Semestre 4, Année 2005-2006 SESSION 1

UE : PIN401

Date : 29 Mai 2006, Durée : 3h

Texte (en italiques) et corrigé (en roman)

Partie I - ALGEBRE

Soit N un nombre entier strictement positif. On désigne par E_N le $\mathbb{R}$-espace vectoriel des fonctions continues affines par morceaux sur $[0, 2^N]$, avec "raccords" aux points $0, 1, 2, 3, ..., 2^N$, ce qui signifie que le graphe d'un élément de E_N est une ligne brisée dont les "brisures" éventuelles sont exactement au dessus des points $0, 1, 2, 3, ..., 2^N$. Dessiner le graphe d'une telle fonction (par exemple pour $N = 3$).

Le graphe d'une telle fonction est représenté sur la figure 1 ci-dessous.

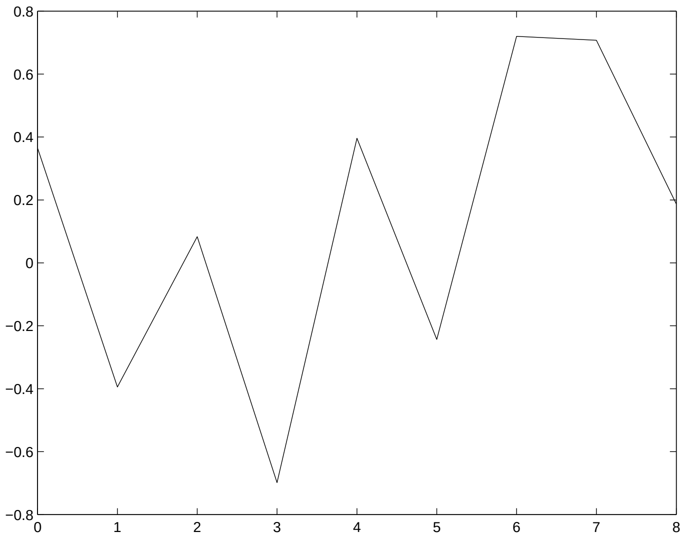

FIG. 1 – Le graphe d'un élément de E_N pour $N = 3$

I.1 *Pour $j = 0, 1, 2, 3, ..., 2^N$, on note u_j l'élément de E_N défini par*

$$\forall\, t \in [0, 2^N]\,,\ u_j(t) := \max(0\,,\, 1 - |t - j|)\,.$$

Montrer que le système $(u_0, u_1, u_2, u_3, ..., u_{2^N})$ constitue une base de E_N et qu'une fonction f quelconque de E_N s'écrit :

$$\forall\, t \in [0, 2^N]\,,\ f(t) = \sum_{j=0}^{2^N} f(j)\, u_j(t)\,;$$

que valent les coordonnées de f dans la base $(u_0, u_1, u_2, u_3, ..., u_{2^N})$? Lorsque l'on choisit $N = 3$, dessiner en trait plein et sur un même graphique les graphes des 9 fonctions $u_0, u_1, u_2, u_3, u_4, u_5, u_6, u_7, u_8$.

Si l'on a

$$\sum_{j=0}^{2^N} \lambda_j u_j(t) \equiv 0 \ \forall t \in [0, 2^N]\,,$$

on trouve, en spécialisant t aux points $0, 1, 2, 3, ..., 2^N$,

$$\lambda_0 = \cdots = \lambda_{2^N} = 0\,.$$

Ceci montre que le système $(u_0, ..., u_{2^N})$ est bien un système libre.
Il est d'autre part évident que la donnée d'un élément de E_N dépend de $2^N + 1$ degrés de liberté (en l'occurrence les valeurs qu'il convient d'interpoler aux points entiers pour construire une fonction affine par morceaux sur $[0, 2^N]$ avec "raccords" au dessus des entiers $0, 1, ..., 2^N$). La dimension de E_N est donc au plus égale à $2^N + 1$; le système libre $(u_0, u_1, ..., u_{2^N})$ qui contient $2^N + 1$ éléments est donc une base de E_N.
Si f est un élément de E_N de la forme

$$t \longmapsto f(t) = \sum_{j=0}^{2^N} \lambda_j\, u_j(t)\,,$$

on trouve, en spécifiant $t = j$ $(j = 0, ..., 2^N)$

$$f(j) = \lambda_j\,.$$

Le vecteur de coordonnées de f dans la base $(u_0, u_1, ..., u_{2^N})$ est donc le vecteur $(f(0), f(1), ..., f(2^N))$.
Sur la figure 2 ci-dessous, on a représenté (en trait plein) les graphes des 9 fonctions $u_0, ..., u_8$ pour $N = 3$.

