modifier leurs lois.

1.

a) Méthode 1 — Conditionnement par le premier numéro

Une fois le premier numéro tiré, chacun des $k - 1$ tirages suivants doit reproduire ce numéro. Par indépendance, la probabilité de ces répétitions est $(1/N)^{k-1}$, quelle que soit la valeur du premier numéro. Ainsi :

$$P(\text{les } k \text{ premiers numéros sont identiques}) = N^{1-k}.$$

Méthode 2 — Dénombrement des suites de tirages

Les N^k suites de k numéros sont équiprobables, chacune de probabilité N^{-k} . Parmi elles, exactement N sont constantes : une pour chaque numéro possible. La probabilité demandée vaut donc $N/N^k = N^{1-k}$.

Le conditionnement laisse le premier tirage libre ; le dénombrement compte directement les suites constantes. Les deux méthodes donnent le même résultat pour tout $k \geq 2$.

b) Après le premier tirage, un nouveau numéro apparaît à chaque essai avec probabilité $p_1 = (N - 1)/N$. Tant qu'il n'apparaît pas, le seul numéro déjà connu est répété avec probabilité $q_1 = 1/N$.

Conditionnellement au premier numéro, ces essais sont indépendants, et leur probabilité de succès ne dépend pas de ce numéro. Pour chaque entier $k \geq 1$, $Y_1 = k$

signifie donc que les $k - 1$ essais suivant le premier tirage reproduisent le premier numéro, puis que le suivant est différent. Ainsi :

$$P(Y_1 = k) = \left(\frac{1}{N}\right)^{k-1} \frac{N-1}{N}.$$

Le support est $\mathbb{N}^*$ et Y_1 suit la loi géométrique sur les entiers strictement positifs de paramètre $(N-1)/N$.

c) Pour une loi géométrique sur $\mathbb{N}^*$ de paramètre p , l'espérance et la variance existent et valent $1/p$ et $(1-p)/p^2$. En appliquant ces résultats avec $p = (N-1)/N$:

$$E(Y_1) = \frac{N}{N-1}, \quad V(Y_1) = \frac{N}{(N-1)^2}.$$

Comme $X_1 = 1 + Y_1$, la linéarité de l'espérance et la variance d'une transformation affine donnent :

$$E(X_1) = 1 + \frac{N}{N-1}, \quad V(X_1) = \frac{N}{(N-1)^2}.$$

Pour $u \in [0, 1]$, la fonction génératrice de Y_1 se calcule par une série géométrique de raison $u/N < 1$:

$$g_{Y_1}(u) = \sum_{k=1}^{\infty} \frac{N-1}{N} \left(\frac{1}{N}\right)^{k-1} u^k = \frac{(N-1)u}{N-u}.$$

La relation $X_1 = 1 + Y_1$ donne directement $u^{X_1} = u u^{Y_1}$; en prenant l'espérance :

$$f_1(u) = g_{X_1}(u) = u g_{Y_1}(u) = \frac{(N-1)u^2}{N-u}.$$

Ces expressions valent aussi aux extrémités : elles sont nulles en 0 et égales à 1 en 1.

2.

a) On suppose $2 \leq n \leq N-1$, $p \geq n$ et $k \geq 1$. L'événement $[X_{n-1} = p]$ a une probabilité strictement positive. En effet, un chemin possible consiste à obtenir $n-1$ numéros distincts lors des $n-1$ premiers tirages, à répéter ensuite le premier jusqu'au tirage $p-1$, puis à obtenir un nouveau numéro au tirage p . Tous les choix prescrits ont une probabilité positive.

Sur cet événement, exactement n numéros distincts ont été obtenus au rang p . Fixons un historique complet des p premiers tirages compatible avec cet événement, et notons S l'ensemble de ces n numéros. Les tirages suivants sont indépendants de cet historique.

