



PRÉPASIMPLE

prepasimple.fr

ESSEC

Mathématiques 2

2001

Corrigé de mathématiques approfondies

Mathématiques approfondies — Prépa ECG

Annales de concours · Corrigé

PARTIE I

Toutes les espérances utilisées dans cette partie existent. En effet, l'existence de variances finies entraîne celle des moments d'ordre deux. L'inégalité $2|XY| \leq X^2 + Y^2$ assure l'intégrabilité de XY . De même,

$$2|(X - E(X))(Y - E(Y))| \leq (X - E(X))^2 + (Y - E(Y))^2. \text{ Enfin,}$$

$(\lambda X + Y)^2 \leq 2\lambda^2 X^2 + 2Y^2$. On peut donc développer les espérances ci-dessous et utiliser leur linéarité.

1°) Covariance des variables aléatoires X et Y

1°) a) Développement de la variance

Pour toute variable aléatoire Z possédant un moment d'ordre deux, la définition donne $\text{Cov}(Z, Z) = V(Z)$. En particulier :

$$\text{Cov}(\lambda X + Y, \lambda X + Y) = V(\lambda X + Y).$$

Posons $X_0 = X - E(X)$ et $Y_0 = Y - E(Y)$. Par linéarité de l'espérance, le centrage de $\lambda X + Y$ est $\lambda X_0 + Y_0$. Ainsi :

$$V(\lambda X + Y) = E[(\lambda X_0 + Y_0)^2] = \lambda^2 E(X_0^2) + 2\lambda E(X_0 Y_0) + E(Y_0^2).$$

On obtient donc, pour tout $\lambda \in \mathbb{R}$:

$$V(\lambda X + Y) = \lambda^2 V(X) + 2\lambda \text{Cov}(X, Y) + V(Y).$$

1°) b) Inégalité et condition d'égalité

Notons $a = V(X) > 0$, $b = \text{Cov}(X, Y)$ et $c = V(Y) \geq 0$. La variance est positive ou nulle ; la question précédente donne, pour tout $\lambda \in \mathbb{R}$:

$$0 \leq a\lambda^2 + 2b\lambda + c = a \left(\lambda + \frac{b}{a} \right)^2 + c - \frac{b^2}{a}.$$

En choisissant $\lambda = -b/a$, on obtient $c - b^2/a \geq 0$, soit :

$$\text{Cov}(X, Y)^2 \leq V(X)V(Y).$$

L'égalité équivaut à $V(Y - (b/a)X) = 0$. D'après la propriété rappelée dans l'énoncé, elle équivaut à l'existence d'une constante réelle β telle que $Y = (b/a)X + \beta$ presque sûrement.

Réciproquement, si $Y = \alpha X + \beta$ presque sûrement, alors le centrage donne $Y - E(Y) = \alpha(X - E(X))$. Par conséquent $\text{Cov}(X, Y) = \alpha V(X)$ et $V(Y) = \alpha^2 V(X)$, ce qui donne l'égalité.

Conclusion

l'égalité a lieu si et seulement s'il existe $\alpha, \beta \in \mathbb{R}$ tels que $Y = \alpha X + \beta$ presque sûrement. La pente α peut être nulle lorsque $V(Y) = 0$.

2°) Coefficient de corrélation linéaire des variables aléatoires X et Y

2°) a) Définition, bornes et cas extrêmes

Les deux variances étant strictement positives, les écarts-types $\sigma(X) = \sqrt{V(X)}$ et $\sigma(Y) = \sqrt{V(Y)}$ sont non nuls. Le coefficient est défini par :

$$\rho = \frac{\text{Cov}(X, Y)}{\sigma(X)\sigma(Y)}.$$

Méthode 1 — Utiliser l'inégalité précédente

L'inégalité de 1°) b), divisée par $V(X)V(Y) > 0$, donne $\rho^2 \leq 1$, donc $-1 \leq \rho \leq 1$.

On a $|\rho| = 1$ si et seulement si $Y = \alpha X + \beta$ presque sûrement. Puisque $V(Y) > 0$, nécessairement $\alpha \neq 0$. Dans ce cas :

$$\rho = \frac{\alpha V(X)}{\sqrt{V(X)}\sqrt{\alpha^2 V(X)}} = \frac{\alpha}{|\alpha|}.$$

Ainsi $\rho = 1$ si et seulement si la relation affine a une pente strictement positive ; $\rho = -1$ si et seulement si elle a une pente strictement négative.

Méthode 2 — Centrer et réduire

Posons $U = (X - E(X))/\sigma(X)$ et $T = (Y - E(Y))/\sigma(Y)$. Ces variables sont centrées, de variance 1, et $E(UT) = \rho$. En développant :

$$E[(U - T)^2] = 2 - 2\rho \geq 0, \quad E[(U + T)^2] = 2 + 2\rho \geq 0.$$

On retrouve $-1 \leq \rho \leq 1$. De plus, $U - T$ et $U + T$ sont centrées ; l'annulation des espérances précédentes signifie que leur variance est nulle. Ainsi :

$$\rho = 1 \iff T = U \text{ presque sûrement}, \quad \rho = -1 \iff T = -U \text{ presque sûrement}.$$

En revenant à X et Y , on retrouve les relations affines de pentes respectives $\sigma(Y)/\sigma(X) > 0$ et $-\sigma(Y)/\sigma(X) < 0$.