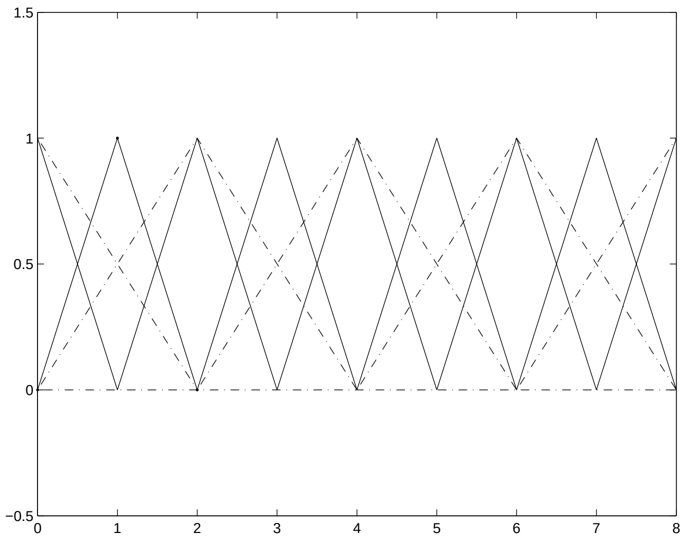

FIG. 2 – Les graphes de $u_0, u_1, ..., u_{2^N}$ pour $N = 3$

I.2 *On considère le sous-espace de* E_N *engendré par les* $2^{N-1} + 1$ *fonctions* v_k, $k = 0, 1, 2, 3, ..., 2^{N-1}$ *définies par*

$$v_k(t) := \max(0, 1 - |t/2 - k|)\,,\ k = 0, 1, 2, 3, ..., 2^{N-1}\,.$$

Lorsque l'on choisit $N = 3$, dessiner en trait pointillé cette fois (ou d'une autre couleur) sur le même graphique qu'au **I.1** *les graphes des 5 fonctions v_k, $k = 0, 1, 2, 3, 4$.*

On a représenté ces 5 graphes sur la figure 2 ci-dessus.

I.3 *Si l'espace E_N est muni du produit scalaire défini par*

$$\langle f \, , \, g \rangle := \int_0^{2^N} f(t)g(t)dt \, ,$$

vérifier que la matrice carrée G_N (de taille $(2^{N-1}+1, 2^{N-1}+1)$, dite de Gram ou de masse) ayant pour entrées les nombres $\langle v_k \, , \, v_l \rangle$, $0 \leq k, l \leq 2^{N-1}$ est donnée par $2G_N$, où :

$$G_N := \begin{pmatrix} 1/3 & 1/6 & 0 & \cdots & 0 & 0 \\ 1/6 & 2/3 & 1/6 & 0 & \cdots & 0 \\ 0 & 1/6 & 2/3 & 1/6 & \cdots & 0 \\ \vdots & . & . & \cdots & \cdots & \vdots \\ 0 & \cdots & 0 & 1/6 & 2/3 & 1/6 \\ 0 & \cdots & 0 & 0 & 1/6 & 1/3 \end{pmatrix}$$

On pourra faire ce calcul dans un premier temps en fixant $N = 3$ (en s'aidant des graphiques tracés au **I.1** *et* **I.2***) puis expliquer ensuite comment on passe au cas N quelconque.*

Nota : *il y avait dans cette question une erreur dans la version initiale de l'énoncé, erreur rectifiée ici ; il en a été bien sûr tenu compte lors de la correction du problème.*

Le calcul de $\langle v_0, v_0 \rangle$ donne

$$\langle v_0, v_0 \rangle = \int_0^2 (1 - t/2)^2 dt = 2 \int_0^1 (1-u)^2 du = \frac{2}{3} = 2 \times \frac{1}{3}$$