L'événement $Y_n = k$ signifie que les $k - 1$ tirages suivants appartiennent à S , puis que le dernier appartient à son complément. La probabilité conditionnelle est donc :

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

Cette expression ne dépend ni de l'historique, ni des numéros constituant S , ni de p . Il existe seulement un nombre fini d'historiques de longueur p ; en les réunissant avec les probabilités totales conditionnelles, on obtient :

$$P(Y_n = k \mid X_{n-1} = p) = \left(\frac{n}{N}\right)^{k-1} \frac{N-n}{N}.$$

Cette probabilité ne dépend donc pas de p .

b) Le temps X_{n-1} est fini presque sûrement. Les événements $[X_{n-1} = p]$, pour $p \geq n$, recouvrent ainsi toutes les possibilités à un événement de probabilité nulle près. Les probabilités totales et a) donnent :

$$P(Y_n = k) = \sum_{p=n}^{\infty} P(X_{n-1} = p) \left(\frac{n}{N}\right)^{k-1} \frac{N-n}{N},$$

$$P(Y_n = k) = \left(\frac{n}{N}\right)^{k-1} \frac{N-n}{N}, \quad k \geq 1.$$

Donc Y_n suit la loi géométrique de paramètre $p_n = (N - n)/N$, avec probabilité d'échec $q_n = n/N$. Ce paramètre appartient à $]0, 1[$, car $1 \leq n \leq N - 1$. Les formules de moments et la somme géométrique donnent :

$$E(Y_n) = \frac{N}{N-n}, \quad V(Y_n) = \frac{nN}{(N-n)^2},$$

$$g_{Y_n}(u) = \frac{(N-n)u}{N-nu}, \quad u \in [0, 1].$$

Pour la fonction génératrice, la raison nu/N est au plus $n/N < 1$, ce qui justifie la convergence et le calcul. Ces formules sont également valables pour $n = 1$, d'après la question précédente.

3. Il faut établir l'indépendance des temps d'attente avant d'additionner leurs variances et de multiplier leurs fonctions génératrices. L'indépendance d'un temps d'attente avec le seul temps total précédent ne suffirait pas, à elle seule, à justifier l'indépendance mutuelle.

Fixons des valeurs $y_1, \dots, y_{r-1} \geq 1$ des temps d'attente précédents. Sur l'événement correspondant, le rang d'apparition du r -ième numéro distinct est :

$$p = 1 + y_1 + \dots + y_{r-1}.$$

Cet événement dépend uniquement des p premiers tirages. Pour chaque historique compatible, exactement r numéros sont présents au rang p . Comme dans II.2 a), la probabilité conditionnelle d'attendre encore y_r tirages est :

$$\left(\frac{r}{N}\right)^{y_r-1} \frac{N-r}{N}.$$

Elle est indépendante des valeurs précédentes et de l'historique. En regroupant les historiques puis en appliquant successivement la formule des probabilités composées, on obtient, pour toute famille $y_1, \dots, y_n \geq 1$:

$$P(Y_1 = y_1, \dots, Y_n = y_n) = \prod_{r=1}^n \left[\left(\frac{r}{N}\right)^{y_r-1} \frac{N-r}{N} \right].$$

Chaque facteur est la probabilité marginale correspondante. C'est la factorisation caractérisant l'indépendance mutuelle de ces variables discrètes. Elle est établie sans invoquer de résultat général sur les tirages à un temps aléatoire.

La définition donne la somme télescopique

$$X_n = 1 + \sum_{r=1}^n Y_r.$$

Les moments de chaque variable géométrique existent. Par linéarité, puis par

indépendance pour la variance :

$$E(X_n) = 1 + \sum_{r=1}^n \frac{N}{N-r},$$

$$V(X_n) = \sum_{r=1}^n \frac{rN}{(N-r)^2}.$$

Pour la fonction génératrice, on applique le préliminaire 2 successivement aux variables indépendantes ; le lemme des coalitions assure à chaque étape l'indépendance d'une somme des variables précédentes et de la suivante. La constante 1 apporte le facteur u . Ainsi :

$$f_n(u) = u \prod_{r=1}^n \frac{(N-r)u}{N-ru} = u^{n+1} \prod_{r=1}^n \frac{N-r}{N-ru}, \quad u \in [0, 1].$$

Tous les dénominateurs sont positifs : $N - ru \geq N - r \geq 1$. En particulier, $f_n(0) = 0$, $f_n(1) = 1$, et $X_n \geq n + 1$ presque sûrement.

4.

a) Rectification de l'énoncé : l'argument est N/x , et non n/x . La formule obtenue directement à partir de II.3 justifie cette correction.

