Université des Sciences et Technologies de Lille
U.F.R. de Mathématiques Pures et Appliquées

M1-MEFM **Année 2010**

Corrigé de l'examen du 5 janvier 2011

Ex 1. *En guise de problème de synthèse*

Soit $(\Omega, \mathcal{F}, (P_\theta)_{\theta\in\Theta})$ un modèle statistique où $\Theta =]0, +\infty[$. On note $X_1, \dots, X_n$ un échantillon associé à ce modèle, les X_i étant des v.a. positives ayant pour densité sous P_θ la fonction :

$$t \mapsto f(t, \theta) := \frac{2t}{\theta^2}\mathbf{1}_{[0,\theta]}(t).$$

1) La densité de la loi de X_1 sous P_θ étant nulle en dehors de $[0, \theta]$, on a $0 \leqslant X_1 \leqslant \theta$ P_θ-presque sûrement. Ainsi X_1 est P_θ-p.s. bornée donc P_θ intégrable.

$$\mathbf{E}_\theta X_1 = \int_{-\infty}^{+\infty} tf(t, \theta)\,\mathrm{d}t = \frac{2}{\theta^2}\int_0^\theta t^2\,\mathrm{d}t = \frac{2}{\theta^2}\frac{\theta^3}{3} = \frac{2}{3}\theta.$$

Ce qui précède est valide pour tout $\theta \in]0, +\infty[$.

Par linéarité de l'espérance et équidistribution sous P_θ des X_i, on a

$$\forall\theta \in \Theta, \quad \mathbf{E}_\theta\Big(\frac{3}{2}\overline{X}\Big) = \frac{3}{2}\mathbf{E}_\theta\left(\frac{1}{n}\sum_{i=1}^n X_i\right) = \frac{3}{2n}(n\mathbf{E}_\theta X_1) = \theta,$$

ce qui montre que la statistique $\frac{3}{2}\overline{X}$ est un estimateur sans biais de θ.

D'autre part en appliquant la loi forte des grands nombres sur chacun des espaces probabilisés $(\Omega, \mathcal{F}, P_\theta)$ aux variables aléatoires P_θ-i.i.d. et intégrables X_i, on obtient :

$$\forall\theta \in \Theta, \quad \overline{X} \xrightarrow[n\to+\infty]{P_\theta-\text{p.s.}} \mathbf{E}_\theta X_1 = \frac{2}{3}\theta,$$

d'où

$$\forall\theta \in \Theta, \quad \frac{3}{2}\overline{X} \xrightarrow[n\to+\infty]{P_\theta-\text{p.s.}} \theta,$$

l'estimateur $\frac{3}{2}\overline{X}$ est donc fortement consistant.

Les v.a. X_i étant P_θ-p.s. bornées sont de carré P_θ-intégrable et par linéarité, il en va de même pour $\frac{3}{2}\overline{X}$. Comme l'estimateur $\frac{3}{2}\overline{X}$ est sans biais, son erreur quadratique moyenne est égale à sa variance :

$$\mathrm{EQM}_\theta\left(\frac{3}{2}\overline{X}\right) = \mathbf{E}_\theta\Big(\frac{3}{2}\overline{X} - \theta\Big)^2 = \mathrm{Var}_\theta\left(\frac{3}{2}\overline{X}\right) = \frac{9}{4}\,\mathrm{Var}_\theta\,\overline{X} = \frac{9}{4n}\,\mathrm{Var}_\theta\,X_1.$$

La variance de X_1 peut se calculer comme suit :

$$\begin{aligned}\operatorname{Var}_\theta X_1 = \mathbf{E}_\theta X_1^2 - (\mathbf{E}_\theta X_1)^2 &= \int_{-\infty}^{+\infty} t^2 f(t,\theta)\,\mathrm{d}t - \left(\frac{2}{3}\theta\right)^2 \\ &= \frac{2}{\theta^2}\int_0^\theta t^3\,\mathrm{d}t - \frac{4}{9}\theta^2 \\ &= \frac{2}{\theta^2}\frac{\theta^4}{4} - \frac{4}{9}\theta^2 = \frac{\theta^2}{18}.\end{aligned}$$

Finalement,

$$\forall \theta \in \Theta, \quad \mathrm{EQM}_\theta\left(\frac{3}{2}\overline{X}\right) = \frac{\theta^2}{8n}.$$

2) Pour construire un intervalle de confiance pour θ basé sur l'estimateur $\frac{3}{2}\overline{X}$, il suffit de construire un intervalle de confiance pour $\theta' = \frac{2}{3}\theta = \mathbf{E}_\theta X_1$ et de multiplier les bornes par $\frac{3}{2}$. On est ainsi ramené à l'estimation par intervalle de confiance d'une espérance inconnue dans le cas d'un « grand » échantillon, exemple traité en cours comme application du théorème limite central avec autonormalisation.

D'après ce théorème, pour des X_i de carré intégrable (ce qui est vrai ici pour tout θ), en notant S^2 la variance empirique de l'échantillon, on a pour tout $\theta \in \Theta$,

$$T_n = \frac{\sqrt{n}}{\sqrt{S^2}}\left(\overline{X} - \theta'\right) \xrightarrow[n\to+\infty]{P_\theta-\text{loi}} Y$$

où Y suit la loi gaussienne $\mathfrak{N}(0,1)$.