La première méthode relie la corrélation à l'inégalité de covariance ; la seconde interprète les cas extrêmes comme l'égalité ou l'opposition des variables centrées réduites.

2°) b) Indépendance

Si X et Y sont indépendantes, le théorème sur l'espérance d'un produit de variables indépendantes intégrables donne $E(XY) = E(X)E(Y)$. Donc $\text{Cov}(X, Y) = 0$ et, puisque les variances sont strictement positives :

$$\rho = 0.$$

2°) c) Un coefficient nul sans indépendance

La loi $\mathcal{N}(0, 1)$ donne $E(X) = 0$ et $V(X) = E(X^2) = 1$. Sa densité est $\varphi(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}$.

Justifions les autres moments nécessaires. Pour chaque entier $m \leq 4$, $|x|^m \varphi(x)$ est continue sur tout segment et, à l'infini, est dominée par une constante fois $1/x^2$, par les croissances comparées. Les moments absolus d'ordre au plus quatre existent. La densité est paire, donc l'intégrale de la fonction impaire $x^3 \varphi(x)$ sur $\mathbb{R}$ est nulle : $E(X^3) = 0$.

Pour calculer le quatrième moment, on utilise $\varphi'(x) = -x\varphi(x)$ et une intégration par parties sur $[-A, A]$, avec $A > 0$:

$$\int_{-A}^A x^4 \varphi(x) dx = -[x^3 \varphi(x)]_{-A}^A + 3 \int_{-A}^A x^2 \varphi(x) dx.$$

Les termes de bord tendent vers zéro et les intégrales convergent absolument. En faisant tendre A vers $+\infty$, on obtient $E(X^4) = 3E(X^2) = 3$.

Le transfert et la formule de Koenig-Huygens donnent alors :

$$E(Y) = E(X^2) = 1, \quad V(Y) = E(X^4) - E(X^2)^2 = 3 - 1 = 2.$$

$$\text{Cov}(X, Y) = E(X^3) - E(X)E(X^2) = 0, \quad \rho = 0.$$

Méthode 1 — Utiliser deux événements identiques

Les événements $A = \{|X| \leq 1\}$ et $B = \{Y \leq 1\}$ sont identiques puisque $Y = X^2$. Leur probabilité commune r est strictement entre 0 et 1 : la densité normale est strictement positive sur $[-1, 1]$ et sur des intervalles extérieurs à ce segment.

Si X et Y étaient indépendantes, A et B seraient indépendants. Il faudrait alors $r = P(A \cap B) = P(A)P(B) = r^2$, ce qui contredit $0 < r < 1$.

Méthode 2 — Utiliser la variance de Y

Si X et Y étaient indépendantes, leurs fonctions X^2 et Y le seraient également. Or $X^2 = Y$, donc Y serait indépendante d'elle-même. Le moment d'ordre deux de Y existant, la factorisation de l'espérance du produit donnerait $E(Y^2) = E(Y)^2$, soit $V(Y) = 0$, en contradiction avec $V(Y) = 2$.

La première preuve exhibe des événements dépendants ; la seconde exploite une transformation des variables et un moment déjà calculé. Dans les deux cas, X et Y ne sont pas indépendantes. La réciproque de 2°) b) est donc fausse.

PARTIE II

Toutes les variables de cette partie sont bornées par n : leurs moments, ceux de leurs produits et ceux des polynômes considérés existent. Les sommes sont finies ; leur réorganisation et leur dérivation terme à terme sont donc autorisées.

1°) Calculs préliminaires

1°) a) Somme de coefficients binomiaux

On fixe $q \in \mathbb{N}$ et on raisonne par récurrence sur $n \geq q$, comme demandé.

Initialisation : pour $n = q$,

$$\sum_{k=q}^q \binom{k}{q} = 1 = \binom{q+1}{q+1}.$$

Hérédité : supposons

$$\sum_{k=q}^n \binom{k}{q} = \binom{n+1}{q+1}.$$

Alors, par l'identité de Pascal :

$$\sum_{k=q}^{n+1} \binom{k}{q} = \binom{n+1}{q+1} + \binom{n+1}{q} = \binom{n+2}{q+1}.$$

La formule est donc établie pour tout $n \geq q$:

$$\sum_{k=q}^n \binom{k}{q} = \binom{n+1}{q+1}.$$

1°) b) Trois sommes factorisées

On utilise $k = \binom{k}{1}$, $k(k-1) = 2\binom{k}{2}$ et $k(k-1)(k-2) = 6\binom{k}{3}$. La question précédente donne :

$$\sum_{k=1}^n k = \binom{n+1}{2} = \frac{n(n+1)}{2}.$$

$$\sum_{k=2}^n k(k-1) = 2 \binom{n+1}{3} = \frac{n(n-1)(n+1)}{3}.$$

$$\sum_{k=3}^n k(k-1)(k-2) = 6 \binom{n+1}{4} = \frac{n(n-1)(n-2)(n+1)}{4}.$$

Ces formules restent valables lorsque les sommes sont vides, avec la convention de l'énoncé.

A] Cas $p = 2$

A.2°) Lois des variables aléatoires X et Y

A.2°) a) Dénombrement et lois marginales

Il y a $\binom{j}{2} = j(j-1)/2$ parties à deux éléments d'un ensemble à j éléments, et $\binom{n}{2} = n(n-1)/2$ tirages possibles. Tous ces tirages sont équiprobables.

