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

IS Math314 **Année 2008**

Corrigé de l'examen du 26 mai 2008.

Ex 1. *Échauffement gaussien*

Soient X et Y, deux variables aléatoires définies sur le même espace probabilisé, indépendantes et de même loi gaussienne standard $\mathfrak{N}(0,1)$. On pose $Z := X + Y$.

1) La covariance de deux variables aléatoires existe si ces deux variables sont de carré intégrable. Les variables aléatoires gaussiennes X et Y étant de carré intégrable, il en est de même pour leur somme Z, ce qui justifie l'existence de $\operatorname{Cov}(Z,Y)$. Cette covariance se calcule comme suit :

$$\operatorname{Cov}(Z,Y) = \operatorname{Cov}(X+Y,Y) = \operatorname{Cov}(X,Y) + \operatorname{Cov}(Y,Y) = 0 + \operatorname{Var}(Y) = 1.$$

Ce calcul utilise successivement la bilinéarité de la covariance, l'indépendance de X et Y et le fait qu'une gaussienne standard a pour variance 1.

2) Si deux variables aléatoires de carré intégrable sont indépendantes, leur covariance est nulle. Comme $\operatorname{Cov}(Z,Y) = 1$, Z et Y ne sont pas indépendantes.

3) Le vecteur (X,Y) est gaussien car à composantes indépendantes gaussiennes. Son vecteur espérance est $(0,0)$ et sa matrice de covariance est la matrice identité. On remarque que (Z,Y) est l'image de (X,Y) par l'application *linéaire* $\varphi : (x,y) \mapsto (x+y,y)$. Il est donc encore gaussien puisque la famille des lois de vecteurs aléatoires gaussiens est conservée par image linéaire. Son espérance est $\varphi(0,0) = (0,0)$ et sa matrice de covariance est

$$K = \begin{pmatrix} \operatorname{Var} Z & \operatorname{Cov}(Z,Y) \\ \operatorname{Cov}(Z,Y) & \operatorname{Var} Y \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix},$$

en notant que par indépendance de X et Y, $\operatorname{Var}(X+Y) = \operatorname{Var} X + \operatorname{Var} Y = 2$.

Ex 2. *Intervalle de confiance*

Dans une verrerie industrielle, une chaîne de production fournit des bouteilles vides. On s'intéresse à la masse moyenne m (en grammes) d'une bouteille produite par cette chaîne. Celle-ci peut s'interpréter comme l'espérance d'une variable aléatoire X de loi inconnue. Pour estimer m, on prélève un échantillon de 400 bouteilles que l'on pèse une par une. On obtient ainsi les données numériques x_i $(i = 1\dots,400)$ où x_i est la masse de la i^{e} bouteille pesée. On donne les résultats intermédiaires :

$$\sum_{i=1}^{400} x_i = 79\ 882 \text{ g}, \qquad \sum_{i=1}^{400} x_i^2 = 15\ 963\ 824 \text{ g}^2.$$

Nous allons proposer un intervalle de confiance au niveau 95% pour m. On interprète pour cela les valeurs x_i comme des réalisations $x_i = X_i(\omega)$ d'une suite de variables aléatoires X_i, indépendantes et de même loi que X, d'espérance m. Ces variables aléatoires sont bornées donc de carré intégrable et les hypothèses du théorème limite central avec autonormalisation sont satisfaites[1]. En notant $\overline{X}$ la moyenne empirique du n échantillon, S la racine carrée positive de sa variance empirique, on a ainsi :

$$T_n := \frac{\sqrt{n}}{S}(\overline{X} - m) \xrightarrow[n\to+\infty]{\text{loi}} Z,$$

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

Ce théorème légitime pour n « grand » (ici $n = 400$) l'approximation gaussienne :

$$P(|T_n| \leq t_\varepsilon) \simeq P(|Z| \leq t_\varepsilon) = 2\Phi(t_\varepsilon) - 1.$$

Pour un niveau de confiance $1-\varepsilon$, on résout l'équation $2\Phi(t_\varepsilon) - 1 = 1 - \varepsilon$, ce qui donne $t_\varepsilon = \Phi^{-1}(1-\varepsilon/2)$ et en particulier pour $1 - \varepsilon = 0{,}95$, $t_\varepsilon = \Phi^{-1}(0{,}975) = 1{,}96$. Ceci nous permet de proposer pour m l'intervalle de confiance I suivant au niveau $1-\varepsilon$:

$$I = \left[\overline{X} - \frac{\Phi^{-1}(1-\varepsilon/2)S}{\sqrt{n}}, \overline{X} + \frac{\Phi^{-1}(1-\varepsilon/2)S}{\sqrt{n}}\right].$$

Pour passer à l'application numérique avec $t_\varepsilon = 1{,}96$, il nous faut calculer $\overline{X}$ et S en utilisant les calculs intermédiaires fournis par l'énoncé et en rappelant que $n = 400$.