Par symétrie, c'est aussi la valeur de $\langle v_{2^{N-1}}, v_{2^{N-1}} \rangle$. Pour $i = 0, ..., 2^{N-1} - 1$, on a

$$\langle v_i \, , \, v_{i+1} \rangle = \int_0^2 (t/2)(1-t/2) \, dt = 2 \int_0^1 u(1-u) \, du = 2 \times (1/2 - 1/3) = 2 \times \frac{1}{6}$$

et

$$\langle v_i, \, , \, v_j \rangle = 0 \quad \forall j > i + 1 \, .$$

Enfin, on a, pour $i = 1, ..., 2^{N-1} - 1$,

$$\langle v_i \, , \, v_i \rangle = \int_0^4 v_1^2(t)dt = 2 \int_0^2 v_1^2(t)dt = 2\langle v_{2^{N-1}}, v_{2^{N-1}} \rangle = 2 \times \frac{2}{3} \, .$$

Comme la matrice de Gram est symétrique, elle est bien de la forme mentionnée dans l'énoncé.

I.4 *Vérifier que le système $(v_0, v_1, v_2, v_3, ..., v_{2^{N-1}})$ constitue un système libre de E_N. Quelle est la dimension du sous-espace vectoriel F_N de E_N engendré par les fonctions $v_0, v_1, v_2, v_3, ..., v_{2^{N-1}}$?*

C'est exactement la même preuve qu'au **I.1** que l'on utilise our montrer que le système $(v_0, ..., v_{2^{N-1}})$ est libre dans F_N. Comme il est par hypothèses (définition de F_N) générateur dans F_N, c'est bien une base de ce $\mathbb{R}$-espace vectoriel. L'espace F_N est donc de dimension le cardinal de cette famille $\{v_0, ..., v_{2^{N-1}}\}$, soit $\dim F_N = 2^{N-1}+1$.

I.5 *Pourquoi la matrice G_N est-elle inversible ? On ne demande pas de calculer l'inverse, que l'on note G_N^{-1}.*

La matrice G_N est inversible puisque le seul vecteur $w = x_0 v_0 + \cdots + x_{2^{N-1}} v_{2^{N-1}}$ de F_N tel que

$$\langle w\,,\, v_j \rangle = 0\,,\ \forall j = 0, ..., 2^{N-1}$$

est le vecteur de coordonnées $(x_0, ..., x_{2^{N-1}}) = (0, ..., 0)$ (la forme $\langle\,,\,\rangle$ étant en effet une forme définie positive). Ceci prouve que le noyau de l'application linéaire de $\mathbb{R}^{2^{N-1}+1}$ dans lui-même dont la matrice dans la base canonique est G_N est réduit au vecteur nul, donc que G_N est une matrice inversible.

I.6 *Soit $f \in E_N$. Exprimer en fonction des nombres $\langle f\,,\, v_k \rangle$, $k = 0, 1, 2, 3, ..., 2^{N-1}$ et de la matrice G_N^{-1} le vecteur*

$$\begin{pmatrix} x_0 \\ \vdots \\ x_{2^{N-1}} \end{pmatrix}$$

des coordonnées suivant $(v_0, v_1, v_2, v_3, ..., v_{2^{N-1}})$ de la projection orthogonale de f sur F_N.

La projection orthogonale $w = x_0 v_0 + \cdots + x_{2^{N-1}} v_{2^{N-1}}$ de f sur F_N est caractérisée par la propriété

$$\forall g \in F_N\,,\ \langle f - w\,,\, g \rangle = 0\,,$$

ou encore, puisque $v_0, ..., v_{2^{N-1}}$ engendrent F_N,

$$\forall j = 0, ..., 2^{N-1}\,,\ \langle w\,,\, v_j = \langle f\,,\, v_j \rangle\,.$$

