ECE 2 - Année 2021-2022
ENC Bessières
Mathématiques - F. Gaunard
http://frederic.gaunard.com

Devoir surveillé n°3 - Sujet B

Samedi 4 Décembre
Durée : 4 heures

Problème 1

Cet exercice est extrait du sujet **ECRICOME 2018**.

Partie I : Étude de deux suites

Pour tout entier naturel n non nul, on pose :

$$u_n = \sum_{k=1}^{n} \frac{1}{k} - \ln(n), \quad \text{et} \quad v_n = u_n - \frac{1}{n}.$$

(1) (a) La limite en 0 ne pose aucun problème, c'est celle du log, on a donc

$$\lim_{x \to 0} f(x) = -\infty.$$

En $+\infty$, on a une forme indéterminée; mais on met tout sous un même logarithme et on sait que $x/(x+1)$ tend vers 1 donc

$$\lim_{x \to +\infty} f(x) = \lim_{x \to +\infty} \frac{1}{x+1} + \ln\left(\frac{x}{x+1}\right) = 0.$$

(b) Sur $\mathbb{R}_+^*$, f est dérivable comme combinaison de fonctions usuelles dérivables. On a d'ailleurs

$$f'(x) = -\frac{1}{(x+1)^2} + \frac{1}{x} - \frac{1}{x+1} = \frac{-x + (x+1)^2 - x(x+1)}{x(x+1)^2} = \frac{1}{x(x+1)^2} > 0.$$

On en déduit le tableau de variations de f

x	0	$+\infty$
$f'(x)$	‖	$+$
f	‖ $-\infty$	$\nearrow$ 0

(c) Par définition de la suite (u_n)

$$\begin{aligned}
u_{n+1} - u_n &= \sum_{k=1}^{n+1} \frac{1}{k} - \ln(n+1) - \sum_{k=1}^{n} \frac{1}{k} + \ln(n) \\
&= \frac{1}{n+1} - \ln(n+1) + \ln(n) \\
&= f(n).
\end{aligned}$$

(d) D'après le tableau de variations de f, on voit que $f(x) < 0$ pour tout $x > 0$. En particulier, $f(n) < 0$ pour tout entier $n \in \mathbb{N}^*$, donc $u_{n+1} - u_n = f(n) < 0$ et la suite (u_n) est (strictement) décroissante.

(e) C'est un petit programme sans réelle difficulté que l'on peut faire avec une boucle `for` ou avec une opération pointée. On propose les deux versions.

```
function y=u(n)
    y=0;
    for k=1:n
        y=y+1/k;
    end
    y=y-log(n)
endfunction
```

ou bien

```
function y=u(n)
    y=sum([1:n].^(-1))-log(n)
endfunction
```

(2) (a) Soit $n \in \mathbb{N}^*$,

$$\begin{aligned} v_{n+1} - v_n &= u_{n+1} - \frac{1}{n+1} - u_n + \frac{1}{n} \\ &= \frac{1}{n} - \frac{1}{n+1} + \frac{1}{n+1} + \ln(n) - \ln(n+1) \\ &= \frac{1}{n} + \ln\left(\frac{n}{n+1}\right) \\ &= \frac{1}{n} - \ln\left(\frac{n+1}{n}\right) \\ &= \frac{1}{n} - \ln\left(1 + \frac{1}{n}\right), \end{aligned}$$

et c'est ce qu'on voulait.

(b) On peut étudier la fonction différence ou utiliser un argument de convexité. En effet, la fonction $x \mapsto \ln(x+1)$ est concave sur $]-1; +\infty[$ (elle est de classe $\mathcal{C}^2$ et sa dérivée seconde est strictement négative). Sa courbe représentative se trouve donc au dessous de toutes ses tangentes, y compris celle en $x = 0$ qui a pour équation $y = x$, et qui donne bien l'inégalité attendue.

En appliquant cette inégalité à $x = 1/n$, on voit que $v_{n+1} - v_n \geq 0$ ou encore que (v_n) est croissante.