Pour $x \geq N$, on a $0 < N/x \leq 1$. En remplaçant u par N/x dans le produit de II.3 :

$$f_n\left(\frac{N}{x}\right) = \frac{N}{x} \prod_{r=1}^n \frac{(N-r)N/x}{N-rN/x}.$$

Or $N - rN/x = N(x-r)/x$. Chaque facteur du produit vaut donc $(N-r)/(x-r)$, d'où :

$$f_n\left(\frac{N}{x}\right) = \frac{N(N-1) \cdots (N-n)}{x(x-1) \cdots (x-n)}.$$

Comme $n \leq N-1$ et $x \geq N$, aucun facteur du dénominateur ne s'annule.

b) Soit $u \in]0, 1]$, et posons $x = N/u \geq N$. Le résultat de I.1 b) donne :

$$\frac{1}{x(x-1)\cdots(x-n)} = \sum_{k=0}^n \frac{(-1)^{n-k} \binom{n}{k}}{n!} \frac{1}{x-k}.$$

En multipliant par le numérateur de a), puis en utilisant

$1/(N/u - k) = u/[N(1 - ku/N)]$, il vient :

$$f_n(u) = \frac{(N-1)(N-2)\cdots(N-n)}{n!} \sum_{k=0}^n (-1)^{n-k} \binom{n}{k} \frac{u}{1 - ku/N}.$$

Le facteur devant la somme est $\binom{N-1}{n}$. Donc :

$$f_n(u) = \binom{N-1}{n} \sum_{k=0}^n (-1)^{n-k} \frac{\binom{n}{k} u}{1 - \frac{k}{N}u}.$$

La formule vaut aussi pour $u = 0$, car les deux membres sont alors nuls. Pour $u \in [0, 1]$ et $k \leq n < N$, les dénominateurs $1 - ku/N$ sont strictement positifs.

c) Chaque fraction de b) est une somme géométrique, puisque $0 \leq k/N < 1$:

$$\frac{u}{1 - ku/N} = \sum_{i=1}^{\infty} \left(\frac{k}{N}\right)^{i-1} u^i.$$

Pour $k = 0$, le membre de droite vaut simplement u : la puissance d'exposant 0 est 1, et les puissances d'exposant strictement positif sont nulles. La somme extérieure en k est finie et chacune de ces séries converge absolument, y compris pour $u = 1$. On peut donc

regrouper leurs termes et écrire :

$$f_n(u) = \sum_{i=1}^{\infty} c_i u^i,$$

$$c_i = \binom{N-1}{n} \sum_{k=0}^n (-1)^{n-k} \binom{n}{k} \left(\frac{k}{N}\right)^{i-1}.$$

Pour identifier ces coefficients avec les probabilités sans utiliser de théorie générale des séries entières, établissons l'unicité nécessaire. On a :

$$\sum_{i=1}^{\infty} |c_i| \leq \binom{N-1}{n} \sum_{k=0}^n \binom{n}{k} \frac{1}{1 - k/N} < +\infty.$$

La suite des probabilités $p_i = P(X_n = i)$ est aussi absolument sommable, sa somme valant 1. Supposons que les coefficients c_i et p_i diffèrent, et soit $r \geq 1$ le premier indice où ils diffèrent. L'égalité des deux expressions de $f_n(u)$ donne, pour $0 < u \leq 1$, après division par u^r :

$$0 = (c_r - p_r) + \sum_{i=r+1}^{\infty} (c_i - p_i) u^{i-r}.$$

La valeur absolue de la dernière somme est au plus

$$u \sum_{i=r+1}^{\infty} |c_i - p_i|,$$

qui tend vers 0 lorsque $u \rightarrow 0^+$. On obtient donc $c_r - p_r = 0$, contradiction. Les coefficients sont bien identiques. Finalement :

$$P(X_n = i) = \binom{N-1}{n} \sum_{k=0}^n (-1)^{n-k} \binom{n}{k} \left(\frac{k}{N}\right)^{i-1}, \quad i \geq 1.$$

La remarque importante est que $X_n \geq n + 1$: pour $1 \leq i \leq n$, cette somme alternée vaut donc 0. Cela fournit les identités :

$$\sum_{k=0}^n (-1)^{n-k} \binom{n}{k} k^{i-1} = 0, \quad 1 \leq i \leq n,$$

avec $k^0 = 1$ dans le cas $i = 1$. Pour $i \geq n + 1$, les probabilités sont strictement positives : on peut obtenir n numéros distincts dans les n premiers tirages, répéter ensuite un numéro connu jusqu'au rang $i - 1$, puis en obtenir un nouveau au rang i . La formule décrit donc une loi sur $\{n + 1, n + 2, \dots\}$, malgré les signes alternés de son expression.