Cette liste de conditions se lit matriciellement

$$2G_N \cdot \begin{pmatrix} x_0 \\ \vdots \\ x_{2^{N-1}} \end{pmatrix} = \begin{pmatrix} \langle f\,,\, v_0 \rangle \\ \vdots \\ \langle f\,,\, v_{2^{N-1}} \rangle \end{pmatrix}\,,$$

soit encore

$$\begin{pmatrix} x_0 \\ \vdots \\ x_{2^{N-1}} \end{pmatrix} = \frac{1}{2} G_N^{-1} \cdot \begin{pmatrix} \langle f\,,\, v_0 \rangle \\ \vdots \\ \langle f\,,\, v_{2^{N-1}} \rangle \end{pmatrix}\,.$$

I.7 *On fixe à partir de maintenant et jusquà la fin du problème $N = 2$. Montrer (sans calcul) que $1/3$ est valeur propre de G_2. Calculer le polynôme caractéristique de G_3 et montrer qu'il admet trois racines réelles distinctes. En déduire que G_2 est diagonalisable. Pouvait-on avant même de se lancer dans les calculs prédire ce résultat ?*

La matrice G_2 est la matrice

$$\begin{pmatrix} 1/3 & 1/6 & 0 \\ 1/6 & 2/3 & 1/6 \\ 0 & 1/6 & 1/3 \end{pmatrix}\,.$$

Le polynôme caractéristique de cette matrice est

$$P(X) = \begin{vmatrix} 1/3 - X & 1/6 & 0 \\ 1/6 & 2/3 - X & 1/6 \\ 0 & 1/6 & 1/3 - X \end{vmatrix}\,.$$

Si l'on prend $X = 1/3$, le déterminant représentant $P(X) = P(1/3)$ a sa première et sa troisième colonnes égales ; il est donc nul et $1/3$ est racine du polynôme caractéristique de G_2, donc valeur propre de G_2.

Le calcul du polynôme caractéristique de G_2 donne (on développe par rapport à la première colonne)

$$\begin{aligned} P(X) &= (1/3 - X)\Big[(1/3 - X)(2/3 - X) - 1/36\Big] - (1/36)(1/3 - X) \\ &= (1/3 - X)(X^2 - X + 1/6)\,. \end{aligned}$$

Comme le discriminant du trinôme $X^2 - X + 1/6$ vaut $1 - 4/6 = 1/3 > 0$, ce trinôme admet deux racines réelles distinctes

$$x = \frac{1 \pm 3^{-1/2}}{2}\,,$$

distinctes de $1/3$. Le polynôme P est donc scindé et la matrice G_2 est par conséquent diagonalisable. On aurait pu prévoir ce résultat car il s'agit d'une matrice symétrique réelle, donc diagonalisable sur $\mathbb{R}$ d'après le cours.

I.8 *Soient $(\alpha_0, \alpha_1, \alpha_2)$, $(\beta_0, \beta_1, \beta_2)$, $(\gamma_0, \gamma_1, \gamma_2)$, trois vecteurs propres dans $\mathbb{R}^3$ de la matrice G_2 correspondant à des valeurs propres distinctes de G_2 (on ne demande pas de calculer les $\alpha_j, \beta_j, \gamma_j$). En vous souvenant que G_3 est la matrice de Gram (ou de masse) du système (v_0, v_1, v_2), construire à partir des données dont vous disposez (les trois fonctions v_0, v_1, v_2, et les 9 nombres $\alpha_j, \beta_j, \gamma_j$ pour $j = 1, 2, 3$) une base orthonormée du sous-espace F_2. Quel procédé algorithmique aurait-il aussi permis cette construction ? Quelle est des deux stratégies la meilleure (surtout si l'on travaille avec N quelconque et non plus $N = 2$) ?*

La matrice G_2, matrice symétrique réelle, s'écrit sous la forme

$$G_2 = PDP^{-1}\,,$$

où $P^{-1} = {}^tP$. Dans la base constituée des vecteurs $\alpha_0 v_0 + \alpha_1 v_1 + \alpha_2 v_2$, $\beta_0 v_0 + \beta_1 v_1 + \beta_2 v_2$, $\gamma_0 v_0 + \gamma_1 v_1 + \gamma_2 v_2$, la matrice de la forme bilinéaire $\langle\,,\,\rangle$ (restreinte au sous-espace F_2) s'exprime donc comme une matrice diagonale, ce qui prouve que les trois vecteurs $\alpha_0 v_0 + \alpha_1 v_1 + \alpha_2 v_2$, $\beta_0 v_0 + \beta_1 v_1 + \beta_2 v_2$, $\gamma_0 v_0 + \gamma_1 v_1 + \gamma_2 v_2$ sont deux à deux orthogonaux relativement au produit scalaire $\langle\,,\,\rangle$. On obtient donc une base orthonormée de F_2 en considérant les trois éléments de F_2 :