(c) La formule de Taylor-Young, ou une connaissance du cours permet d'écrire de développement limité à l'ordre 2 en 0 de cette fonction usuelle

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

Comme $1/n \to 0$, si $n \to +\infty$, on peut utiliser ce DL dans l'expression de $v_{n+1} - v_n$:

$$v_{n+1} - v_n = \frac{1}{n} - \ln\left(1 + \frac{1}{n}\right) + \frac{1}{n} - \frac{1}{n} + \frac{1}{2n^2} + o\left(\frac{1}{n^2}\right) = \frac{1}{2n^2} + o\left(\frac{1}{n^2}\right)$$

ou encore

$$v_{n+1} - v_n \underset{n \to +\infty}{\sim} \frac{1}{2n^2}.$$

(d) La série de terme général $1/2n^2$ est convergente comme multiple d'une série de Riemann convergente. Par critère d'équivalence pour les séries à termes positifs, la série de terme générale $v_{n+1} - v_n$ est donc de même nature c'est à dire convergente. On note alors γ la valeur de sa somme

$$\gamma = \sum_{n=1}^{+\infty}(v_{n+1} - v_n).$$

(e) Le calcul de la somme partielle de la série susnommée fait apparaitre une somme télescopique;

$$\sum_{k=1}^{n-1}(v_{k+1} - v_k) = v_n - v_1 = v_n - v_n - u_1 + 1 = v_n$$

Ainsi, (v_n) est convergente et

$$\lim_{n\to+\infty} v_n = \lim_{n\to+\infty}\sum_{k=1}^{n-1}(v_{k+1} - v_k) = \sum_{k=1}^{n+\infty}(v_{k+1} - v_k) = \gamma.$$

(3) (a) Cette question n'étant pas vraiment formulée; on interprète l'énoncé comme une demande de justification de la convergence de (u_n) et le calcul de sa limite. On voit que

$$v_n = u_n - \frac{1}{n} \iff u_n = v_n + \frac{1}{n}.$$

Comme (v_n) converge et que $1/n \to 0$, on en déduit que (u_n) converge et a la même limite que (v_n), c'est à dire γ.

(b) (v_n) étant croissante et convergente vers γ, (u_n) étant décroissante et convergente vers γ, on a bien l'encadrement demandé

$$v_n \leq \gamma \leq u_n.$$

Ceci permet de voir que

$$0 \leq u_n - \gamma = v_n - \gamma + \frac{1}{n} \leq \frac{1}{n},$$

ce qui donne bien

$$|u_n - \gamma| \leq \frac{1}{n}.$$

(c) Le programme proposé permet de calculer une approximation de γ à la précision `eps` près (rentrée par l'utilisateur). En effet, une telle approximation sera réalisée par un terme u_n tel que $|u_n - \gamma| <$ `eps`, ce qui, d'après la question précédente a lieu dès que $1/n <$ `eps`. Il suffit de prendre le premier entier n tel que $n > 1/$`eps`, donné par $\lfloor$`1/eps`$\rfloor + 1$.

Partie II: Étude d'une série

Pour tout entier naturel n, on pose $a_n = \dfrac{1}{n(2n-1)}$.

(4) Le terme général de la série est équivalent à celui d'une série convergente. En effet,

$$a_n = \frac{1}{n(2n-1)} \underset{n\to+\infty}{\sim} \frac{1}{2n^2},$$

et la convergence de ce terme a été justifiée ci-avant.

(5) (a) Observons que

$$\sum_{k=1}^{n}\frac{1}{2k-1} = \sum_{k=1}^{n}\frac{1}{k} - \sum_{k=1}^{n}\frac{1}{2k} \iff \sum_{k=1}^{n}\frac{1}{2k-1} + \sum_{k=1}^{n}\frac{1}{2k} = \sum_{k=1}^{2n}\frac{1}{k}$$

et on reconnait une décomposition des indices de sommation selon leur parité. Plus précisément,

$$\sum_{k=1}^{2n} \frac{1}{k} \;=\; \sum_{\substack{k=1\\ k \text{ pair}}}^{2n} \frac{1}{k} + \sum_{\substack{k=1\\ k \text{ impair}}}^{2n} \frac{1}{k}$$

Or,

$$\begin{cases} 1 \leq k \leq 2n \\ k \text{ pair} \end{cases} \iff \begin{cases} k = 2j \\ 1 \leq j \leq n \end{cases}$$

et

$$\begin{cases} 1 \leq k \leq 2n \\ k \text{ impair} \end{cases} \iff \begin{cases} k = 2j-1 \\ 1 \leq j \leq n \end{cases}$$

Ainsi,

$$\sum_{k=1}^{2n} \frac{1}{k} = \sum_{\substack{k=1\\ k \text{ pair}}}^{2n} \frac{1}{k} + \sum_{\substack{k=1\\ k \text{ impair}}}^{2n} \frac{1}{k} = \sum_{j=1}^{n} \frac{1}{2j} + \sum_{j=1}^{n} \frac{1}{2j-1}$$