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{79\ 882}{400} = 199{,}705 \text{ g}.$$

$$S^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - (\overline{X})^2 = \frac{15\ 963\ 824}{400} - 199{,}705^2 = 39\ 909{,}56 - 39\ 882{,}087\ 025 = 27{,}472\ 975.$$

$$S = \sqrt{27{,}472\ 975} \simeq 5{,}241\ 47 \text{ g}.$$

Finalement, l'intervalle de confiance recherché est

$$I = \left[199{,}191 \text{ g}, 200{,}219 \text{ g}\right].$$

Arrière-cuisine : en fait les données numériques proposées par l'énoncé étaient issues de la simulation d'un 400-échantillon de la loi gaussienne $\mathfrak{N}(200, 5)$.

1. On peut clairement supposer ici que $\sigma^2 = \operatorname{Var} X$ est *strictement* positif. En effet s'il était nul, X serait presque sûrement constante, de même que les X_i. En pratique, cela se traduirait par l'égalité de *toutes* les observations x_i. Or si l'on cherche un intervalle de confiance pour m, c'est bien parce que les données expérimentales x_i présentent une certaine variabilité. Sinon, il suffirait de prendre pour m la valeur commune observée de tous les x_i.

Ex 3. *Un exemple de T.L.C. sans indépendance*

Le but de cet exercice est de présenter un exemple de somme de n variables aléatoires non indépendantes ni de même loi, qui avec normalisation $n^{-1/2}$, converge en loi vers une gaussienne.

1) On dit que la variable aléatoire Y est symétrique (ou de loi symétrique) si Y et $-Y$ ont même loi. Si Y est symétrique et intégrable ($\mathbf{E}|Y| < +\infty$), alors $\mathbf{E}Y = 0$. En effet, Y et $-Y$ ayant même loi et étant intégrables, ont même espérance : $\mathbf{E}Y = \mathbf{E}(-Y)$. Or par linéarité de l'espérance $\mathbf{E}(-Y) = -\mathbf{E}Y$. Par conséquent $\mathbf{E}Y = -\mathbf{E}Y$, ce qui impose $\mathbf{E}Y = 0$.

Si X est une v.a. symétrique quelconque (pas forcément intégrable), X et $-X$ ont même loi, donc $h(X)$ et $h(-X)$ ont même loi pour toute fonction borélienne h. Si h est de plus impaire, $h(X)$ et $h(-X) = -h(X)$ ont même loi, donc la v.a. $h(X)$ a une loi symétrique. Ceci s'applique en particulier en prenant pour h la fonction sinus. Comme cette fonction est en outre bornée, $Y = \sin X$ est intégrable comme variable aléatoire bornée et d'après le paragraphe précédent, $\mathbf{E}(\sin X) = 0$.

Dans toute la suite, $(a_k)_{k\geq 1}$ désigne une suite de nombres réels vérifiant :

$$\frac{1}{n}\sum_{k=1}^{n} a_k^2 \xrightarrow[n\to+\infty]{} 0. \tag{1}$$

Notons que cette condition est réalisée en particulier si $(a_n)_{n\geq 1}$ converge vers 0 par application du théorème de Césaro [2] à la suite $(a_n^2)_{n\geq 1}$ qui converge vers 0. On note $(X_k)_{k\geq 0}$, une suite de variables aléatoires réelles, définies sur le même espace de probabilité, indépendantes et de même loi *symétrique*. On pose pour tout $k \geq 1$,

$$Y_k := a_k \sin(X_{k-1}), \qquad Z_k = X_k + Y_k.$$

2) La suite $(Y_k)_{k\geq 1}$ hérite de l'indépendance de la suite $(X_k)_{k\geq 0}$, mais pas de son équidistribution car la loi de Y_k dépend de k *via* le coefficient a_k, alors que celle de X_k est la même que celle de X_1. Comme les Y_k sont bornées donc de carré intégrable, cette indépendance implique la nullité de $\operatorname{Cov}(Y_j, Y_k)$ pour tout couples d'indices distincts j et k. On a donc

$$\operatorname{Var}\left(\frac{1}{\sqrt{n}}\sum_{k=1}^{n} Y_k\right) = \frac{1}{n}\sum_{j,k=1}^{n}\operatorname{Cov}(Y_j, Y_k) = \frac{1}{n}\sum_{k=1}^{n}\operatorname{Var} Y_k = \frac{1}{n}\sum_{k=1}^{n} a_k^2\mathbf{E}\sin^2(X_{k-1}),$$

en rappelant que d'après la question 1), $\mathbf{E}\sin(X_{k-1})$ est nulle. En majorant $\mathbf{E}\sin^2(X_{k-1})$ par 1, on obtient ainsi

$$0 \leq \operatorname{Var}\left(\frac{1}{\sqrt{n}}\sum_{k=1}^{n} Y_k\right) \leq \frac{1}{n}\sum_{k=1}^{n} a_k^2.$$

2. Selon ce théorème, si $(u_n)_{n\geq 1}$ converge vers une limite finie ℓ, la suite des moyennes arithmétiques $n^{-1}\sum_{k=1}^{n} u_k$ converge elle aussi vers ℓ. Il est impossible que vous ayiez échappé à cet exercice classique dans votre parcours en analyse.