$$\frac{\alpha_0 v_0 + \alpha_1 v_1 + \alpha_2 v_2}{\sqrt{\Big\langle \alpha_0 v_0 + \alpha_1 v_1 + \alpha_2 v_2\,,\, \alpha_0 v_0 + \alpha_1 v_1 + \alpha_2 v_2 \Big\rangle}}$$
$$\frac{\beta_0 v_0 + \beta_1 v_1 + \beta_2 v_2}{\sqrt{\Big\langle \beta_0 v_0 + \beta_1 v_1 + \beta_2 v_2\,,\, \beta_0 v_0 + \beta_1 v_1 + \beta_2 v_2 \Big\rangle}}$$
$$\frac{\gamma_0 v_0 + \gamma_1 v_1 + \gamma_2 v_2}{\sqrt{\Big\langle \gamma_0 v_0 + \gamma_1 v_1 + \gamma_2 v_2\,,\, \gamma_0 v_0 + \gamma_1 v_1 + \gamma_2 v_2 \Big\rangle}}\,.$$

C'est le procédé algorithmique de Gram-Schmidt qui aurait aussi permis la construction à partir de v_0, v_1, v_2 d'une base orthonormée (relativement au produit scalaire $\langle\,,\,\rangle$) de F_2. Il consiste à prendre

$$w_0 = \frac{v_0}{\sqrt{\langle v_0\,,\, v_0\rangle}}$$

puis à construire

$$\tilde{v}_1 := v_1 - \langle v_1, w_0\rangle w_0$$

et

$$w_1 = \frac{\tilde{v}_1}{\sqrt{\langle \tilde{v}_1\,,\,\tilde{v}_1\rangle}}\,,$$

puis

$$\tilde{v}_2 := v_2 - \langle v_2, w_0\rangle w_0 - \langle v_2, w_1\rangle w_1$$

et finalement

$$w_2 = \frac{\tilde{v}_2}{\sqrt{\langle \tilde{v}_2\,,\,\tilde{v}_2\rangle}}\,.$$

Si $N > 2$, cette dernière démarche (algorithmique et directe) est bien sûr moins coûteuse que celle qui passe par la diagonalisation de la matrice symétrique G_N (qui oblige, elle, à la recherche des valeurs propres, puis la détermination des vecteurs propres et n'est d'ailleurs sans ambigüité que si le polynôme caractéristique de G_N est scindé, ce qu'il faudrait au préalable prouver pour disposer immédiatement avec $2^{N-1}+1$ vecteurs propres correspondant à des valeurs propres distinctes et que l'on normalise relativement au produit scalaire comme ci-dessus dans le cas particulier $N = 2$, d'une base orthonormée de F_N).

Partie II - PROBABILITES

Exercice A

Zigue et Willie jouent aux fléchettes. La cible est circulaire, de rayon $R = 2$. Le point d'impact d'un lancer de Zigue est représenté par la variable aléatoire conjointe $Z = (X, Y)$, de densité

$$f(x,y) \;= \frac{1}{2\pi\sigma^2}\, e^{-\dfrac{x^2+y^2}{2\,\sigma^2}}$$

avec $\sigma = \sqrt{2}$ et, pour un lancer de Willie, par la variable aléatoire conjointe $W = (U, V)$, de densité

$$g(u,v) \;= \frac{1}{2\pi}\, e^{-\sqrt{u^2+v^2}}\,.$$

A.1 *Les variables aléatoires marginales X et Y de Z sont elles indépendantes ?*
La variable X a pour densité

$$f(x) = \int f(x,y)\,dy = \frac{e^{-x^2/(2\sigma^2)}}{\sqrt{2\pi}\,\sigma}\int_{\mathbb{R}} \frac{e^{-y^2/(2\sigma^2)}}{\sqrt{2\pi}\,\sigma}\,dy = \frac{1}{\sqrt{2\pi}\,\sigma}\,e^{-x^2/(2\sigma^2)}$$

tandis que la variable Y a pour densité

$$g(y) = \int f(x,y)\,dx = \frac{e^{-y^2/(2\sigma^2)}}{\sqrt{2\pi}\,\sigma}\int_{\mathbb{R}} \frac{e^{-x^2/(2\sigma^2)}}{\sqrt{2\pi}\,\sigma}\,dx = \frac{1}{\sqrt{2\pi}\,\sigma}\,e^{-y^2/(2\sigma^2)}\,.$$