En approximant pour n grand la loi de T_n par celle de Y, on peut donc écrire pour tout $\theta \in \Theta$:

$$P_\theta(|T_n| \leqslant t) \simeq P_\theta(|Y| \leqslant t) = 2\Phi(t) - 1,$$

où Φ est la fonction de répartition de la loi $\mathfrak{N}(0,1)$. Pour $t = 1{,}96$, $2\Phi(t) - 1 = 0{,}95$, ce qui nous donne en négligeant l'erreur d'approximation gaussienne :

$$P_\theta\left(\overline{X} - \frac{1{,}96S}{\sqrt{n}} \leqslant \theta' \leqslant \overline{X} + \frac{1{,}96S}{\sqrt{n}}\right) = 0{,}95.$$

On en déduit que

$$P_\theta\left(\frac{3}{2}\overline{X} - \frac{3}{2} \times \frac{1{,}96S}{\sqrt{n}} \leqslant \theta \leqslant \frac{3}{2}\overline{X} + \frac{3}{2} \times \frac{1{,}96S}{\sqrt{n}}\right) = 0{,}95.$$

Ceci étant vrai pour tout $\theta \in \Theta$, on peut donc proposer comme intervalle de confiance au niveau (approximatif) 95% pour θ :

$$I_n(\omega) = \left[\frac{3}{2}\overline{X}(\omega) - \frac{5{,}88S(\omega)}{2\sqrt{n}}\ ;\ \frac{3}{2}\overline{X}(\omega) - \frac{5{,}88S(\omega)}{2\sqrt{n}}\right].$$

Pour l'échantillon simulé présenté à la table 1, on a $n = 100$, $\overline{X} = 2{,}068\,985$, $S^2(\omega) = 0{,}638\,195$, d'où $S(\omega) = 0{,}798\,871$, ce qui nous donne finalement :

$$I_n(\omega) = [2{,}868\,609\ ;\ 3{,}338\,346].$$

3) On pose
$$M_n := \max(X_1, \dots, X_n).$$
Pour calculer la fonction de répartition $F_n(., \theta)$ de M_n, on commence par remarquer que l'inégalité $M_n \leqslant x$ équivaut à « pour tout $i = 1, \dots, n$, $X_i \leqslant x$ », ce qui permet d'écrire l'évènement $\{M_n \leqslant x\}$ comme l'intersection des n évènements P_θ-indépendants et de même P_θ probabilité $\{X_i \leqslant x\}$, $i = 1, \dots, n$. Ainsi, pour tout x réel,
$$F_n(x, \theta) = P_\theta(M_n \leqslant x) = \prod_{i=1}^{n} P_\theta(X_i \leqslant x) = P_\theta(X_1 \leqslant x)^n = F_1(x, \theta)^n.$$
La fonction de répartition $F_1(\,.\, , \theta)$ de X_1 se calcule en intégrant la densité :
$$F_1(x, \theta) = \int_{-\infty}^{x} f(t, \theta)\,\mathrm{d}t = \begin{cases} 0 & \text{si } x < 0, \\ \frac{2}{\theta^2}\int_0^x t\,\mathrm{d}t = \frac{x^2}{\theta^2} & \text{si } 0 \leqslant x \leqslant \theta, \\ 1 & \text{si } x > \theta. \end{cases}$$
On en déduit que
$$F_n(x, \theta) = \begin{cases} 0 & \text{si } x < 0, \\ (\frac{x}{\theta})^{2n} & \text{si } 0 \leqslant x \leqslant \theta, \\ 1 & \text{si } x > \theta. \end{cases}$$

4) Pour montrer que M_n est un estimateur fortement consistant de θ, c'est-à-dire que M_n converge P_θ-p.s. vers θ, il suffit d'établir la convergence de $\sum_{n=1}^{+\infty} P_\theta(|M_n - \theta| > \delta)$ pour tout [1] $\delta > 0$. Comme $0 \leqslant M_n \leqslant \theta$ P_θ-p.s., d'où $P_\theta(|M_n - \theta| > \delta) = 0$ pour $\delta \geqslant \theta$, il suffit de traiter le cas $0 < \delta < \theta$. Le terme général de la série s'exprime à l'aide de la fonction de répartition *continue* $F_n(\,.\, , \theta)$:
$$\begin{aligned} P_\theta(|M_n - \theta| > \delta) &= P_\theta(M_n < \theta - \delta) + P_\theta(M_n > \theta + \delta) \\ &= P_\theta(M_n \leqslant \theta - \delta) + 1 - P_\theta(M_n \leqslant \theta + \delta) \\ &= \left(\frac{\theta - \delta}{\theta}\right)^{2n} + 0. \end{aligned}$$
On voit ainsi que la série est géométrique de raison $q = (\frac{\theta-\delta}{\theta})^{2n} \in]0, 1[$ donc convergente. On en déduit que
$$\forall \theta \in \Theta, \quad M_n \xrightarrow[n \to +\infty]{P_\theta-\text{p.s.}} \theta.$$
M_n est donc un estimateur fortement consistant de θ.

5) Comme M_n est une v.a. P_θ-p.s. positive, on peut calculer $\mathbf{E}_\theta M_n$ en intégrant sa fonction de survie [2] $1 - F_n(\,.\, , \theta)$ sur $[0, +\infty[$:
$$\mathbf{E}_\theta M_n = \int_0^{+\infty} (1 - F_n(x, \theta))\,\mathrm{d}x = \int_0^{\theta} \left(1 - \left(\frac{x}{\theta}\right)^{2n}\right)\mathrm{d}x = \theta - \frac{1}{\theta^{2n}}\int_0^{\theta} x^{2n}\,\mathrm{d}x = \theta - \frac{1}{\theta^{2n}}\frac{\theta^{2n+1}}{2n+1},$$

1. Clairement, si cette série à termes positifs converge pour un $\delta_0 > 0$, elle converge aussi pour tout $\delta > \delta_0$, par majoration terme à terme . On peut donc se contenter d'établir la convergence pour tout δ suffisamment petit.

2. Alternativement, on peut remarquer que sa fonction de répartition $F_n(\,.\, , \theta)$ est C^1 par morceaux pour en déduire que sa loi a une densité $f_n(\,.\, , \theta) = F_n'(\,.\, , \theta)$ et appliquer la formule de calcul de l'espérance d'une v.a. à densité.

d'où
$$\mathbf{E}_\theta M_n = \frac{2n}{2n+1}\theta.$$
Ceci montre que la statistique M_n est un estimateur biaisé de θ, mais comme $\frac{2n}{2n+1}$ tend vers 1 quand n tend vers l'infini, cet estimateur est asymptotiquement sans biais.

L'estimateur
$$M'_n := \frac{2n+1}{2n} M_n$$
vérifie $\mathbf{E}_\theta M'_n = \theta$ pour tout $\theta \in \Theta$, il est donc sans biais. Comme $\frac{2n+1}{2n}$ tend vers 1 quand n tend vers l'infini et M_n converge P_θ-p.s. vers θ pour tout $\theta \in \Theta$ d'après la question précédente, M'_n converge lui aussi P_θ-p.s. vers θ pour tout $\theta \in \Theta$. C'est donc un estimateur fortement consistant de θ.