(b) Il suffit de mettre au même dénominateur et de procéder par identification

$$\begin{aligned}
\frac{1}{n(2n-1)} = \frac{\alpha}{n} + \frac{\beta}{2n-1} \quad &\iff \quad \frac{1}{n(2n-1)} = \frac{\alpha(2n-1)+\beta n}{n(2n-1)} \\
&\iff \begin{cases} 2\alpha + \beta &= 0 \\ -\alpha &= 1 \end{cases} \\
&\iff \begin{cases} \alpha &= -1 \\ \beta &= 2 \end{cases}
\end{aligned}$$

(c) On utilise les résultats des deux dernières questions

$$\begin{aligned}
\sum_{k=1}^{n} a_k &= -\sum_{k=1}^{n} \frac{1}{k} + 2\sum_{k=1}^{n} \frac{1}{2k-1} && \text{(d'après 5b.)} \\
&= -\sum_{k=1}^{n} \frac{1}{k} + 2\left(\sum_{k=1}^{2n} \frac{1}{k} - \frac{1}{2}\sum_{k=1}^{n} \frac{1}{k}\right) && \text{(d'après 5a.)} \\
&= 2\sum_{k=1}^{2n} \frac{1}{k} - 2\sum_{k=1}^{n} \frac{1}{k} \\
&= 2\sum_{k=n+1}^{2n} \frac{1}{k}.
\end{aligned}$$

(6) (a) On revient à la définition de u_n

$$\begin{aligned}
u_{2n} - u_n + \ln(2) &= \sum_{k=1}^{2n} \frac{1}{k} - \ln(2n) - \sum_{k=1}^{n} \frac{1}{k} + \ln(n) + \ln(2) \\
&= \sum_{k=n+1}^{2n} \frac{1}{k} - \ln(2) - \ln(n) + \ln(n) + \ln(2) \\
&= \sum_{k=n+1}^{2n} \frac{1}{k},
\end{aligned}$$

ce qu'on attendait.

(b) D'après 5c. et 6a., on a

$$\begin{aligned}\sum_{k=1}^{+\infty} a_k &= \lim_{n\to+\infty} \sum_{k=1}^{n} a_k \\ &= \lim_{n\to+\infty} 2 \sum_{k=n+1}^{2n} \frac{1}{k} \\ &= 2\times \lim_{n\to+\infty} (u_{2n} - u_n + \ln(2)) \\ &= 2\ln(2),\end{aligned}$$

car, comme (u_n) converge, u_{2n} et u_n ont même limite et leur différence tend vers 0.

(7) (a) On voit que

$$\begin{aligned}\sum_{k=n+1}^{2n} \frac{1}{k} &= \sum_{k=1}^{n} \frac{1}{n+k} \\ &= \sum_{k=1}^{n} \frac{1}{n\left(1+\frac{k}{n}\right)} \\ &= \frac{1}{n}\sum_{k=1}^{n} \frac{1}{1+\frac{k}{n}}.\end{aligned}$$

(b) On sait plus ou moins[1] que, si g est de classe $\mathcal{C}^1$ sur $[0;1]$, alors

$$\lim_{n\to+\infty} \frac{1}{n}\sum_{k=1}^{n} g\left(\frac{k}{n}\right) = \int_0^1 g(x)\mathrm{d}x.$$

En prenant $g(x) = \ln(x+1)$, on voit que

$$\lim_{n\to+\infty} \frac{1}{n}\sum_{k=1}^{n} \frac{1}{1+\frac{k}{n}} = \int_0^1 \frac{\mathrm{d}x}{x+1} = \ln(2),$$

et on retrouve bien la valeur précédente

$$\sum_{k=1}^{+\infty} a_k = 2\lim_{n\to+\infty} \sum_{k=n+1}^{2n} \frac{1}{k} = 2\lim_{n\to+\infty} \frac{1}{n}\sum_{k=1}^{n} \frac{1}{1+\frac{k}{n}} = 2\ln(2).$$

Remarque importante: Sommes de Riemann

Dans cet exercice, on a entre autres utilisé le résultat suivant:

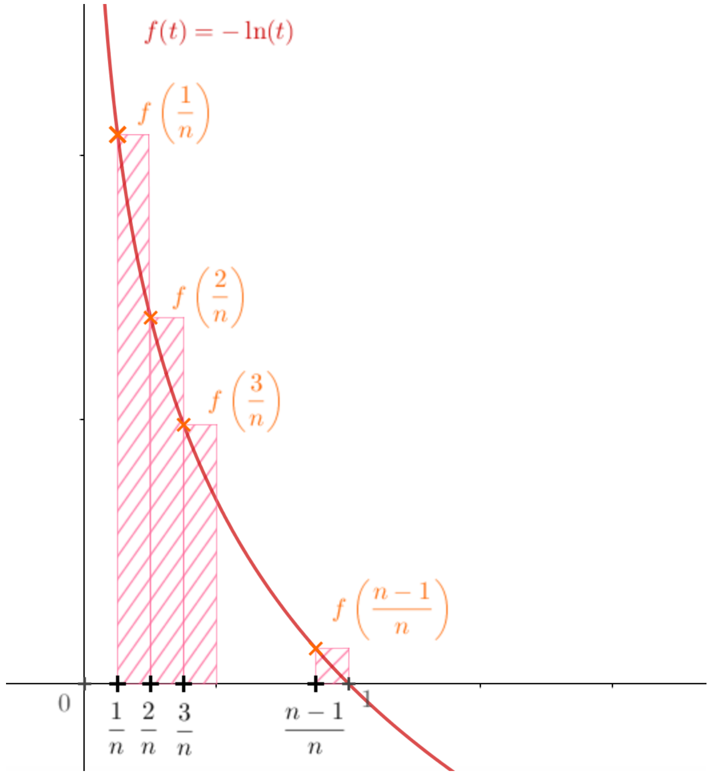
Si f est une fonction de classe $\mathcal{C}^1$ sur $[0;1]$, alors

$$\lim_{n\to+\infty} \frac{1}{n}\sum_{k=1}^{n} f\left(\frac{k}{n}\right) = \int_0^1 f(t)\mathrm{d}t.$$

Ce résultat apparait parfois sous le nom de *sommes de Riemann*[2] et fait partie du cours de première année. La preuve est **admise**, mais certains sujets peuvent avoir pour but de la démontrer ou d'en démontrer une variante (**EDHEC 2008**, ou le très chouette Exercice 2 de ce DM). Lorsque le pas de $1/n$ tend vers 0, la somme des aires des rectangles tend vers l'intégrale.

[1] C'est par exemple démontré dans le sujet **EDHEC 2008**

[2] À ne pas confondre avec les séries de Riemann ni même avec les intégrales de Riemann....



Problème 2

Partie I - Préliminaires

L'objectif de cette première partie est notamment l'obtention de la formule de Vandermonde $(*)$ ci-dessous afin de pouvoir légitimer la formule de la loi hypergéométrique qui suit et obtenir son espérance.

(1) On décide avec les conventions $a_i = 0$ si $i > m$ et $b_j = 0$ si $j > n$ d'écrire

$$P(t) = \sum_{i=0}^{m+n} a_i t^i, \quad \text{et} \quad Q(t) = \sum_{j=0}^{m+n} b_j t^j.$$

On peut obtenir une puissance k de t en faisant le produit d'une puissance r et de la puissanc $k-r$, et ceci pour n'importe quel r entre 0 et k. Le coefficient devant t^r dans le développement de $P(t)$ est a_r, celui devant t^{k-r} dans $Q(t)$ est b_{k-r}. La somme de ces produits donne donc le coefficient devant t^k dans le développement de $R(t)$. Par identification, on a bien

$$c_k = \sum_{r=0}^{k} a_r b_{k-r}.$$

(2) Soient m, n deux entiers naturels non nuls et $0 \leq p \leq m$ et $p \leq n$.

(a) Par la formule du binôme de Newton, on peut écrire

$$\begin{aligned} P(t) &= (1+t)^m = \sum_{i=0}^{m} \binom{m}{i} t^i = \sum_{i=0}^{m+n} \binom{m}{i} t^i \\ Q(t) &= \sum_{j=0}^{n} \binom{n}{j} t^j = \sum_{j=0}^{m+n} \binom{n}{j} t^j \\ R(t) &= (1+t)^m (1+t)^n = (1+t)^{m+n} = \sum_{k=0}^{m+n} \binom{m+n}{k} t^k \end{aligned}$$