Comme le couple (X, Y) est à densité $f(x,y) = f(x)g(y)$, les variables X et Y sont indépendantes.

A.2 *Les variables aléatoires marginales U et V de W sont elles indépendantes ?*

La variable U a pour densité

$$r(u) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-\sqrt{u^2+v^2}}\, dv$$

tandis que la variable V a pour densité

$$s(v) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-\sqrt{u^2+v^2}}\, du\,.$$

La probabilité de $\{U \in [0, du]\}$ vaut $r(0)\, du = du/\pi$, tandis que celle de l'évènement $\{V \in [0, dv]\}$ vaut dv/π et que celle de $\{(U, V) \in [0, du] \times [0, dv]\}$ vaut

$$(1/(2\pi))\, dudv \neq ((1/\pi)\, du) \times ((1/\pi)\, dv)\,.$$

Les deux variables U et V ne sont donc pas indépendantes.

A.3 On note respectivement $F(r)$, pour Zigue, et $G(r)$, pour Willie, la probabilité d'un impact à une distance au plus r (≥ 0) du centre de la cible. Montrer que

$$F(r) \;=\; 1 \;-\; e^{-\dfrac{r^2}{2\sigma^2}} \qquad , \qquad G(r) = 1 - (1 + r)e^{-r}\,.$$

On a

$$\begin{aligned} F(r) \;&=\; \int\!\!\int_{D(0,r)} f(x, y)\, dxdy = \frac{2\pi}{2\pi\sigma^2} \int_0^r \rho e^{-\rho^2/(2\sigma^2)}\, d\rho = \frac{1}{2\sigma^2} \int_0^{r^2} e^{-t/(2\sigma^2)}\, dt \\ &=\; 1 - \exp\Big(-\frac{r^2}{2\sigma^2}\Big) \end{aligned}$$

(en passant en cordonnées polaires). On a aussi (toujours en passant en coordonnées polaires)

$$\begin{aligned} G(r) \;&=\; \int\!\!\int_{D(0,r)} g(u, v)\, dudv = \int_0^r \rho e^{-\rho}\, d\rho = \Big[-\rho e^{-\rho}\Big]_0^r + \int_0^r e^{-\rho}\, d\rho \\ &=\; 1 - (1 + r)e^{-r}\,. \end{aligned}$$

A.4 *Quelle est la probabilité de voir Zigue atteindre la cible ? Même question pour Willie.*

La probabilité que Zigue atteigne la cible vaut (puisque celle ci est de rayon 2)

$$F(2) = 1 - e^{-1} = 0.6321.$$

La probabilité que Willie atteigne la cible vaut

$$G(2) = 1 - 3e^{-2} = 0.5940.$$

A.5 *On compte 10 points pour un impact dans le cercle de rayon 1 centré au centre de la cible, et 1 point pour un impact dans le reste de la cible. Quels sont les espérances de gain respectives de Zigue et de Willie, sur un seul lancer ?*

L'espérance de gain de Zigue est

$$10F(1) + (F(2) - F(1)) = F(2) + 9F(1) = 2.623\,.$$

L'espérance de gain de Willie est

$$10G(1) + (G(2) - G(1)) = G(2) + 9G(1) = 2.972.$$

A.6 *Après une partie de 36 lancers (chacun), on note S_Z le score de Zigue et S_W le score de Willie. Estimer la probabilité de voir $S_Z \geq 100$, puis estimer la probabilité de voir $S_W \geq 100$ (une table des valeurs de la loi normale réduite est jointe à ce texte en page 4).*

Le score de Zigue à chaque partie est une variable aléatoire s_Z prenant les valeurs 10 (avec la probabilité $F(1)$), 1 (avec la probabilité $F(2)-F(1)$) et 0 (avec la probabilité $1-F(2)$). Cette variable aléatoire s_z a pour espérance 2.623 et pour variance :

$$V(s_Z) = 100F(1) + (F(2) - F(1)) - (2.623)^2 = 15.65.$$

L'écart type de s_Z vaut donc $\sigma_Z = 3.956$.