5.

a) Pour $u \in]0, 1]$, posons encore $x = N/u \geq N$. D'après II.4 a), puis la formule rectifiée de I.3 c) :

$$\sum_{r=1}^{N-1} f_r(u) = \sum_{r=1}^{N-1} \frac{N(N-1) \cdots (N-r)}{x(x-1) \cdots (x-r)} = \frac{N(N-1)}{x(x-N+1)}.$$

En remplaçant x par N/u , on obtient :

$$\sum_{r=1}^{N-1} f_r(u) = \frac{(N-1)u^2}{N - (N-1)u}.$$

La formule vaut aussi à $u = 0$, où tous les termes sont nuls. En posant $q = (N-1)/N$, on peut l'écrire :

$$\sum_{r=1}^{N-1} f_r(u) = \frac{qu^2}{1-qu} = \sum_{p=2}^{\infty} q^{p-1} u^p.$$

La dernière série est géométrique de raison qu , strictement inférieure à 1 en valeur absolue sur tout $[0, 1]$.

b) Méthode 1 — Utilisation des temps d'apparition et de la somme des fonctions génératrices

Notons A_p l'événement « le tirage de rang p apporte un numéro nouveau ». Pour $p \geq 2$, si r numéros étaient présents avant ce tirage, ce nouveau numéro est le $r + 1$ -ième numéro distinct ; l'événement correspondant est $[X_r = p]$. Les événements $[X_r = p]$, pour $1 \leq r \leq N - 1$, sont disjoints et leur réunion est A_p . Donc :

$$P(A_p) = \sum_{r=1}^{N-1} P(X_r = p).$$

La somme des fonctions génératrices de a) a précisément ces sommes de probabilités pour coefficients, car la somme en r est finie. La justification élémentaire d'unicité des coefficients de II.4 c) s'applique également ici : la somme absolue des coefficients est au plus $N - 1$, et celle de la série géométrique vaut aussi $N - 1$. On en déduit :

$$P(A_p) = q^{p-1} = \left(1 - \frac{1}{N}\right)^{p-1}.$$

Méthode 2 — Conditionnement par la valeur du tirage courant

Pour un numéro $a \in \{1, \dots, N\}$, la probabilité qu'il soit tiré au rang p est $1/N$. Par indépendance, les $p - 1$ tirages précédents évitent tous a avec probabilité $(1 - 1/N)^{p-1}$. Les N événements ainsi définis sont disjoints, car le numéro du tirage courant est unique. En les sommant :

$$P(A_p) = N \times \frac{1}{N} \left(1 - \frac{1}{N}\right)^{p-1} = \left(1 - \frac{1}{N}\right)^{p-1}.$$

La première méthode relie la question à la somme de fonctions génératrices obtenue en a) ; la seconde calcule directement la probabilité d'un nouveau numéro.

c) Notons D_n le nombre de numéros distincts obtenus en n tirages. Cette variable est comprise entre 1 et N , donc son espérance existe.

Méthode 1 — Indicatrices des tirages apportant un numéro nouveau

Posons $I_p = \mathbf{1}_{A_p}$, avec $I_1 = 1$, puisque le premier numéro est nécessairement nouveau. Chaque numéro distinct est compté exactement à son premier tirage. Ainsi :

$$D_n = \sum_{p=1}^n I_p.$$

Par linéarité de l'espérance, sans supposer ces indicatrices indépendantes, et par b) :

$$E(D_n) = \sum_{p=1}^n q^{p-1} = \frac{1-q^n}{1-q} = N \left[1 - \left(1 - \frac{1}{N} \right)^n \right].$$