6) On a P_θ-p.s. l'inégalité $0 \leqslant M_n \leqslant \theta$, d'où $0 \leqslant M'_n \leqslant \frac{2n+1}{2n}\theta$, ce qui montre que M'_n est P_θ-p.s. bornée, donc en particulier de carré P_θ-intégrable. Calculons son erreur quadratique moyenne, égale à sa variance puisque M'_n est sans biais :
$$\mathrm{EQM}_\theta(M'_n) = \mathbf{E}_\theta(M'_n - \theta)^2 = \mathrm{Var}_\theta\, M'_n = \left(\frac{2n+1}{2n}\right)^2 \mathrm{Var}_\theta\, M_n.$$
On est ainsi ramené au calcul de $\mathrm{Var}_\theta\, M_n$:
$$\mathrm{Var}_\theta\, M_n = \mathbf{E}_\theta M_n^2 - (\mathbf{E}_\theta M_n)^2 = \mathbf{E}_\theta M_n^2 - \left(\frac{2n}{2n+1}\theta\right)^2,$$
puis à celui de $\mathbf{E}_\theta M_n^2$. Ce dernier peut se faire au choix en utilisant la densité de M_n ou sa fonction de survie. Le calcul à l'aide de la densité s'écrit :
$$\mathbf{E}_\theta M_n^2 = \int_{-\infty}^{+\infty} t^2 f_n(t,\theta)\,\mathrm{d}t = \frac{2n}{\theta^{2n}} \int_0^\theta t^{2n+1}\,\mathrm{d}t = \frac{2n}{2n+2}\theta^2.$$
Pour le calcul à l'aide de la fonction de survie, on rappelle que $\mathbf{E}g(X) = g(0) + \int_0^{+\infty} g'(s)P(X > s)\,\mathrm{d}s$ pour toute v.a. positive X et toute fonction g strictement croissante sur $\mathbb{R}_+$ et de classe C^1 sur $\mathbb{R}_+^*$. En appliquant ceci avec $g(s) = s^2$ on obtient :
$$\mathbf{E}_\theta M_n^2 = \int_0^{+\infty} 2tP_\theta(M_n > t)\,\mathrm{d}t = \int_0^\theta 2t\left(1 - \frac{t^{2n}}{\theta^{2n}}\right)\mathrm{d}t = \theta^2 - \frac{2\theta^{2n+2}}{(2n+2)\theta^{2n}} = \left(1 - \frac{n}{n+1}\right)\theta^2.$$
On est rassuré de constater que par les deux méthodes on trouve :
$$\mathbf{E}_\theta M_n^2 = \frac{n}{n+1}\theta^2.$$
Finalement on obtient pour l'erreur quadratique moyenne de M'_n :
$$\mathrm{EQM}_\theta(M'_n) = \left(\frac{2n+1}{2n}\right)^2 \frac{n}{n+1}\theta^2 - \theta^2 = \left(\frac{(2n+1)^2}{4n(n+1)} - 1\right)\theta^2 = \frac{\theta^2}{4n(n+1)}.$$
On remarque que
$$\mathrm{EQM}_\theta\, M'_n \sim \frac{\theta^2}{4n^2}, \quad n \to +\infty,$$

ce qui est bien meilleur asymptotiquement que l'inégalité de Cramér-Rao qui prévoit une borne inférieure de la forme $c(\theta)/n$. Il n'y a pas de contradiction car le « support » de la loi des X_i sous P_θ est

$$A(\theta) = \{t \in \mathbb{R};\ f(t,\theta) > 0\} =]0,\theta]$$

qui dépend de θ. La première hypothèse du théorème de Cramér-Rao n'est donc pas satisfaite.

7) Comme les X_i sont à densité $f(\,.\,,\theta)$ sous P_θ, la fonction de vraisemblance associée au n-échantillon s'écrit ici :

$$L(x_1,\dots,x_n,\theta) = \prod_{i=1}^n f(x_i,\theta) = \prod_{i=1}^n \frac{2x_i}{\theta^2}\mathbf{1}_{[0,\theta]}(x_i) = \frac{2^n x_1\dots x_n}{\theta^{2n}}\prod_{i=1}^n \mathbf{1}_{[0,\theta]}(x_i).$$

Remarquons d'abord que cette fonction est nulle (quel que soit θ) si au moins un des x_i est négatif ou nul, à cause des indicatrices si $x_i < 0$, ou du produit des x_j si $x_i = 0$. Dans la recherche d'un maximum pour L, on peut donc se restreindre au cas où tous les x_i sont strictement positifs.

Le produit des indicatrices peut se simplifier comme suit. C'est un produit de n facteurs valant chacun 0 ou 1. Il ne peut donc prendre que les valeurs :

- 0 si au moins l'un des facteurs est nul, c'est à dire si l'un des x_i au moins est strictement supérieur à θ, autrement-dit si $\theta < \max_{1\leqslant i\leqslant n} x_i$;
- 1 si tous les facteurs valent 1, c'est à dire si $0 \leqslant x_i \leqslant \theta$ pour tout i, ce qui s'écrit encore $\max_{1\leqslant i\leqslant n} x_i \leqslant \theta$ (rappelons que l'on a supposé tous les $x_i > 0$).

On voit ainsi en posant $b = \max_{1\leqslant i\leqslant n} x_i$, que la fonction $\theta \mapsto L(x_1,\dots,x_n,\theta)$ est nulle sur $]0,b[$ et que sur $[b,+\infty[$, elle s'écrit :

$$\theta \longmapsto \frac{2^n x_1\dots x_n}{\theta^{2n}}.$$

Sur cet intervalle, c'est une fonction décroissante de θ, elle atteint donc son maximum absolu sur cet intervalle au point $b = \max_{1\leqslant i\leqslant n} x_i$ borne gauche de cet intervalle et ce maximum $L(x_1,\dots,x_n,b) = 2^n x_1\dots x_n b^{-n}$ est strictement positif. Le maximum absolu de la fonction sur $\Theta =]0,+\infty[$ est aussi atteint en b puisque la fonction est nulle sur $]0,b[$.