Le score de Willie à chaque partie est une variable aléatoire s_W prenant les valeurs 10 (avec la probabilité $G(1)$), 1 (avec la probabilité $G(2) - G(1)$) et 0 (avec la probabilité $1 - G(2)$). Cette variable aléatoire s_W a pour espérance 2.972 et pour variance :

$$V(s_W) = 100G(1) + (G(2) - G(1)) - (2.972)^2 = 17.92$$

L'écart type de s_W vaut donc $\sigma_W = 4.233$.

D'après le théorème limite centrale, on peut assimiler les variables

$$\frac{S_Z - 36E[s_Z]}{6\sigma_Z} \text{ et } \frac{S_W - 36E[s_W]}{6\sigma_W}$$

à des variables suivant une loi normale N réduite centrée.

On a

$$\begin{aligned} P(S_Z \geq 100) &= P\Big(N \geq \frac{100 - 36E[s_Z]}{6\sigma_Z}\Big) = P(N \geq 0.2347) = 1 - F(0.2347) \\ P(S_W \geq 100) &= P\Big(N \geq \frac{100 - 36E[s_W]}{6\sigma_W}\Big) = P(N \geq -0.2753) = F(0.2753)\,. \end{aligned}$$

D'après la table, on a donc

$$P(S_Z \geq 100) \in [0.3821, 0.4207]$$

et

$$P(S_W \geq 100) \in [0.5793, 0.6179]\,.$$

Exercice B

Sur les $N = 128$ matches des deux dernières coupes du monde (1998 et 2002), 344 buts ont été marqués, en cumulant les scores des deux équipes pour chaque match et sans compter les séances de tirs au but.

On représente par la variable aléatoire X le nombre de buts par match. On veut savoir quelle est la meilleure loi à associer à cette variable aléatoire X, entre une loi de Poisson ($P(X = n) = \frac{\lambda^n}{n!}e^{-\lambda}$, $\lambda > 0$) et une loi géométrique ($P(X = n) = a^n(1 - a)$, $0 < a < 1$).

B.1 *Donner une estimation de la moyenne de X notée $m = E(X)$ On suppose dorénavant que cette estimation de m est exacte.*

L'estimation de la moyenne de X est (d'après la loi forte des grands nombres)

$$\overline{x} = \frac{344}{128} = \frac{x_1 + \cdots + x_{128}}{128} = 2.687\,,$$

où x_j désigne le nombre de buts marqués lors du j-ème match.

B.2 *Déterminer λ et a en fonction de m.*

L'espérance d'une variable suivant une loi de Poisson de paramètre λ vaut λ. On a donc, si l'on suppose que X suit une loi de Poisson de paramètre λ,

$$\lambda = m = 2.687\,.$$

L'espérance d'une variable X suivant une loi géométrique de paramètre a vaut

$$\sum_{k=0}^{\infty}(1-a)ka^k = a(1-a)\sum_{k=1}^{n} ka^{k-1} = \frac{a}{1-a}\,.$$

On a donc, si X est supposée suivre une loi géométrique de paramètre a,

$$a = \frac{m}{1+m} = \frac{2.687}{3.687} = 0.729.$$

B.3 *On suppose dans un premier temps que X suit la loi géométrique de paramètre a. Calculer sa variance.*

La variance d'une variable aléatoire suivant une loi géométrique de paramètre a vaut

$$\begin{aligned}
(1-a)\sum_{k=0}^{\infty} k^2a^k - \Big(\frac{a}{1-a}\Big)^2 &= (1-a)a^2\sum_{k=2}^{\infty} k(k-1)a^{k-2} \\
&\quad + \frac{a}{1-a} - \Big(\frac{a}{1-a}\Big)^2 \\
&= 2\Big(\frac{a}{1-a}\Big)^2 + \frac{a}{1-a} - \Big(\frac{a}{1-a}\Big)^2 \\
&= \Big(\frac{a}{1-a}\Big)^2 + \frac{a}{1-a} \\
&= \frac{a}{(1-a)^2}\,.
\end{aligned}$$