En appliquant le résultat précédent avec $a_i = \binom{m}{i}$, $b_j = \binom{n}{j}$ et $c_k = \binom{m+n}{k}$, on peut identifier le coefficient de degré p de $R(t)$ pour obtenir la formule de *Vandermonde*

$$\sum_{k=0}^{p} \binom{n}{k}\binom{m}{p-k} = \binom{m+n}{p} \qquad (*)$$

(b) On peut aussi obtenir la formule précédente avec un argument de dénombrement; pour former un groupe de p personnes parmi $m+n$ personnes, on peut découper le groupe total en deux: un de m personnes et un de n personnes. On veut choisir p personnes: il suffit d'en choisir k dans le premier groupe et de compléter avec $p-k$ choisies dans le deuxième groupe, et la valeur de k peut varier entre 0 et p.

(3) D'après la formule de *Vandermonde* précédente (avec $n = p$, $m = n - p$), on a

$$\sum_{k=0}^{p} P(X=k) = \sum_{k=0}^{p} \frac{\binom{p}{k}\binom{n-p}{p-k}}{\binom{n}{p}} = \frac{\binom{n}{p}}{\binom{n}{p}} = 1.$$

Comme chacune des quantités

$$\frac{\binom{p}{k}\binom{n-p}{p-k}}{\binom{n}{p}}$$

est clairement positive (quotient de quantités positives) et que la somme fait 1, il est clair que $P(X=k) \in [0;1]$. On a bien une variable aléatoire.

(4) (a) Soit $k \in [\![1,p]\!]$, on a :

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

(b) On calcule alors l'espérance de X, avec la question précédente.

$$\begin{aligned}
E(X) &= \sum_{k=0}^{p} kP(X=k) = \sum_{k=0}^{p} k\frac{\binom{p}{k}\binom{n-p}{p-k}}{\binom{n}{p}} \\
&= \frac{1}{\binom{n}{p}} \sum_{k=1}^{p} k\binom{p}{k}\binom{n-p}{p-k} \\
&= \frac{p}{\binom{n}{p}} \sum_{k=1}^{p} \binom{p-1}{k-1}\binom{n-p}{p-k} = \frac{p}{\binom{n}{p}} \sum_{k=1}^{p} \binom{p-1}{k-1}\binom{n-p}{p-1-(k-1)} \\
&= \frac{p}{\binom{n}{p}} \sum_{j=0}^{p-1} \binom{p-1}{j}\binom{n-p}{p-1-j} = \frac{p}{\binom{n}{p}} \times \binom{n-1}{p-1} \qquad \text{(d'après } (*)) \\
&= \frac{p(n-1)!p!(n-p)!}{n!(p-1)!(n-1-(p-1))!} \\
&= \frac{p^2}{n},
\end{aligned}$$

ce qui est bien le résultat attendu.

Partie II - Simulation sous `SciLab`

(4) Parmi les arguments de cette fonction, U représente la liste de tous les tickets dans l'urne; L la liste des tickets gagnants et p le nombre de tickets à retirer de la liste U (on a d'ailleurs `p=length(L)`).

L'instruction `find("condition")` renvoie la liste des composantes de la liste prise en argument et qui satisfont à la "condition".
Ici, `find(L==U(j)))` renvoie la liste des composantes de L qui sont égales à $U(j)$. On sait qu'il y en a au plus une. Le but étant de voir si le terme $U(j)$ sélectionné au hasard dans la liste U est un élément de L car dans ce cas on ne veut pas l'enlever et il faut en sélectionner un autre. Si `length(find(L==U(j)))==1`, cela signifie donc que le terme $U(j)$ fait partie de la liste L. Tant que c'est le cas, on choisit un nouvel indice j au hasard. Une fois que le terme de la liste $U(j)$ n'est pas un élément gagnant de la liste L, on le retire de la liste U par concaténation avec la commande `U=[U(1:j-1), U(j+1:N)]` qui renvoie la liste U privée du terme $U(j)$. On poursuis le processus jusqu'à avoir enlevé p mauvais numéros et on renvoie la liste ainsi tronquée avec `S=U`.