L'estimateur par maximum de vraisemblance de θ est donc la statistique

$$\max_{1\leqslant i\leqslant n} X_i = M_n.$$

8) Nous allons maintenant utiliser M_n pour construire un intervalle de confiance pour θ. On commence pour cela par remarquer que pour $0 < x < \theta$,

$$P_\theta(M_n \leqslant \theta < M_n + x) = P_\theta(\theta - x < M_n \leqslant \theta) = F_n(\theta,\theta) - F_n(\theta - x,\theta) = 1 - \left(\frac{\theta - x}{\theta}\right)^{2n}.$$

En posant $x = a_n\theta$, $a_n \in]0,1[$, ceci se réécrit :

$$P_\theta(M_n \leqslant \theta < M_n + a_n\theta) = 1 - (1 - a_n)^{2n}.$$

Cette probabilité est égale à $1-\varepsilon$, pour a_n vérifiant

$$(1-a_n)^{2n} = \varepsilon \quad \text{soit} \quad a_n = 1 - \varepsilon^{\frac{1}{2n}}.$$

Remarquons ensuite que

$$\theta < M_n + a_n\theta \iff (1-a_n)\theta < M_n \iff \theta < \frac{M_n}{1-a_n}.$$

Ceci est valable pour tout $a_n \in]0,1[$. Avec le choix de a_n précisé ci-dessus, on obtient finalement :

$$\forall \theta \in \Theta, \quad P_\theta(M_n \leqslant \theta < \varepsilon^{-\frac{1}{2n}} M_n) = 1-\varepsilon.$$

Ceci nous permet de proposer :

$$J_n(\omega) = \left[M_n(\omega), \varepsilon^{-\frac{1}{2n}} M_n(\omega)\right]$$

comme intervalle de confiance au niveau exact $1-\varepsilon$ pour θ.

Pour l'échantillon simulé de la table (1), on a $M_n(\omega) = 3{,}138\,464$ et l'intervalle de confiance $J_n(\omega)$ au niveau 95%, donc avec $\varepsilon = 0{,}05$, est alors

$$J_n(\omega) = \left[3{,}138\,464\,; (0{,}05)^{-1/200} \times 3{,}138\,464\right] = [3{,}138\,464\,; 3{,}185\,828]\,.$$

Cet intervalle est bien meilleur que $I_n(\omega)$, ce qui était prévisible au vu de l'erreur

Table 1 – Un 100-échantillon simulé pour l'exercice 1

2,373 641	2,539 755	1,317 328	2,927 648	3,075 983	3,089 054	2,079 839	1,866 325
3,065 378	1,423 575	2,486 000	0,372 537	2,179 226	3,025 487	3,035 052	2,226 735
2,677 645	1,869 972	2,992 573	2,095 179	1,694 912	2,885 349	1,950 677	1,867 663
1,240 670	2,277 816	3,095 005	1,399 811	3,034 904	3,049 601	2,769 146	3,065 194
1,240 670	2,277 816	3,095 005	1,399 811	3,034 904	3,049 601	2,769 146	3,065 194
2,585 347	1,042 526	3,083 546	3,126 235	0,510 191	1,435 366	1,610 284	2,613 401
0,161 609	2,996 090	2,824 511	2,331 914	2,360 276	0,946 244	3,019 824	0,291 257
1,498 035	2,999 015	2,900 211	1,567 625	2,611 896	2,642 432	2,938 963	2,011 620
1,240 892	2,002 642	1,543 602	2,935 429	1,732 009	0,687 088	2,202 109	1,519 365
2,616 795	3,138 464	3,014 825	2,130 727	2,979 077	0,916 038	1,579 642	1,959 762
1,968 556	2,737 566	1,610 758	0,293 354	2,049 023	2,179 292	0,535 313	1,472 453
2,761 824	1,157 541	2,926 185	1,329 376	2,880 823	1,204 144	2,058 592	0,683 237
1,424 086	1,939 582	1,229 418	1,451 298	2,233 886	2,456 869	2,439 421	2,141 528
2,321 064	1,014 335	1,889 860	1,152 523				

Quelques statistiques associées à cet échantillon $x_1, \dots, x_{100}$:

$$\overline{x} = \frac{1}{100}\sum_{i=1}^{100} x_i = 2{,}068\,985.$$

variance empirique ($s^2 = S^2(\omega)$) $\quad s^2 = 0{,}638\,195, \quad s = 0{,}798\,871.$

$\min(x_1, \dots, x_{100}) = 0{,}161\,609, \quad \max(x_1, \dots, x_{100}) = 3{,}138\,464.$

intervalle médian : $[2{,}130\,727;\ 2{,}141\,528]$.

quadratique moyenne de M_n qui est de l'ordre de n^{-2} comme pour M'_n (vérifiez), alors que celle de $\frac{3}{2}\overline{X}$ n'est que de l'ordre de n^{-1}. De plus le niveau de confiance de $J_n(\omega)$ est exact, alors que celui de $I_n(\omega)$ est entâché de l'erreur d'approximation gaussienne.

Ex 2. *Inversions d'une permutation aléatoire*

On note Ω_n l'ensemble des permutations de $[\![1,n]\!]$, c'est à dire des bijections $\omega : [\![1,n]\!] \to [\![1,n]\!]$. Soient $k = \omega(i) \in [\![1,n]\!]$. On dit que k présente une inversion avec $\ell = \omega(j)$ si $j > i$ et $\ell < k$. Ainsi le nombre d'inversions dans ω de $k = \omega(i)$ est le nombre d'éléments inférieurs à $\omega(i)$ et situés à sa droite dans la liste $\omega(1), \dots, \omega(i), \dots, \omega(n)$. Le nombre d'inversions de k est ainsi toujours compris entre 0 et $k-1$, le nombre d'inversions de 1 est toujours nul.

Le but de cet exercice est d'établir une loi faible des grands nombres et un théorème limite central pour le nombre total d'inversions d'une permutation aléatoire. Notons pour cela $X_k(\omega)$ le nombre d'inversions de k dans ω, pour $k \in [\![1,n]\!]$ et $S_n(\omega) = \sum_{k=1}^n X_k(\omega)$.