Pour $2 \leq j \leq n$, l'événement $Y \leq j$ signifie que les deux jetons appartiennent à $\{1, \dots, j\}$.
Donc :

$$P(Y \leq j) = \frac{\binom{j}{2}}{\binom{n}{2}} = \frac{j(j-1)}{n(n-1)}.$$

Pour avoir $Y = j$, on tire j et un jeton parmi les $j-1$ précédents. Ainsi :

$$P(Y = j) = \frac{j-1}{\binom{n}{2}} = \frac{2(j-1)}{n(n-1)}.$$

Pour $1 \leq i \leq n-1$, $X \geq i$ signifie que les deux jetons appartiennent à $\{i, \dots, n\}$, de cardinal $n-i+1$. Il vient :

$$P(X \geq i) = \frac{\binom{n-i+1}{2}}{\binom{n}{2}} = \frac{(n-i+1)(n-i)}{n(n-1)}.$$

Pour avoir $X = i$, on tire i et un jeton parmi les $n - i$ suivants. Donc :

$$P(X = i) = \frac{n-i}{\binom{n}{2}} = \frac{2(n-i)}{n(n-1)}.$$

Les deux probabilités ponctuelles valent bien 0 pour $j = 1$ et $i = n$. Les supports sont $Y(\Omega) = \{2, \dots, n\}$ et $X(\Omega) = \{1, \dots, n-1\}$; les probabilités sont nulles ailleurs.

A.2°) b) Symétrie des lois

Méthode 1 — Comparer les probabilités ponctuelles

Pour $2 \leq j \leq n$:

$$P(n+1-X = j) = P(X = n+1-j) = \frac{2(j-1)}{n(n-1)} = P(Y = j).$$

Les deux variables ont le même support; elles ont donc la même loi. Leurs moments correspondants sont égaux :

$$E(n+1-X) = E(Y), \quad V(n+1-X) = V(Y).$$

Comme $E(n+1-X) = n+1-E(X)$ et $V(n+1-X) = V(X)$, on conclut :

$$E(X) = n + 1 - E(Y), \quad V(X) = V(Y).$$

Méthode 2 — Transformer chaque tirage

L'application $k \mapsto n + 1 - k$ est une bijection de l'ensemble des jetons. Elle induit une bijection sur les tirages à deux éléments et préserve leur équiprobabilité. Le minimum du tirage transformé vaut $n + 1 - Y$ et son maximum vaut $n + 1 - X$. Le maximum du tirage transformé a donc à la fois la loi de Y et celle de $n + 1 - X$.

On retrouve l'égalité des lois, puis les deux identités de moments ci-dessus. Cette méthode explique la symétrie sans calcul des probabilités ; la première la vérifie directement à partir des lois marginales.

A.3°) Espérances et variances des variables aléatoires X et Y

A.3°) a) Espérances

Méthode 1 — Utiliser la loi de Y

Le transfert et II.1°) b) donnent :

$$E(Y) = \frac{2}{n(n-1)} \sum_{j=2}^n j(j-1) = \frac{2}{n(n-1)} \frac{n(n-1)(n+1)}{3} = \frac{2(n+1)}{3}.$$

La symétrie de A.2°) b) donne :

$$E(X) = n + 1 - E(Y) = \frac{n+1}{3}.$$

Méthode 2 — Sommer les probabilités de dépassement de X

Pour une variable entière Z à valeurs dans $\{1, \dots, m\}$, l'identité

$$Z = \sum_{i=1}^m \mathbf{1}_{\{Z \geq i\}}$$

entraîne, par linéarité,

$$E(Z) = \sum_{i=1}^m P(Z \geq i).$$

En l'appliquant à X :

$$E(X) = \frac{1}{\binom{n}{2}} \sum_{i=1}^{n-1} \binom{n-i+1}{2} = \frac{1}{\binom{n}{2}} \sum_{k=2}^n \binom{k}{2} = \frac{\binom{n+1}{3}}{\binom{n}{2}} = \frac{n+1}{3}.$$

Puis $E(Y) = n+1 - E(X) = 2(n+1)/3$. Cette seconde méthode est utile quand les probabilités de dépassement sont plus simples que la loi ponctuelle.

A.3°) b) Moments d'ordre deux et variances

En utilisant la loi de Y et la troisième somme préliminaire :

$$E[Y(Y-2)] = \frac{2}{n(n-1)} \sum_{j=2}^n j(j-1)(j-2) = \frac{(n+1)(n-2)}{2}.$$

Le terme correspondant à $j=2$ est nul. Cette identité reste vraie pour $n=2$, car la somme est alors nulle.

Comme $Y^2 = Y(Y-2) + 2Y$:

$$E(Y^2) = \frac{(n+1)(n-2)}{2} + \frac{4(n+1)}{3} = \frac{(n+1)(3n+2)}{6}.$$

Par Koenig-Huygens :

$$V(Y) = \frac{(n+1)(3n+2)}{6} - \frac{4(n+1)^2}{9} = \frac{(n+1)[3(3n+2) - 8(n+1)]}{18} = \frac{(n+1)(n-2)}{18}.$$

La symétrie donne donc :

$$V(X) = V(Y) = \frac{(n+1)(n-2)}{18}.$$

Pour $n = 2$, ces deux variances sont nulles : les variables sont respectivement constantes égales à 1 et 2.