(5) Le principe de concaténation des listes qui apparait dans le programme précédent permet sans difficulté d'écrire celui-demandé. Pour information, ce programme était notamment demandé dans le sujet **ESSEC II 2020**.

```
function [y,L]=selection(U)
    N=length(U)
    j=grand(1,1,'uin', 1, N)
    y=U(j)
    L=[U(1:j-1), U(j+1: N)]
 endfunction
```

(6) On répète le processus de sélection jusqu'à avoir sélectionné p éléments de la liste U. On retire ensuite avec le tout premier programme p mauvais tickets *au hasard* de l'urne et on sélectionne à nouveau p tickets parmi ceux restant dans l'urne. La commande `length(find(L==Y(k)))` renvoie 1 ou 0 et permet de vérifier si le nouveau ticket pioché est un ticket gagnant ou non et vient s'ajouter au nombre de tickets gagnants déjà piochés.

```
function [A, B]=strategies(p, n, L)
    U=1:n
    Y=zeros(1,p)
    A=0
    for k=1:p
        [Y(k), U]=selection(U)
        A=A+length(find(L==Y(k)))
    end
    U=retrait(p, U, L)
    B=0
    Z=zeros(1,p)
    for k=1:p
        [Z(k), U]=selection(U)
        B=B+length(find(L==Z(k)))
    end
endfunction
```

Partie III : Étude de cas particuliers avec $p = 1$ et $p = 2$

(7) Dans cette question et dans cette question uniquement, on suppose que $n = 3$ et $p = 1$.

Il y a donc trois tickets dans l'urne dont un seul gagnant. C'est la situation connue sous le nom de Problème de Monty Hall.

(a) On pioche donc un seul ticket dans cette première partie d'expérience. Ce ticket est gagnant ou perdant donc $A(\Omega) = \{0; 1\}$. On a une chance sur trois que le ticket pioché soit gagnant, on a donc

$$A \hookrightarrow \mathcal{B}(1/3).$$

(b) Sachant $[X_1 = 0]$ signifie qu'on a pioché au premier coup un ticket perdant. Comme l'animateur du jeu retire ensuite un mauvais ticket de l'urne qui n'en contenait plus que 2, il est alors certain de piocher au deuxième coup le ticket gagnant. Ainsi

$$P_{[X_1=0]}(Y_1 = 1) = 1.$$

La suite du même raisonnement donne clairement

$$P_{[X_1=1]}(Y_1 = 1) = 0.$$

Il suit, par la formule des probabilités totales appliquée au s.c.e $\{[X_1 = 1]; [X_1 = 0]\}$,

$$\begin{aligned} P(B = 1) &= P(Y_1 = 1) = P_{[X_1=0]}(Y_1 = 1)P(X_1 = 0) + P_{[X_1=1]}(Y_1 = 1)P(X_1 = 1) \\ &= P(X_1 = 0) = \frac{2}{3} \end{aligned}$$

Ainsi, on a

$$B \hookrightarrow \mathcal{B}(2/3).$$

(c) On constate que, dans cette première situation,

$$E(A) = \frac{1}{3} < \frac{2}{3} = E(B)$$

et la deuxième stratégie semble donc meilleure.

(8) Dans cette question et dans cette question uniquement, on suppose $p = 2$ et $n \geq 6$.

(a) Le numéro pioché en premier est gagnant (et $X_1 = 1$) ou perdant (et $X_1 = 0$); on a donc une loi de Bernoulli. Il y a 2 tickets gagnant et n tickets en tout, on a donc

$$X_1 \hookrightarrow \mathcal{B}\left(\frac{2}{n}\right).$$

(b) Sachant $[X_1 = 1]$, il reste 1 bons tickets dans l'urne qui en contient $n - 1$. Sachant $[X_1 = 0]$ il en reste 2. Il suit que

$$\begin{aligned} P(X_1 = 0 \cap X_2 = 0) &= P(X_1 = 0)P_{[X_1=0]}(X_2 = 0) = \frac{n-2}{n} \times \frac{n-2-1}{n-1} \\ P(X_1 = 0 \cap X_2 = 1) &= P(X_1 = 0)P_{[X_1=0]}(X_2 = 1) = \frac{n-2}{n} \times \frac{2}{n-1} \\ P(X_1 = 1 \cap X_2 = 0) &= P(X_1 = 1)P_{[X_1=1]}(X_2 = 0) = \frac{2}{n} \times \frac{n-1-1)}{n-1} \\ P(X_1 = 1 \cap X_2 = 1) &= P(X_1 = 1)P_{[X_1=1]}(X_2 = 1) = \frac{2}{n} \times \frac{1}{n-1} \end{aligned}$$

On résume cette loi conjointe sous forme d'un tableau

$X_1 \backslash X_2$	0	1
0	$\dfrac{(n-2)(n-3)}{n(n-1)}$	$\dfrac{2(n-2)}{n(n-1)}$
1	$\dfrac{2(n-2)}{n(n-1)}$	$\dfrac{2}{n(n-1)}$