Comme $a = 0.729$, cette variance vaut ici

$$\sigma_X^2 = 9.926.$$

B.4 Parmi ces 128 matches, 52 n'ont compté que 0, 1 ou 2 buts. On note Y la variable de Bernoulli définie par

$$Y = \begin{cases} 1 & \text{pour } X \leq 2, \\ 0 & \text{pour } X \geq 3. \end{cases}$$

Etudier la variable aléatoire Y (moyenne, écart-type).

On a

$$P(Y=1) = P(X=0)+P(X=1)+P(X=2) = (1-a)(1+a+a^2) = 1-a^3 = 0.613$$

puisque $a = 0.729$. La variable Y suit une loi de Bernouilli de paramètre $p = 0.613$, donc d'espérance $p = 0.613$ et de variance $p(1-p) = 0.237$, donc d'écart-type $\sigma_Y = 0.487$.

B.5 *On note T le nombre de matches de score ≤ 2 buts. Evaluer la probabilité $P(45 \leq T \leq 65)$ (une table des valeurs de la loi normale réduite est jointe à ce texte en page 4).*

Selon le théorème limite centrale, on peut considérer que la variable

$$N = \frac{T - 128E[Y]}{8\sqrt{2}\,\sigma_Y} = \frac{T - 78.46}{5.51}$$

Dire que $T \in [45, 65]$ équivaut à dire que

$$N \in [-6.07, -2.44]\,.$$

La probabilité que N soit dans cet intervalle est majorée par $1 - F(2.5) = 0.0062$.

B.6-B.7-B.8 *Reprendre les questions* **B.3-B.4-B.5** *pour X suivant la loi de Poisson de paramètre λ, et proposer la meilleure loi à associer à la variable aléatoire X .*

Solution de B.6. La variance d'une loi de Poisson de paramètre λ vaut $V = \lambda$. Ici, on a donc $V = \lambda = m = 2.687$.

Solution de B.7. Le paramètre p de la variable de Bernouilli Y vaut maintenant

$$p = e^{-2.687}(1 + 2.687 + (2.687)^2/2) = 0.497.$$

La variance de Y vaut $p(1-p) = 0.25$ et l'écart-type vaut donc 0.5.

Solution de B.8. Selon le théorème limite centrale, on peut considérer que la variable

$$N = \frac{T - 128E[Y]}{8\sqrt{2}\,\sigma_Y} = \frac{T - 63.61}{5.66}$$

Dire que $T \in [45, 65]$ équivaut à dire que

$$N \in [-3.29, 0.245]\,.$$

La probabilité que N tombe dans cet intervalle est entre 0.5793 et 0.6179 ; elle est de l'ordre de 0.6.

Conclusion : comme on a constaté 52 matches avec un score de 2 buts ou moins et que $52 \in [45, 65]$, il est raisonnable (au vu des résultats comparés des questions **B.5** et **B.8**) de penser que la loi de Poisson est la meilleure loi à associer à X.

Table de la loi normale réduite de fonction de répartition

$$F(x) \ = \ \frac{1}{\sqrt{2\pi}} \ \int_{-\infty}^{x} \ e^{-\frac{\xi^2}{2}} \ d\xi$$

x	F(x)	x	F(x)	x	F(x)
0.0	0.5	1.0	0.8413	2.0	0.9773
0.1	0.5398	1.1	0.8643	2.1	0.9821
0.2	0.5793	1.2	0.8849	2.2	0.9861
0.3	0.6179	1.3	0.9032	2.3	0.9893
0.4	0.6554	1.4	0.9192	2.4	0.9918
0.5	0.6915	1.5	0.9332	2.5	0.9938
0.6	0.7257	1.6	0.9452	2.6	0.9953
0.7	0.7580	1.7	0.9554	2.7	0.9965
0.8	0.7881	1.8	0.9641	2.8	0.9974
0.9	0.8159	1.9	0.9713	2.9	0.9981