A.4°) Covariance et coefficient de corrélation linéaire des variables aléatoires X et Y

A.4°) a) Loi conjointe

Pour $1 \leq i < j \leq n$, un seul tirage à deux éléments réalise simultanément $X = i$ et $Y = j$: la partie $\{i, j\}$. Par équiprobabilité :

$$P(\{X = i\} \cap \{Y = j\}) = \frac{1}{\binom{n}{2}} = \frac{2}{n(n-1)}.$$

Pour les autres couples, la probabilité est nulle.

A.4°) b) Espérances mixtes

Méthode 1 — Somme sur la loi conjointe

Le transfert donne :

$$E[X(Y-2)] = \frac{2}{n(n-1)} \sum_{j=2}^n (j-2) \sum_{i=1}^{j-1} i.$$

La somme intérieure vaut $j(j-1)/2$. Ainsi :

$$E[X(Y-2)] = \frac{1}{n(n-1)} \sum_{j=2}^n j(j-1)(j-2) = \frac{(n+1)(n-2)}{4}.$$

On utilise ensuite $XY = X(Y-2) + 2X$:

$$E(XY) = \frac{(n+1)(n-2)}{4} + \frac{2(n+1)}{3} = \frac{(n+1)(3n+2)}{12}.$$

Méthode 2 — Conditionner par la valeur de Y

Pour $2 \leq j \leq n$, $P(Y = j) > 0$. D'après la loi conjointe :

$$P(X = i \mid Y = j) = \frac{1}{j-1}, \quad 1 \leq i \leq j-1.$$

La loi de X conditionnellement à $Y = j$ est uniforme sur $\{1, \dots, j-1\}$. Son espérance vaut $(1+j-1)/2 = j/2$. En sommant sur ces événements :

$$E[X(Y-2)] = \sum_{j=2}^n (j-2) \frac{j}{2} P(Y = j) = \frac{1}{2} E[Y(Y-2)] = \frac{(n+1)(n-2)}{4}.$$

De même :

$$E(XY) = \frac{1}{2}E(Y^2) = \frac{(n+1)(3n+2)}{12}.$$

La première méthode utilise directement un dénombrement ; la seconde fait apparaître la moyenne du minimum lorsque le maximum est fixé.

A.4°) c) Covariance et corrélation

Question adaptée au programme actuel : le cas $n = 2$, où les variances sont nulles, est distingué avant de définir la corrélation.

La formule de la covariance donne :

$$\text{Cov}(X, Y) = \frac{(n+1)(3n+2)}{12} - \frac{n+1}{3} \frac{2(n+1)}{3}.$$

En mettant au même dénominateur :

$$\text{Cov}(X, Y) = \frac{(n+1)[3(3n+2) - 8(n+1)]}{36} = \frac{(n+1)(n-2)}{36}.$$

Si $n \geq 3$, les deux variances sont strictement positives et égales. Donc :

$$\rho = \frac{(n+1)(n-2)/36}{(n+1)(n-2)/18} = \frac{1}{2}.$$

Ce coefficient est indépendant de n sur le domaine $n \geq 3$. Pour $n = 2$, la covariance vaut 0 et le coefficient de corrélation n'est pas défini ; on ne doit pas simplifier un quotient 0/0.

B] Cas général $1 \leq p \leq n$

B.5°) Lois des variables aléatoires X et Y

B.5°) a) Dénombrement des tirages

Le nombre total de tirages est $D = \binom{n}{p} > 0$, chacun de probabilité $1/D$.

Pour $p \leq j \leq n$, les tirages réalisant $Y \leq j$ sont les parties à p éléments de $\{1, \dots, j\}$. Pour réaliser $Y = j$, il faut inclure j et choisir les $p - 1$ autres jetons parmi les $j - 1$ précédents. Ainsi :

$$P(Y \leq j) = \frac{\binom{j}{p}}{\binom{n}{p}}, \quad P(Y = j) = \frac{\binom{j-1}{p-1}}{\binom{n}{p}}.$$

Pour $1 \leq i \leq n - p + 1$, les tirages réalisant $X \geq i$ sont les parties à p éléments de $\{i, \dots, n\}$. Pour avoir $X = i$, on inclut i et on choisit les $p - 1$ autres jetons parmi les $n - i$ suivants. Donc :

$$P(X \geq i) = \frac{\binom{n-i+1}{p}}{\binom{n}{p}}, \quad P(X = i) = \frac{\binom{n-i}{p-1}}{\binom{n}{p}}.$$

Les supports sont $\{p, \dots, n\}$ pour Y et $\{1, \dots, n - p + 1\}$ pour X . Les probabilités ponctuelles sont nulles hors de ces supports.

Si $p = 1$, les deux variables ont la loi uniforme sur $\{1, \dots, n\}$ et sont égales. Si $p = n$, $X = 1$ et $Y = n$ sont constantes. Les formules ci-dessus couvrent ces deux situations.

B.5°) b) Symétrie des lois

Pour $p \leq j \leq n$:

$$P(n + 1 - X = j) = P(X = n + 1 - j) = \frac{\binom{j-1}{p-1}}{\binom{n}{p}} = P(Y = j).$$

Les supports coïncident ; les variables ont la même loi. Les formules de transformation affine donnent :

$$E(X) = n + 1 - E(Y), \quad V(X) = V(Y).$$

La réflexion des jetons $k \mapsto n + 1 - k$, décrite en A.2°) b), fournit également une preuve directe pour un tirage à p éléments : elle préserve l'équiprobabilité des parties et transforme le maximum en $n + 1 - X$.