1) Lorsque n n'est pas trop grand, on peut représenter ω en donnant la liste explicite des $\omega(i)$. Par commodité, nous écrirons chaque $k \in [\![1,n]\!]$ sur un jeton carré et chaque disposition sur une même ligne des n jetons représentera une permutation ω. Par exemple

$$\boxed{4}\,\boxed{3}\,\boxed{5}\,\boxed{1}\,\boxed{2}$$

représente la permutation ω définie par : $1 \mapsto 4$, $2 \mapsto 3$, $3 \mapsto 5$, $4 \mapsto 1$, $5 \mapsto 2$. L'élément $4 = \omega(1)$ a 3 inversions, puisqu'à sa droite on trouve 3, 1 et 2 qui lui sont inférieurs, donc $X_4(\omega) = 3$. De proche en proche, on voit ainsi que le vecteur d'inversions associé à ω est $(X_1(\omega), \dots, X_5(\omega)) = (0,0,2,3,2)$.

Il y a une seule permutation ayant pour vecteur d'inversions $(0,1,0,3,2)$. Sa construction par placements successifs des 5 jetons est donnée par le tableau :

Jeton n° k	$X_k(\omega)$	Placement
1	0	$\boxed{1}$
2	1	$\boxed{2}\,\boxed{1}$
3	0	$\boxed{2}\,\boxed{1}\,\boxed{3}$
4	3	$\boxed{4}\,\boxed{2}\,\boxed{1}\,\boxed{3}$
5	2	$\boxed{4}\,\boxed{2}\,\boxed{5}\,\boxed{1}\,\boxed{3}$

De même, il y a une seule permutation ayant pour vecteur d'inversions $(0,0,1,2,0)$. Sa construction est donnée par le tableau :

Jeton n° k	$X_k(\omega)$	Placement
1	0	$\boxed{1}$
2	0	$\boxed{1}\,\boxed{2}$
3	1	$\boxed{1}\,\boxed{3}\,\boxed{2}$
4	2	$\boxed{1}\,\boxed{4}\,\boxed{3}\,\boxed{2}$
5	0	$\boxed{1}\,\boxed{4}\,\boxed{3}\,\boxed{2}\,\boxed{5}$

Sur ces deux exemples, on remarque que lors du placement du jeton n° k, le nombre d'inversions $X_k(\omega)$ ne laisse qu'une possibilité de placement du jeton n° k, par rapport aux jetons numérotés de 1 à $k-1$ déjà placés.

2) On note Δ_n le produit cartésien

$$\Delta_n = \{0\} \times [\![0,1]\!] \times [\![0,2]\!] \times \cdots \times [\![0,n-1]\!].$$

Tout vecteur d'inversions d'une permutation $\omega \in \Omega_n$ est dans Δ_n puisque, pour chaque $k = 1, \dots, n$, le nombre d'inversions de k dans ω est le nombre d'entiers situés à sa droite dans le liste $\omega(1), \dots, \omega(n)$ et inférieurs strictement à k. Ce nombre d'inversions de k est donc au plus $k-1$. Autrement dit, pour $k = 1, \dots, n$, $X_k(\omega) \in [\![0, k-1]\!]$, d'où

$$(X_1(\omega), \dots, X_n(\omega)) \in \prod_{k=1}^{n} [\![0, k-1]\!] = \Delta_n$$

et ceci est valide pour tout $\omega \in \Omega_n$.

Nous allons montrer que pour tout $x \in \Delta_n$, il existe *un unique* $\omega \in \Omega_n$ tel que $(X_1(\omega), \dots, X_n(\omega)) = x$. Autrement dit que l'application $\omega \mapsto (X_1(\omega), \dots, X_n(\omega))$ est une bijection de Ω_n sur Δ_n.

Les deux exemples étudiés à la question précédente indiquent la voie à suivre : à partir d'un vecteur x quelconque de Δ_n, on construit pas à pas une permutation ω ayant x pour vecteur d'inversions et on voit que cette construction nous donne à la fois l'existence et l'unicité.

- On commence par poser le jeton $\boxed{1}$.
- Si $x_2 = 0$, $\boxed{2}$ n'a pas d'inversion et on ne peut le placer qu'à la droite de $\boxed{1}$. Sinon, on le place à sa gauche.
- Si $x_3 = 0$, $\boxed{3}$ n'a pas d'inversion et on ne peut le placer qu'à la droite des deux premiers jetons déjà placés ; si $x_3 = 1$, $\boxed{3}$ a une seule inversion et on ne peut le placer qu'entre les deux jetons déjà placés ; si $x_3 = 2$, $\boxed{3}$ a deux inversions et on ne peut le placer qu'à la gauche des deux jetons déjà placés.
- ...
- Les jetons numérotés de 1 à $k-1$ étant déjà placés, la valeur j de x_k $(0 \leqslant j \leqslant k-1)$ nous indique la seule possibilité de placement pour $\boxed{k}$: comme il doit présenter j inversions, on ne peut le mettre qu'en j^{e} position à partir de la droite, c'est-à-dire à la doite des $k-1$ jetons déjà placés si $j = 0$, intercalé entre les deux plus à droite si $j = 1$, etc.
- ...

On peut ainsi construire ω de proche en proche jusqu'à épuisement des n jetons. À chaque étape k, parmi les k positions relatives du jeton $\boxed{k}$ par rapport aux $k-1$ jetons déjà placés, une seule est compatible avec la valeur de x_k. Ceci nous donne bien l'existence et l'unicité de ω ayant pour vecteur d'inversions x.

3) On munit Ω_n de la tribu $\mathcal{F}_n = \mathcal{P}(\Omega_n)$ de tous ses sous-ensembles et de la probabilité P_n, loi uniforme sur Ω_n. On a donc $P_n(\{\omega\}) = \frac{1}{n!}$ pour tout $\omega \in \Omega_n$. On en déduit immédiatement que :

$$\forall x \in \Delta_n, \quad P_n((X_1, \dots, X_n) = x) = \frac{1}{n!}$$

puisque pour chaque $x \in \Delta_n$, il existe un unique $\omega \in \Omega_n$ tel que $(X_1(\omega), \dots, X_n(\omega)) = x$, donc pour cet ω, $\{(X_1, \dots, X_n) = x\} = \{\omega\}$. Or $P_n(\{\omega\}) = \frac{1}{n!}$ pour tout $\omega \in \Omega_n$.