(c) On obtient la loi marginale de X_2 en sommant les termes du tableau, par la formule des probabilités totales appliquée au s.c.e $\{[X_1 = 0];\ [X_1 = 1]\}$. Plus précisément,

$$\begin{aligned} P(X_2 = 1) &= P(X_1 = 1 \cap X_2 = 1) + P(X_1 = 0 \cap X_2 = 1) \\ &= \frac{2(n-2)+2}{n(n-1)} = \frac{2}{n} \end{aligned}$$

$$\begin{aligned} P(X_2 = 0) &= P(X_1 = 0 \cap X_2 = 0) + P(X_1 = 1 \cap X_2 = 0) \\ &= \frac{n-2}{n} \end{aligned}$$

et on observe que X_2 suit la même loi que X_1. En revanche, la loi conjointe n'étant pas égale à la loi produit, les variables X_1 et X_2 ne sont pas indépendantes:

$$P(X_1 = 1)P(X_2 = 1) = \frac{4}{n^2} \neq \frac{2}{n(n-1)} = P(X_1 = 1 \cap X_2 = 1).$$

(d) La variable A compte le nombre de tickets gagnants obtenus après deux pioches, on a donc clairement

$$A = X_1 + X_2$$

et $A(\Omega) = \{0, 1, 2\}$. Il suit que

$$\begin{aligned} P(A = 0) &= P(X_1 = 0 \cap X_2 = 0) = \frac{(n-2)(n-3)}{n(n-1)} \\ P(A = 2) &= P(X_1 = 1 \cap X_2 = 1) = \frac{2}{n(n-1)} \\ P(A = 1) &= P(X_1 = 1 \cap X_2 = 0) + P(X_1 = 0 \cap X_2 = 1) \\ &= \frac{4(n-1)}{n(n-1)} \end{aligned}$$

Il suit que

$$E(A) = \frac{4(n-2)}{n(n-1)} + 2 \times \frac{2}{n(n-1)} = \frac{4}{n}.$$

(e) Soit $X \hookrightarrow \mathcal{H}(2, 2/n, n)$, où cette loi est celle de la Partie I. D'après la formule, on a donc, pour $k \in \{0, 1, 2\}$,

$$P(X = k) = \frac{\binom{2}{k}\binom{n-2}{2-k}}{\binom{n}{2}} = \frac{2!(n-2)!2!(n-2)!}{k!(2-k)!(2-k)!(n+k-4)!n!} = \frac{4(n-2)!}{[(2-k)!]^2 n(n-1)k!(n+k-4)!}$$

ce qui donne le tableau suivant

k	0	1	2
$P(X = k)$	$\dfrac{(n-2)(n-3)}{n(n-1)}$	$\dfrac{4(n-2)}{n(n-1)}$	$\dfrac{2}{n(n-1)}$

On a bien la même loi que celle trouvée pour A ce qui répond bien à la question posée.

(f) Soit $k \in [\![0;2]\!]$. Sachant $A = k$, il reste $2-k$ tickets gagnants dans l'urne qui contient $n-2-2 = n-4$ tickets en tout. On a donc

$$P_{[A=k]}(Y_1 = 1) = \frac{2-k}{n-4}, \qquad P_{[A=k]}(Y_1 = 0) = \frac{n+6-k}{n-4}.$$

(g) Par la formule des probabilités totales appliquée au s.c.e $\{[A=k] : \ 0 \leq k \leq 2\}$, on obtient

$$\begin{aligned}
P(Y_1 = 1) &= \sum_{k=0}^{2} P_{[A=k]}(Y_1 = 1) P(A = k) \\
&= \frac{1}{n-4} \sum_{k=0}^{2} (2-k) P(A=k) = \frac{1}{n-4} \left(\sum_{k=0}^{2} 2P(A=k) - \sum_{k=0}^{2} kP(A=k) \right) \\
&= \frac{1}{n-4} \left(2 - E(A)\right),
\end{aligned}$$

car

$$\sum_{k=0}^{2} P(A=k) = 1 \qquad \text{et} \qquad \sum_{k=0}^{2} kP(A=k) = E(A).$$

On a bien (car $p = 2$)

$$Y_1 \hookrightarrow \mathcal{B}\left(\frac{p - E(A)}{n - 2p}\right).$$

(h) On admet que Y_2 suit la même loi que Y_1. Comme, clairement $B = Y_1 + Y_2$, on a

$$E(B) = E(Y_1) + E(Y_2) = 2E(Y_1) = 2P(Y_1 = 1) = \frac{2}{n-4} \times \left(2 - \frac{4}{n}\right) = \frac{4(n-2)}{n(n-4)}$$

ce qui est bien la formule attendue.