Méthode 2 — Indicatrices des numéros déjà sortis

Pour chaque numéro a , posons $J_{a,n} = 1$ s'il est apparu parmi les n premiers tirages, et $J_{a,n} = 0$ sinon. On a

$$D_n = \sum_{a=1}^N J_{a,n}.$$

Par indépendance des tirages, la probabilité que le numéro a soit absent est q^n . Donc $E(J_{a,n}) = 1 - q^n$. La linéarité donne :

$$E(D_n) = \sum_{a=1}^N (1 - q^n) = N(1 - q^n).$$

La première méthode compte les instants d'apparition ; la seconde compte les numéros ayant déjà été vus. Les deux décompositions en indicatrices sont utiles dans les problèmes de dénombrement aléatoire.

Pour N fixé, $0 < q < 1$, donc $q^n \rightarrow 0$ et $E(D_n) \rightarrow N$. Cette limite était prévisible : chaque numéro finit par apparaître presque sûrement, et le nombre de numéros distincts ne peut dépasser N . La formule précise quantitativement cette convergence.

6. Dans cette question, on fait varier n et N sous la contrainte $N = \alpha(n+1)$. Si α n'est pas entier, seuls les entiers n pour lesquels $\alpha(n+1)$ est entier sont considérés. Comme α est rationnel, il existe une infinité de tels indices.

a) Supposons $\alpha > 1$ et posons $m = n+1$. Les formules de II.3 deviennent :

$$E(X_n) = 1 + \sum_{r=1}^{m-1} \frac{\alpha}{\alpha - r/m},$$

$$V(X_n) = \sum_{r=1}^{m-1} \frac{\alpha(r/m)}{(\alpha - r/m)^2}.$$

Les deux fonctions $A(t) = \alpha/(\alpha - t)$ et $B(t) = \alpha t/(\alpha - t)^2$ sont continues sur $[0, 1]$, car $\alpha - t \geq \alpha - 1 > 0$. Les sommes de Riemann associées donnent :

$$\frac{1}{m} \sum_{r=1}^{m-1} A(r/m) \rightarrow \int_0^1 A(t) dt,$$

$$\frac{1}{m} \sum_{r=1}^{m-1} B(r/m) \rightarrow \int_0^1 B(t) dt.$$

L'absence du terme d'indice 0 ne change pas les limites : par rapport aux sommes de Riemann à gauche avec m termes, la différence est respectivement $A(0)/m$ ou $B(0)/m$, qui tend vers 0.

Comme $m/n \rightarrow 1$ et $1/n \rightarrow 0$, il vient :

$$\ell(\alpha) = \lim_{n \rightarrow \infty} \frac{E(X_n)}{n} = \int_0^1 \frac{\alpha}{\alpha - t} dt = \alpha \ln \left(\frac{\alpha}{\alpha - 1} \right).$$

Pour la seconde intégrale, on décompose l'intégrande :

$$\frac{\alpha t}{(\alpha - t)^2} = \frac{\alpha^2}{(\alpha - t)^2} - \frac{\alpha}{\alpha - t}.$$

Une primitive est $\alpha^2/(\alpha - t) + \alpha \ln(\alpha - t)$. Donc :

$$v(\alpha) = \lim_{n \rightarrow \infty} \frac{V(X_n)}{n} = \frac{\alpha^2}{\alpha - 1} - \alpha - \alpha \ln \left(\frac{\alpha}{\alpha - 1} \right),$$

$$v(\alpha) = \frac{\alpha}{\alpha - 1} - \alpha \ln \left(\frac{\alpha}{\alpha - 1} \right).$$

L'intégrale qui la définit est strictement positive, donc $v(\alpha) > 0$.