Ainsi le vecteur aléatoire discret $(X_1, \dots, X_n)$ peut prendre toute valeur vectorielle x dans Δ_n et nulle part ailleurs (d'après la question précédente). Il prend chacune de ces valeurs x avec *la même* probabilité $\frac{1}{n!}$. Sa loi est donc la *loi uniforme* (discrète) sur l'ensemble fini Δ_n.

Une autre façon de voir que $(X_1, \dots, X_n)$ suit la loi uniforme sur Δ_n est de remarquer que :

$$\operatorname{card} \Delta_n = \prod_{k=1}^{n} \operatorname{card}[\![0, k-1]\!] = n!$$

d'où

$$\forall x \in \Delta_n, \quad P_n((X_1, \dots, X_n) = x) = \frac{1}{\operatorname{card} \Delta_n}.$$

4) Soit un espace probabilisé $(\Omega', \mathcal{F}', P')$ sur lequel on puisse définir des variables aléatoires discrètes $U_1, \dots, U_n$ indépendantes et telles que pour tout $k \in [\![1, n]\!]$, U_k suive la loi uniforme sur $[\![0, k-1]\!]$. L'énoncé ne vous demandait pas de vérifier l'existence d'un tel espace[3]. Pour identifier la loi du vecteur aléatoire discret $U = (U_1, \dots, U_n)$, on calcule pour tout $x = (x_1, \dots, x_n) \in \Delta_n$, $P'((U_1, \dots, U_n) = (x_1, \dots, x_n))$:

$$\begin{aligned}
P'\left((U_1, \dots, U_n) = (x_1, \dots, x_n)\right) &= P'\left(U_1 = x_1, \dots, U_n = x_n\right) \\
&= \prod_{k=1}^{n} P'(U_k = x_k) \quad \text{(par indépendance des } U_k\text{)} \\
&= \prod_{k=1}^{n} \frac{1}{k} \quad \text{(car } U_k \text{ suit la loi uniforme sur } [\![0, k-1]\!]\text{)} \\
&= \frac{1}{n!} = \frac{1}{\operatorname{card} \Delta_n}.
\end{aligned}$$

Ceci étant vrai pour tout $x = (x_1, \dots, x_n) \in \Delta_n$, on en conclut que le vecteur aléatoire U suit la loi uniforme sur Δ_n.

Ainsi les vecteurs aléatoires $(X_1, \dots, X_n)$ et $(U_1, \dots, U_n)$ ont même loi. Ceci implique que chaque X_k a même loi que U_k, donc suit la loi uniforme sur $[\![0, k-1]\!]$. De plus comme les composantes de $(U_1, \dots, U_n)$ sont indépendantes, il en va de même pour celles de $(X_1, \dots, X_n)$. Cela devrait vous paraître évident si vous êtes familier avec les lois produits : un vecteur aléatoire est à composantes indépendantes si et seulement si sa loi est le produit de ses lois marginales, donc le fait d'être à composantes indépendantes est une propriété qui ne dépend que de la loi du vecteur. Sinon, il suffit d'écrire :

$$\begin{aligned}
P((X_1, \dots, X_n) = (x_1, \dots, x_n)) &= P'((U_1, \dots, U_n) = (x_1, \dots, x_n)) \\
&= \prod_{k=1}^{n} P'(U_k = x_k) \quad \text{(par indépendance des } U_k\text{)} \\
&= \prod_{k=1}^{n} P(X_k = x_k) \quad \text{(car } U_k \text{ et } X_k \text{ ont même loi).}
\end{aligned}$$

3. Vous pourrez vérifier à titre d'exercice qu'une solution possible est de prendre $\Omega' = \Delta_n$, $\mathcal{F}'$ l'ensemble des parties de Δ_n, P' la probabilité uniforme sur Δ_n (donnée par $P'(\{x\}) = \frac{1}{n!}$) avec pour $U = (U_1, \dots, U_n)$ l'identité sur Δ_n.

5) Justification des formules sommatoires

$$T_{n,1} := \sum_{j=1}^{n} j = \frac{n(n+1)}{2} \tag{1}$$

et

$$T_{n,2} := \sum_{j=1}^{n} j^2 = \frac{n(n+1)(2n+1)}{6}. \tag{2}$$

Ces deux formules peuvent se vérifier par récurrence. Il y a une méthode plus simple pour la première, illustrée par le tableau suivant :

1	2	3	$\dots$	k	$\dots$	$n-2$	$n-1$	n	$T_{n,1}$
n	$n-1$	$n-2$	$\dots$	$n-k+1$	$\dots$	3	2	1	$T_{n,1}$
$n+1$	$n+1$	$n+1$	$\dots$	$n+1$	$\dots$	$n+1$	$n+1$	$n+1$	$2T_{n,1}$

qui nous montre en sommant horizontalement et verticalement que $2T_{n,1} = n(n+1)$, ce qui donne (1).

Pour tout entier r, notons $T_{n,r} = \sum_{j=1}^{n} j^r$. Pour calculer $T_{n,2}$, on cherche une relation de récurrence entre $T_{n+1,3}$ et $T_{n,3}$ faisant intervenir $T_{n,2}$ et $T_{n,1}$. Pour cela, on note que

$$T_{n+1,3} = \sum_{j=1}^{n+1} j^3 = 1 + \sum_{k=1}^{n} (k+1)^3 = 1 + \sum_{k=1}^{n} (k^3+3k^2+3k+1) = T_{n,3} + 3T_{n,2} + 3T_{n,1} + (n+1).$$

En notant que l'on a directement $T_{n+1,3} = T_{n,3} + (n+1)^3$, on en déduit :

$$(n+1)^3 - (n+1) = 3T_{n,2} + 3T_{n,1}$$

d'où

$$3T_{n,2} = (n+1)((n+1)^2 - 1) - \frac{3}{2}n(n+1) = (n+1)(n^2 + 2n - \frac{3}{2}n) = n(n+1)(n + \frac{1}{2}),$$

ce qui nous donne bien (2).