(i) On a obtenu

$$E(A) = \frac{4}{n} < \frac{4(n-2)}{n(n-4)} = E(B)$$

car $n-2 > n-4 > 0$ et donc $\frac{n-2}{n-4} > 1$. C'est encore la deuxième stratégie la plus avantageuse.

Partie IV : Cas général

On admet dans cette partie que toutes les variables X_i suivent la même loi $\mathcal{B}\left(\frac{p}{n}\right)$ (pour $1 \leq i \leq p$).

(9) Comme précédemment, on a

$$A = X_1 + X_2 + ... + X_p.$$

Il suit, par linéarité de l'espérance que

$$E(A) = E(X_1) + ... + E(X_p) = pE(X_1) = pP(X_1 = 1) = p \times \frac{p}{n} = \frac{p^2}{n}.$$

(10) Afin de déterminer la loi de A, on modélise l'expérience par un tirage simultané des p tickets dans l'urne qui en contient toujours n.

(a) Il y a $\binom{n}{p}$ façons de choisir p tickets dans l'urne qui en contient n.

(b) Il y a $\binom{p}{k}$ façons de choisir les k tickets gagnants puis $\binom{n-p}{p-k}$ façons de compléter avec $p-k$ tickets perdants, donc $\binom{p}{k} \times \binom{n-p}{p-k}$ tirages de p tickets dont k sont gagnants.

(c) Il suit des deux dénombrements précédents que

$$P(A = k) = \frac{\dbinom{p}{k} \times \dbinom{n-p}{p-k}}{\dbinom{n}{p}}$$

et on a bien la formule de la Partie I qui définit $A \hookrightarrow \mathcal{H}(p, p/n, n)$.

(11) Comme précédemment, sachant $[A = k]$, il reste $p - k$ tickets gagnants dans l'urne qui contient maintenant $n - p - p = n - 2p$ tickets en tout. Ainsi, comme par hypothèse du texte, toutes les vloi conditionnelles des Y_i sachant $[A = k]$ sont les mêmes lois de Bernoulli, on a

$$P_{[A=k]}(Y_i = 1) = P_{[A=k]}(Y_1 = 1) = \frac{p-k}{n-2p}.$$

(12) Soit $i \in [\![1; p]\!]$. Par la formule des probabilités totales appliquée au s.c.e $\{[A = k];\ k \in [\![0; p]\!]\}$, on a

$$\begin{aligned} P(Y_i = 1) &= \sum_{k=0}^{p} P_{[A=k]}(Y_i = 1)P(A = k) \\ &= \sum_{k=0}^{p} \frac{p-k}{n-2p} \times \frac{\binom{p}{k} \times \binom{n-p}{p-k}}{\binom{n}{p}} \\ &= \frac{1}{(n-2p)\binom{n}{p}} \sum_{k=0}^{p} (p-k)\binom{p}{k}\binom{n-k}{p-k}. \end{aligned}$$

(13) Par linéarité de l'espérance

$$\begin{aligned} E(B) &= E(Y_1 + ... + Y_p) = E(Y_1) + ... + E(Y_p) = pE(Y_1) = pP(Y_1 = 1) \\ &= p\left(\frac{1}{(n-2p)} \sum_{k=0}^{p} (p-k)P(A = k)\right) = \frac{p}{(n-2p)}\left(\sum_{k=0}^{p} pP(A = k) - \sum_{k=0}^{p} kP(A = k)\right) \\ &= \frac{p}{(n-2p)}(p - E(A)) = \frac{p}{(n-2p)}\left(p - \frac{p^2}{n})\right) \\ &= \frac{p^2(n-p)}{n(n-2p)}, \end{aligned}$$

comme attendu.

(14) Comme $0 < n - 2p < n - p$, on a $(n - p)/(n - 2p) > 1$ et il suit que

$$E(A) = \frac{p^2}{n} < \frac{p^2}{n} \times \frac{n-p}{n-2p} = E(B)$$

donc la deuxième stratégie est bien, en toute généralité, la plus avantageuse.