B.6°) Espérances et variances des variables aléatoires X et Y

B.6°) a) Identité binomiale et espérances

Question adaptée au programme actuel : les moments sont exprimés en fonction des deux paramètres n et p .

Pour $p \leq j \leq n$, les factorielles donnent :

$$j \binom{j-1}{p-1} = \frac{j!}{(p-1)!(j-p)!} = p \frac{j!}{p!(j-p)!} = p \binom{j}{p}.$$

Méthode 1 — Pondérer la loi du maximum

Par transfert et par II.1°) a) :

$$E(Y) = \frac{1}{D} \sum_{j=p}^n j \binom{j-1}{p-1} = \frac{p}{D} \sum_{j=p}^n \binom{j}{p} = \frac{p \binom{n+1}{p+1}}{\binom{n}{p}}.$$

Le quotient de coefficients binomiaux vaut $(n+1)/(p+1)$, comme on le vérifie en écrivant les factorielles. Donc :

$$E(Y) = \frac{p(n+1)}{p+1}, \quad E(X) = n + 1 - E(Y) = \frac{n+1}{p+1}.$$

Méthode 2 — Probabilités de dépassement du minimum

L'identité d'indicatrices établie en A.3°) a) donne :

$$E(X) = \sum_{i=1}^{n-p+1} P(X \geq i) = \frac{1}{D} \sum_{i=1}^{n-p+1} \binom{n-i+1}{p} = \frac{1}{D} \sum_{k=p}^n \binom{k}{p} = \frac{n+1}{p+1}.$$

Puis la symétrie donne $E(Y) = p(n+1)/(p+1)$. La première méthode traite le maximum par sa loi ; la seconde traite le minimum par ses probabilités de dépassement.

B.6°) b) Moments d'ordre deux et variances

Question adaptée au programme actuel : la dépendance en p est précisée, comme en a).

L'identité de a) et l'égalité $(j+1)\binom{j}{p} = (p+1)\binom{j+1}{p+1}$ donnent :

$$(j+1)j\binom{j-1}{p-1} = p(j+1)\binom{j}{p} = p(p+1)\binom{j+1}{p+1}.$$

Par transfert et changement d'indice $k = j+1$ dans une somme finie :

$$E[(Y+1)Y] = \frac{p(p+1)}{D} \sum_{j=p}^n \binom{j+1}{p+1} = \frac{p(p+1)}{D} \binom{n+2}{p+2}.$$

Le quotient de coefficients binomiaux est :

$$\frac{\binom{n+2}{p+2}}{\binom{n}{p}} = \frac{(n+2)(n+1)}{(p+2)(p+1)}.$$

Ainsi :

$$E[(Y+1)Y] = \frac{p(n+1)(n+2)}{p+2}.$$

Puisque $Y^2 = (Y+1)Y - Y$:

$$E(Y^2) = \frac{p(n+1)(n+2)}{p+2} - \frac{p(n+1)}{p+1} = \frac{p(n+1)[(p+1)n+p]}{(p+1)(p+2)}.$$

En retranchant $E(Y)^2$ et en factorisant :

$$V(Y) = \frac{p(n+1)\{(p+1)[(p+1)n+p] - p(n+1)(p+2)\}}{(p+1)^2(p+2)}.$$

Dans les accolades, le coefficient de n est $(p+1)^2 - p(p+2) = 1$ et le terme constant est $p(p+1) - p(p+2) = -p$. Donc :

$$V(Y) = V(X) = \frac{p(n+1)(n-p)}{(p+1)^2(p+2)}.$$

Les variances sont strictement positives si $p < n$, et nulles si $p = n$. Pour $p = 1$, elles valent $(n^2 - 1)/12$; pour $p = 2$, elles redonnent $(n+1)(n-2)/18$.

B.7°) Covariance et coefficient de corrélation linéaire des variables aléatoires X et Y

B.7°) a) Loi conjointe et cas d'un seul jeton

Question adaptée au programme actuel : on sépare $p = 1$, car la formule de dénombrement contenant $p - 2$ n'y a pas de sens et les deux extrêmes sont alors égaux.

Si $p \geq 2$, pour réaliser $X = i$ et $Y = j$, il faut inclure les deux jetons i et j et choisir les $p - 2$ autres dans $\{i + 1, \dots, j - 1\}$, de cardinal $j - i - 1$. Pour les indices de l'énoncé, ce cardinal est au moins $p - 2$. Par conséquent :

$$P(\{X = i\} \cap \{Y = j\}) = \frac{\binom{j-i-1}{p-2}}{\binom{n}{p}}, \quad p \leq j \leq n, \quad 1 \leq i \leq j - p + 1.$$

La probabilité est nulle hors de ce domaine.