On aurait pu utliser la même méthode pour calculer $T_{n,1}$ en comparant $T_{n+1,2}$ et $T_{n,2}$.

6) Ces formules sommatoires vont nous servir à calculer l'espérance et la variance de $S_n = X_1 + \cdots + X_n$.

Puisque X_k suit la loi uniforme sur $[\![0, k-1]\!]$, son espérance est :

$$\mathbf{E}X_k = \sum_{j=0}^{k-1} jP(X_k = j) = \frac{1}{k}\sum_{j=1}^{k-1} j = \frac{1}{k}T_{k-1,1} = \frac{1}{k}\frac{(k-1)k}{2} = \frac{k-1}{2}.$$

Par additivité de l'espérance, on en déduit :

$$\mathbf{E}S_n = \sum_{k=1}^{n} \mathbf{E}X_k = \sum_{k=1}^{n} \frac{k-1}{2} = \frac{1}{2}T_{n-1,1} = \frac{n(n-1)}{4}.$$

Par indépendance des X_k, $\operatorname{Var} S_n = \operatorname{Var} X_1 + \cdots + \operatorname{Var} X_n$. Pour calculer la variance de X_k, on commence par calculer $\mathbf{E}X_k^2$:

$$\mathbf{E}X_k^2 = \sum_{j=0}^{k-1} j^2 P(X_k = j) = \frac{1}{k}\sum_{j=1}^{k-1} j^2 = \frac{1}{k}T_{k-1,2} = \frac{(k-1)(2k-1)}{6}.$$

Ensuite,

$$\begin{aligned}\operatorname{Var} X_k = \mathbf{E}X_k^2 - (\mathbf{E}X_k)^2 &= \frac{(k-1)(2k-1)}{6} - \frac{(k-1)^2}{4}\\ &= \frac{k-1}{2}\Big(\frac{2k-1}{3} - \frac{k-1}{2}\Big)\\ &= \frac{(k-1)(k+1)}{12} = \frac{k^2-1}{12}.\end{aligned}$$

Finalement,

$$\begin{aligned}\operatorname{Var} S_n = \sum_{k=1}^{n}\frac{k^2-1}{12} = \frac{1}{12}(T_{n,2}-n) &= \frac{1}{12}\left(\frac{n(n+1)(2n+1)}{6} - n\right)\\ &= \frac{n}{72}\big((n+1)(2n+1) - 6\big)\\ &= \frac{n}{72}(2n^2+3n-5)\\ &= \frac{n(n-1)(2n+5)}{72}.\end{aligned}$$

7) Puisque $\mathbf{E}S_n$ est équivalent à $n^2/4$ quand n tend vers l'infini, il est naturel de se demander si S_n normalisé par n^2 satisfait une loi faible des grands nombres avec limite $1/4$. Plus précisément on regarde si la convergence « en probabilité[4] » suivante est vérifiée :

$$\forall\varepsilon > 0, \quad P_n\left(\left|\frac{S_n}{n^2} - \frac{1}{4}\right| > \varepsilon\right) \xrightarrow[n\to+\infty]{} 0. \tag{3}$$

L'idée est de se ramener à une inégalité de Tchebycheff en procédant comme suit :

$$\begin{aligned}P_n\left(\left|\frac{S_n}{n^2} - \frac{1}{4}\right| > \varepsilon\right) &= P_n\left(\left|S_n - \frac{n^2}{4}\right| > \varepsilon n^2\right)\\ &\leqslant P_n\left(|S_n - \mathbf{E}S_n| + \left|\mathbf{E}S_n - \frac{n^2}{4}\right| > \varepsilon n^2\right)\\ &\leqslant P_n\left(|S_n - \mathbf{E}S_n| > \frac{\varepsilon n^2}{2}\right) + P_n\left(\left|\mathbf{E}S_n - \frac{n^2}{4}\right| > \frac{\varepsilon n^2}{2}\right)\end{aligned}$$

Dans le deuxième terme de ce majorant, la « variable aléatoire » $|\mathbf{E}S_n - n^2/4|$ est constante et vaut $n/4$. La probabilité corespondante vaut donc 1 si $n/4 > \varepsilon n^2/2$ et 0

4. Il s'agit ici d'une notion généralisée de convergence en probabilité puisque chaque S_n est définie sur un espace probabilisé dépendant de n et muni d'une probabilité P_n.

sinon. Il est donc clair qu'elle vaudra toujours zéro à partir d'un certain rang n_1. On peut prendre $n_1 \geqslant (2\varepsilon)^{-1}$.

Pour majorer le premier terme, on utilise l'inégalité de Tchebycheff :

$$P_n\left(|S_n - \mathbf{E}S_n| > \frac{\varepsilon n^2}{2}\right) \leqslant \frac{4 \operatorname{Var} S_n}{\varepsilon^2 n^4} = \frac{n(n-1)(2n+5)}{18\varepsilon^2 n^4} \sim \frac{1}{9\varepsilon^2 n},$$

ce qui montre que $P_n\left(|S_n - \mathbf{E}S_n| > \frac{\varepsilon n^2}{2}\right)$ tend vers 0 quand n tend vers l'infini et complète la preuve de (3).

8) On rappelle le théorème limite central de Lindeberg que l'on énonce ici sous une forme adaptée à notre problème.

Théorème 1 (Lindeberg)**.** *Considérons le « tableau triangulaire » de variables aléatoires de ligne numéro n ($n \in \mathbb{N}^*$) :*

$$X_{n,1}, \dots, X_{n,i}, \dots, X_{n,n},$$

où ces n variables sont définies sur le même $(\Omega_n, \mathcal{F}_n, P_n)$, indépendantes, de carré intégrable. On note

$$S_n := \sum_{i=1}^{n} X_{n,i}, \qquad \sigma_{n,i}^2 := \operatorname{Var} X_{n,i}, \qquad s_n^2 := \operatorname{Var} S_n = \sum_{i=1}^{n} \sigma_{n,i}^2.$$

On suppose $s_n > 0$ pour tout n. Si de plus le tableau vérifie la condition de Lindeberg *:*

$$\forall \varepsilon > 0, \quad R_{n,\varepsilon} := \frac{1}{s_n^2} \sum_{i=1}^{n} \mathbf{E}\Big((X_{n,i} - \mathbf{E}X_{n,i})^2 \mathbf{1}_{\{|X_{n,i} - \mathbf{E}X_{n,i}| > \varepsilon s_n\}}\Big) \xrightarrow[n \to +\infty]{} 0, \tag{4}$$

alors $s_n^{-1}(S_n - \mathbf{E}S_n)$ converge en loi vers une v.a. Z de loi $\mathfrak{N}(0,1)$.