Pour étudier ensuite $\alpha \rightarrow +\infty$, posons $h = 1/\alpha$, qui tend vers 0^+ . On exprime directement les fonctions en cette nouvelle variable :

$$\ell(\alpha) = \frac{-\ln(1 - h)}{h}, \quad v(\alpha) = \frac{1}{1 - h} - \frac{-\ln(1 - h)}{h}.$$

Les développements limités usuels en 0 donnent :

$$-\ln(1 - h) = h + \frac{h^2}{2} + o(h^2), \quad \frac{1}{1 - h} = 1 + h + o(h).$$

Après division de la première égalité par $h > 0$:

$$\ell(\alpha) = 1 + \frac{h}{2} + o(h), \quad v(\alpha) = \frac{h}{2} + o(h).$$

Par conséquent :

$$L = 1, \quad V = 0, \quad \ell(\alpha) - L \sim \frac{1}{2\alpha}, \quad v(\alpha) - V \sim \frac{1}{2\alpha}.$$

Les sommes de Riemann sont utilisées pour $\alpha > 1$ fixé; l'étude de $\alpha \rightarrow \infty$ porte ensuite sur les fonctions limites. Ces deux passages à la limite sont distincts.

b) Supposons $\alpha = 1$, donc $N = n + 1$. Posons

$$H_n = \sum_{p=1}^n 1/p$$

et

$$S_n = \sum_{p=1}^n 1/p^2.$$

Le sujet admet $H_n \sim \ln n$ et $S_n \rightarrow \pi^2/6$; ces résultats sont utilisés sous ces hypothèses, sans autre formule asymptotique.

Dans l'espérance de II.3, le changement d'indice $p = N - r$ parcourt les entiers de 1 à n . Ainsi :

$$E(X_n) = 1 + (n + 1)H_n.$$

Comme $H_n \sim \ln n$, $n + 1 \sim n$ et $1/(n \ln n) \rightarrow 0$, on obtient :

$$E(X_n) \sim n \ln n.$$

Plus explicitement, $E(X_n)/n = 1/n + (1 + 1/n)H_n \sim \ln n$, ce qui donne $E(X_n)/n \rightarrow +\infty$.

Pour la variance, le même changement d'indice et la relation $r = N - p$ donnent :

$$V(X_n) = \sum_{p=1}^n \frac{N(N-p)}{p^2} = N^2 S_n - N H_n,$$

$$V(X_n) = (n + 1)^2 S_n - (n + 1) H_n.$$

On divise cette égalité par n^2 :

$$\frac{V(X_n)}{n^2} = \left(1 + \frac{1}{n}\right)^2 S_n - \left(1 + \frac{1}{n}\right) \frac{H_n}{n}.$$

Le premier terme tend vers $\pi^2/6$. Le second tend vers 0, puisque $H_n \sim \ln n$ et $\ln n/n \rightarrow 0$. Par conséquent :

$$V(X_n) \sim \frac{\pi^2}{6} n^2.$$

Il s'ensuit $V(X_n)/n \sim (\pi^2/6)n \rightarrow +\infty$. Les deux quotients demandés divergent donc bien vers $+\infty$ dans le cas $\alpha = 1$.

Relevé des adaptations

- I.1 c) : explicitation de la convention de matrice de passage, dont les colonnes sont les coordonnées de la nouvelle base dans la base canonique ; le sens du changement de coordonnées est ainsi fixé.
- I.2 a) : ajout de la convention $L_{0,0} = 1$, nécessaire puisque la définition initiale était donnée pour un entier strictement positif.
- I.3 c) : précision du domaine $[0, n - k]$ des différences d'ordre k et rectification de $x \leq N$ en $x \geq N$ dans l'identité finale. La condition originale incluait des pôles ; la condition rectifiée est celle nécessaire à l'application de II.5 a) et elle satisfait $a > n$ dans la démonstration.
- II, préliminaire 1 : précision de la valeur $u^0 = 1$ en $u = 0$, pour rendre la définition de la fonction génératrice valable sur tout $[0, 1]$.
- II, préambule des tirages : explicitation de l'indépendance et de l'équiprobabilité des tirages au hasard avec remise, conformément au modèle utilisé dans le problème.
- II.4 a) : rectification de $f_n(n/x)$ en $f_n(N/x)$. La formule rectifiée découle du produit de fonctions génératrices et assure la cohérence avec II.4 b) et II.5 a).
- II.6 : précision que les limites sont prises le long des entiers n pour lesquels $N = \alpha(n + 1)$ est entier, afin de conserver un nombre entier de boules.

Aucune question n'a été supprimée et aucun résultat supplémentaire n'a été admis. Les deux résultats admis en II.6 b) sont ceux du document source. Aucune figure ni aucun

code informatique n'est présent dans ce sujet.