Si $p = 1$, on tire un jeton uniforme et $X = Y$. La loi conjointe est :

$$P(\{X = i\} \cap \{Y = j\}) = \begin{cases} 1/n & \text{si } 1 \leq i = j \leq n, \\ 0 & \text{sinon.} \end{cases}$$

B.7°) b) Écart entre les extrêmes et moment mixte

Méthode 1 — Somme sur la loi conjointe

Supposons d'abord $p \geq 2$. Le transfert donne :

$$E[(Y + 1)(Y - X)] = \frac{1}{D} \sum_{j=p}^n (j + 1) \sum_{i=1}^{j-p+1} (j - i) \binom{j-i-1}{p-2}.$$

Pour j fixé, on pose $k = j - i$; les indices parcourus sont $p - 1 \leq k \leq j - 1$. L'identité déjà utilisée en B.6°) a), avec $p - 1$ à la place de p , donne $k \binom{k-1}{p-2} = (p - 1) \binom{k}{p-1}$. On obtient :

$$\sum_{i=1}^{j-p+1} (j - i) \binom{j-i-1}{p-2} = (p - 1) \sum_{k=p-1}^{j-1} \binom{k}{p-1} = (p - 1) \binom{j}{p}.$$

Ainsi, en utilisant $(j + 1) \binom{j}{p} = (p + 1) \binom{j+1}{p+1}$, puis II.1°) a) :

$$E[(Y+1)(Y-X)] = \frac{(p-1)(p+1)}{D} \sum_{j=p}^n \binom{j+1}{p+1} = \frac{(p-1)(p+1)}{D} \binom{n+2}{p+2}.$$

On en déduit la forme factorisée :

$$E[(Y+1)(Y-X)] = \frac{(p-1)(n+1)(n+2)}{p+2}.$$

Lorsque $p = 1$, $Y - X = 0$ identiquement ; la même formule est donc vraie, sans utiliser la loi conjointe du cas $p \geq 2$.

L'identité $(Y+1)X = (Y+1)Y - (Y+1)(Y-X)$ entraîne maintenant, pour tout $1 \leq p \leq n$:

$$E[(Y+1)X] = \frac{p(n+1)(n+2)}{p+2} - \frac{(p-1)(n+1)(n+2)}{p+2} = \frac{(n+1)(n+2)}{p+2}.$$

Enfin, $(Y+1)X = XY + X$, donc :

$$E(XY) = \frac{(n+1)(n+2)}{p+2} - \frac{n+1}{p+1} = \frac{(n+1)[(p+1)n+p]}{(p+2)(p+1)}.$$

Méthode 2 — Étudier le minimum conditionnellement au maximum

Supposons $p \geq 2$ et fixons $j \in \{p, \dots, n\}$. L'événement $Y = j$ a une probabilité strictement positive. Conditionnellement à cet événement, le jeton j est présent et les $p-1$ autres jetons forment une partie uniforme de $\{1, \dots, j-1\}$. En effet, chacune de ces parties correspond à un tirage de même probabilité.

On applique le résultat sur l'espérance du minimum de B.6°) a) à une urne de $j-1$ jetons et un tirage de $p-1$ jetons ; ses hypothèses $1 \leq p-1 \leq j-1$ sont satisfaites. Il donne :

$$E(X | Y = j) = \frac{(j-1) + 1}{(p-1) + 1} = \frac{j}{p}.$$

Pour $p = 1$, l'égalité $X = Y$ donne également $E(X | Y = j) = j = j/p$. En sommant sur le système complet des événements $\{Y = j\}$:

$$E[(Y+1)(Y-X)] = \sum_{j=p}^n (j+1) \left(j - \frac{j}{p}\right) P(Y = j) = \frac{p-1}{p} E[(Y+1)Y].$$

Comme $p \geq 1$, la division est licite. Le moment calculé en B.6°) b) donne

$$E[(Y+1)(Y-X)] = (p-1)(n+1)(n+2)/(p+2).$$

De même :

$$E(XY) = \sum_{j=p}^n j \frac{j}{p} P(Y = j) = \frac{1}{p} E(Y^2) = \frac{(n+1)[(p+1)n+p]}{(p+2)(p+1)}.$$

La première méthode suit directement la loi conjointe et les sommes binomiales ; la seconde réutilise le problème du minimum dans une urne plus petite.

B.7°) c) Covariance et corrélation

Question adaptée au programme actuel : le cas $p = n$, où les variances s'annulent, est distingué.

Les espérances précédentes donnent :

$$\text{Cov}(X, Y) = \frac{(n+1)[(p+1)n+p]}{(p+1)(p+2)} - \frac{p(n+1)^2}{(p+1)^2}.$$

Le calcul du numérateur est le même qu'en B.6°) b), sans le facteur p extérieur. On obtient :

$$\text{Cov}(X, Y) = \frac{(n+1)(n-p)}{(p+1)^2(p+2)}.$$

Pour $1 \leq p < n$, les variances sont strictement positives et égales. Par conséquent :

$$\rho = \frac{\text{Cov}(X, Y)}{V(Y)} = \frac{1}{p}.$$

Pour un p fixé, ce coefficient ne dépend pas de n sur le domaine $n > p$. Si $p = 1 < n$, on retrouve $\rho = 1$, puisque $X = Y$ et la variable uniforme n'est pas constante.

Si $p = n$, le tirage contient tous les jetons : $X = 1$ et $Y = n$. La covariance est nulle et la corrélation n'est pas définie. Ce cas inclut $n = p = 1$.

C] Retour au cas $p = 2$

On pose $N = n(n-1) > 0$. La fonction G est un polynôme en deux variables réelles : aucune théorie des séries entières ou des fonctions génératrices infinies n'est nécessaire.