Dans notre problème, les espaces probabilisés $(\Omega_n, \mathcal{F}_n, P_n)_{n \geqslant 1}$ sont ceux définis à la question 3. La v.a. notée X_k, quand on travaillait à n fixé, est notée maintenant $X_{n,k}$ (elle dépend de n puisqu'elle s'applique à des permutations ω sur $[\![1, n]\!]$).

L'écart type s_n est ici équivalent à $n^{3/2}/6$ quand n tend vers l'infini d'après le calcul de $\operatorname{Var} S_n$ effectué ci-dessus. Or $X_{n,i}$ est à valeurs dans $[\![0, i-1]\!]$, ce qui implique que

$$|X_{n,i} - \mathbf{E}X_{n,i}| \leqslant i - 1 \leqslant n$$

Comme s_n tend vers l'infini plus vite que ce majorant n, il en va de même pour εs_n. Par conséquent il existe un rang $n_0 = n_0(\varepsilon)$ à partir duquel l'inégalité $|X_{n,i} - \mathbf{E}X_{n,i}| > \varepsilon s_n$ n'est vraie pour aucun $n \geqslant n_0$, aucun $i \leqslant n$ et aucun $\omega \in \Omega_n$. Il en résulte que pour $n \geqslant n_0(\varepsilon)$, toutes les indicatrices dans $R_{n,\varepsilon}$ sont nulles (sur Ω_n tout entier) et donc $R_{n,\varepsilon} = 0$ pour tout $n \geqslant n_0(\varepsilon)$. Comme ceci est valide pour tout $\varepsilon > 0$, la condition de Lindeberg (4) est trivialement vérifiée. Par le théorème de Lindeberg on a donc la convergence en loi :

$$S_n^\star := \frac{1}{s_n}(S_n - \mathbf{E}S_n) \xrightarrow[n \to +\infty]{\text{loi}} Z,$$

où Z est gaussienne $\mathfrak{N}(0,1)$.

Pour en déduire la convergence en loi vers Z de $6n^{-3/2}(S_n - n^2/4)$, on rappelle que $\mathbf{E}S_n - n^2/4 = -n/4$, d'où

$$\frac{6}{n^{3/2}}\left(S_n - \frac{n^2}{4}\right) = \frac{6}{n^{3/2}}\left(S_n - \mathbf{E}S_n - \frac{n}{4}\right) = \frac{6}{n^{3/2}}(S_n - \mathbf{E}S_n) - \frac{3}{2n^{1/2}} = \frac{6s_n}{n^{3/2}}S_n^\star - \frac{3}{2n^{1/2}}.$$

On obtient ainsi une expression de la forme :

$$\frac{6}{n^{3/2}}\left(S_n - \frac{n^2}{4}\right) = a_n S_n^\star + b_n, \quad \lim_{n\to+\infty} a_n = 1, \quad \lim_{n\to+\infty} b_n = 0.$$

En considérant a_n et b_n comme des variables aléatoires constantes, les limites ci-dessus nous donnent trivialement la convergence en probabilité de a_n vers 1 et la convergence en probabilité de b_n vers 0. En appliquant deux fois le lemme de Slutsky, on en déduit successivement que $a_n S_n^\star$ converge en loi vers Z, puis que $a_n S_n^\star + b_n$ converge en loi vers Z.

9) Nous venons de prouver la convergence en loi

$$\frac{6}{n^{3/2}}\left(S_n - \frac{n^2}{4}\right) \xrightarrow[n\to+\infty]{\text{loi}} Z,$$

qui implique, puisque la fonction de répartition Φ de la gaussienne standard Z est continue sur $\mathbb{R}$ que

$$\forall t > 0, \quad P_n\left(\frac{6}{n^{3/2}}\left|S_n - \frac{n^2}{4}\right| \leqslant t\right) \xrightarrow[n\to+\infty]{\text{loi}} P(|Z| \leqslant t) = \Phi(t) - \Phi(-t) = 2\Phi(t) - 1.$$

Ceci justifie pour n grand, l'approximation :

$$P_n\left(\frac{n^2}{4} - \frac{tn^{3/2}}{6} \leqslant S_n \leqslant \frac{n^2}{4} + \frac{tn^{3/2}}{6}\right) \simeq 2\Phi(t) - 1. \tag{5}$$

Cherchons t tel que $2\Phi(t) - 1 = 0{,}5$, ce qui équivaut à $\Phi(t) = 0{,}75$. La table donne

$$\Phi(0{,}67) = 0{,}748\,6, \quad \Phi(0{,}68) = 0{,}751\,7.$$

On en déduit une valeur approchée de t par interpolation affine :

$$t \simeq 0{,}67 + \frac{0{,}75 - 0{,}748\,6}{0{,}751\,7 - 0{,}748\,6}(0{,}68 - 0{,}67) = 0{,}67 + \frac{14}{31} \times 0{,}01 \simeq 0{,}674\,5.$$

En reportant cette valeur dans (5), on en déduit que pour n grand,

$$P_n\left(\frac{n^2}{4} - 0{,}113\,n^{3/2} \leqslant S_n \leqslant \frac{n^2}{4} + 0{,}113\,n^{3/2}\right) \simeq \frac{1}{2}$$

Et comme P_n est la probabilité *uniforme* sur l'espace Ω_n de toutes les permutations de $[\![1,n]\!]$, cela signifie bien que « pour n grand, environ la moitié des permutations sur $[\![1,n]\!]$ ont leur nombre total d'inversions compris entre les bornes $\dfrac{n^2}{4} \pm 0{,}113\,n^{3/2}$ ».

La source de cet exercice se trouve dans le livre de W. Feller *An introduction to probability theory and its applications*, tome 1, chap. X, section 6, où elle est traitée en une demi page.