C.8°) Utilisation de la fonction génératrice des variables aléatoires X et Y

C.8°) a) Dérivées et moments

La somme définissant G est finie ; on peut la dériver terme à terme. En notant $p_{ij} = P(\{X = i\} \cap \{Y = j\})$:

$$\frac{\partial G}{\partial u}(u, v) = \sum_{j=1}^n \sum_{i=1}^n i p_{ij} (1+u)^{i-1} (1+v)^j.$$

En évaluant en $(0, 0)$ et en regroupant selon i , on obtient :

$$\frac{\partial G}{\partial u}(0, 0) = \sum_{i=1}^n i \sum_{j=1}^n p_{ij} = \sum_{i=1}^n i P(X = i) = E(X).$$

De même, la dérivée par rapport à v donne $\frac{\partial G}{\partial v}(0, 0) = E(Y)$.

En dérivant une seconde fois, les facteurs sont respectivement $i(i-1)$, $j(j-1)$ et ij . Les termes de degré trop faible ont une dérivée seconde nulle ; on ne leur attribue pas une puissance négative. Il vient :

$$\frac{\partial^2 G}{\partial u^2}(0,0) = E[X(X-1)] = E(X^2) - E(X).$$

$$\frac{\partial^2 G}{\partial v^2}(0,0) = E[Y(Y-1)] = E(Y^2) - E(Y).$$

$$\frac{\partial^2 G}{\partial u \partial v}(0,0) = E(XY).$$

Ces cinq identités permettront de retrouver tous les moments demandés en c).

C.8°) b) Sommes géométriques et expression polynomiale

On utilise uniquement la loi conjointe :

$$G(u,v) = \frac{2}{N} \sum_{j=2}^n (1+v)^j \sum_{i=1}^{j-1} (1+u)^i.$$

Pour $u \neq 0$, la somme géométrique intérieure vaut :

$$\sum_{i=1}^{j-1} (1+u)^i = \frac{(1+u)^j - (1+u)}{u}.$$

Le terme correspondant à $j=1$ serait nul, donc on peut ajouter ce terme. Avec $1+w = (1+u)(1+v)$:

$$G(u,v) = \frac{2}{Nu} \left[\sum_{j=1}^n (1+w)^j - (1+u) \sum_{j=1}^n (1+v)^j \right].$$

Pour $w \neq 0$ et $v \neq 0$, les deux sommes géométriques sont :

$$\sum_{j=1}^n (1+w)^j = (1+w) \frac{(1+w)^n - 1}{w}, \quad \sum_{j=1}^n (1+v)^j = (1+v) \frac{(1+v)^n - 1}{v}.$$

En utilisant à nouveau $(1+u)(1+v) = 1+w$, on obtient la formule demandée :

$$G(u, v) = \frac{2(1+w)}{Nu} \left[\frac{(1+w)^n - 1}{w} - \frac{(1+v)^n - 1}{v} \right].$$

Le binôme donne :

$$\frac{(1+w)^n - 1}{w} - \frac{(1+v)^n - 1}{v} = \sum_{k=1}^n \binom{n}{k} (w^{k-1} - v^{k-1}) = \sum_{k=2}^n \binom{n}{k} (w^{k-1} - v^{k-1}).$$

On pourrait donc écrire

$$G(u, v) = \frac{2(1+w)}{Nu} \sum_{k=2}^n \binom{n}{k} (w^{k-1} - v^{k-1}),$$

mais cette expression contient encore une division par u . Pour calculer les dérivées à l'origine, supprimons cette division.

L'identité polynomiale

$$a^m - b^m = (a - b) \sum_{r=0}^{m-1} a^{m-1-r} b^r$$

se vérifie en développant le produit : tous les termes intermédiaires se compensent.

Puisque $w - v = u(1+v)$, elle donne :

$$G(u, v) = \frac{2}{N} (1+w)(1+v) \sum_{k=2}^n \binom{n}{k} \sum_{r=0}^{k-2} w^{k-2-r} v^r, \quad w = u + v + uv.$$

Cette dernière expression est polynomiale et n'a plus de dénominateur variable. Elle vaut pour tous $u, v \in \mathbb{R}$, y compris à l'origine. Pour le justifier sans supposer une propriété générale de prolongement, fixons $v \notin \{0, -1\}$. L'égalité déjà établie vaut pour tous les u sauf au plus deux valeurs ; par continuité des deux polynômes en u , elle vaut pour tout u . Pour chaque u fixé, la continuité en v permet ensuite d'inclure $v = 0$ et $v = -1$.

C.8°) c) Calcul des moments par les coefficients du polynôme

Question adaptée au programme actuel : la conclusion concernant la corrélation distingue $n = 2$ et $n \geq 3$.

On a :

$$\frac{\partial w}{\partial u} = 1 + v, \quad \frac{\partial w}{\partial v} = 1 + u.$$

Pour extraire les dérivées d'ordre au plus deux, il suffit de conserver les monômes de degré total au plus deux dans le polynôme obtenu en b). Il s'agit d'un calcul exact sur un polynôme fini, et non d'une composition de développements limités.

Posons $a = \binom{n}{2}$, $b = \binom{n}{3}$ et $c = \binom{n}{4}$, avec $\binom{n}{k} = 0$ lorsque $k > n$. Les trois coefficients s'écrivent, pour tout $n \geq 2$:

$$a = \frac{N}{2}, \quad b = \frac{N(n-2)}{6}, \quad c = \frac{N(n-2)(n-3)}{24}.$$

Dans la somme polynomiale de b), les termes de degrés zéro, un et deux en w, v sont :

$$a + b(w + v) + c(w^2 + wv + v^2).$$

Les autres termes ont un degré total au moins trois en w, v . Comme $w = u + v + uv$ est sans terme constant, après substitution ils ne produisent aucun monôme de degré inférieur à trois en u, v .

En multipliant par $(1 + w)(1 + v) = 1 + w + v + wv$, la partie de degré au plus deux en w, v devient :

$$a + (a + b)(w + v) + (a + 2b + c)wv + (b + c)(w^2 + v^2).$$

On remplace w par $u + v + uv$. Pour les degrés au plus deux, on a $w + v = u + 2v + uv$, $wv = uv + v^2$ et $w^2 + v^2 = u^2 + 2uv + 2v^2$, les termes omis étant tous de degré au moins trois. On obtient l'identité :

$$\begin{aligned} G(u, v) &= \frac{2}{N} [a + (a + b)u + 2(a + b)v + (b + c)u^2 + (2a + 5b + 3c)uv + (a + 4b + 3c)v^2] \\ &\quad + R(u, v), \end{aligned}$$

où R est un polynôme dont tous les monômes ont un degré total au moins trois. Toutes ses dérivées partielles d'ordre au plus deux s'annulent à l'origine. La constante $2a/N = 1$ concorde avec $G(0, 0) = 1$.

Lecture des dérivées premières

$$\frac{\partial G}{\partial u}(0, 0) = \frac{2(a + b)}{N} = 1 + \frac{n - 2}{3} = \frac{n + 1}{3}.$$

$$\frac{\partial G}{\partial v}(0, 0) = \frac{4(a + b)}{N} = \frac{2(n + 1)}{3}.$$

D'après a), on retrouve donc $E(X) = (n + 1)/3$ et $E(Y) = 2(n + 1)/3$ sans réutiliser les lois marginales.

Lecture des dérivées secondes pures

$$\frac{\partial^2 G}{\partial u^2}(0, 0) = \frac{4(b + c)}{N} = 4 \left[\frac{n - 2}{6} + \frac{(n - 2)(n - 3)}{24} \right] = \frac{(n - 2)(n + 1)}{6}.$$

Pour la dérivée en v , la simplification donne :

$$\frac{\partial^2 G}{\partial v^2}(0,0) = \frac{4(a+4b+3c)}{N} = 2 + \frac{8(n-2)}{3} + \frac{(n-2)(n-3)}{2}.$$

En mettant au dénominateur 6, le numérateur est

$$12 + 16(n-2) + 3(n-2)(n-3) = 3n^2 + n - 2 = (n+1)(3n-2). \text{ Ainsi :}$$

$$\frac{\partial^2 G}{\partial v^2}(0,0) = \frac{(n+1)(3n-2)}{6}.$$

On ajoute les espérances conformément à a) :

$$E(X^2) = \frac{(n-2)(n+1)}{6} + \frac{n+1}{3} = \frac{n(n+1)}{6}.$$

$$E(Y^2) = \frac{(n+1)(3n-2)}{6} + \frac{2(n+1)}{3} = \frac{(n+1)(3n+2)}{6}.$$

La formule de Koenig-Huygens donne alors :

$$V(X) = \frac{n(n+1)}{6} - \frac{(n+1)^2}{9} = \frac{(n+1)(n-2)}{18}.$$

$$V(Y) = \frac{(n+1)(3n+2)}{6} - \frac{4(n+1)^2}{9} = \frac{(n+1)(n-2)}{18}.$$

Lecture de la dérivée mixte

$$\frac{\partial^2 G}{\partial u \partial v}(0,0) = \frac{2(2a+5b+3c)}{N} = 2 + \frac{5(n-2)}{3} + \frac{(n-2)(n-3)}{4}.$$

Au dénominateur 12, le numérateur est

$$24 + 20(n - 2) + 3(n - 2)(n - 3) = 3n^2 + 5n + 2 = (n + 1)(3n + 2). \text{ Par a) :}$$

$$E(XY) = \frac{(n + 1)(3n + 2)}{12}.$$

Finalement :

$$\text{Cov}(X, Y) = E(XY) - E(X)E(Y) = \frac{(n + 1)(n - 2)}{36}.$$

Pour $n \geq 3$, les variances sont strictement positives et le coefficient vaut $\rho = 1/2$. Pour $n = 2$, elles sont nulles ; $X = 1$ et $Y = 2$ sont constantes, et le coefficient n'est pas défini. Les résultats de la partie A sont tous retrouvés à partir de la loi conjointe et du polynôme G .

Relevé des adaptations

- II.A.4°) c) : distinction entre $n = 2$, où la corrélation est indéfinie, et $n \geq 3$, où elle vaut $1/2$. Le sujet original permettait $n = 2$ tout en demandant le coefficient sans réserve.
- II.B.6°) a) et b) : précision « en fonction de n et p » ; le paramètre p intervient dans les moments et ne peut être omis de leur description.
- II.B.7°) a) : traitement séparé de $p = 1$, où $X = Y$ et la loi conjointe est diagonale. Le dénombrement avec deux extrêmes distincts n'est applicable qu'à $p \geq 2$.
- II.B.7°) c) : distinction entre $p = n$, où la corrélation est indéfinie, et $p < n$, où elle vaut $1/p$.
- II.C.8°) c) : même précision sur le domaine de définition de la corrélation que dans II.A.4°) c).