Compte-tenu de la convergence (1), on en déduit :

$$\operatorname{Var}\left(\frac{1}{\sqrt{n}}\sum_{k=1}^{n} Y_k\right) \xrightarrow[n\to+\infty]{} 0. \tag{2}$$

La variable aléatoire $R_n := n^{-1/2}\sum_{k=1}^{n} Y_k$ est centrée donc $\operatorname{Var}(R_n) = \mathbf{E}R_n^2 = \mathbf{E}|R_n - 0|^2$, ce qui nous autorise à interpréter (2) comme la convergence en moyenne quadratique (ou au sens L^2) de R_n vers 0 quand n tend vers l'infini. Cette convergence implique la convergence en probabilité vers la même limite :

$$R_n = \frac{1}{\sqrt{n}}\sum_{k=1}^{n} Y_k \xrightarrow[n\to+\infty]{\Pr} 0. \tag{3}$$

3) On suppose désormais que X_1 est de carré intégrable et d'écart-type σ strictement positif. On peut alors appliquer le théorème limite central (version i.i.d.) à la suite $(X_k)_{k\geq 1}$, ce qui s'écrit ici en rappelant que $\mathbf{E}X_k = 0$ par symétrie et intégrabilité de X_k :

$$\frac{1}{\sigma\sqrt{n}}\sum_{k=1}^{n} X_k = \frac{1}{\sigma}T_n \xrightarrow[n\to+\infty]{\text{loi}} Z'$$

où Z' est gaussienne $\mathfrak{N}(0,1)$, ou de manière équivalente :

$$T_n \xrightarrow[n\to+\infty]{\text{loi}} Z = \sigma Z', \tag{4}$$

la loi de Z étant cette fois $\mathfrak{N}(0,\sigma)$.

4) En se rappelant que $Z_k = X_k + Y_k$, on a

$$\frac{1}{\sqrt{n}}\sum_{k=1}^{n} Z_k = \frac{1}{\sqrt{n}}\sum_{k=1}^{n} X_k + R_n.$$

Les convergences (3) et (4) nous permettent alors d'appliquer le lemme de Slutsky pour conclure que

$$\frac{1}{\sqrt{n}}\sum_{k=1}^{n} Z_k \xrightarrow[n\to+\infty]{\text{loi}} Z, \tag{5}$$

où Z est une v.a. gaussienne $\mathfrak{N}(0,\sigma)$.

Nous avons ainsi construit un exemple de suite $(Z_k)_{k\geq 1}$ de v.a. non indépendantes, non équidistribuées, qui vérifie un théorème limite central.

Ex 4. *Estimation de paramètres*

Soient a et b des constantes strictement positives et X une variable aléatoire positive définie sur un espace probabilisé $(\Omega, \mathcal{F}, P)$ dont la fonction de survie est donnée par

$$P(X > t) = G_{a,b}(t) = \begin{cases} 1 & \text{si } t < a, \\ \exp(-bt) & \text{si } t \geq a. \end{cases} \tag{6}$$

1) Que vaut $P(X = a)$? La fonction de survie $G_{a,b}$ est *discontinue* au point $t = a$ car sa limite à gauche 1 est différente de $G_{a,b}(a) = \exp(-ab)$ puisque ab n'est pas nul. Il en va de même pour la fontion de répartition $F_{a,b} = 1 - G_{a,b}$. Le saut de $F_{a,b}$ au point $t = a$ est la valeur de $P(X = a)$:

$$P(X = a) = F_{a,b}(a) - F_{a,b}(a-) = (1 - \mathrm{e}^{-ab}) - 0 = 1 - \mathrm{e}^{-ab} > 0.$$

L'existence de ce saut interdit à la loi de X d'être à densité. Remarquons au passage que la loi de X n'est pas discrète car la somme des sauts (il n'y en a qu'un c'est $1-\exp(-ab)$) est strictement inférieure à 1.

2) Calculons en fonction de a et b les quantités $\mathbf{E}(X - a)$ et $\mathbf{E}(X - a)^2$. Comme X n'est ni discrète ni à densité, le seul moyen de calculer l'espérance des v.a. $(X - a)$ et $(X - a)^2$ est de revenir à la définition de l'espérance. On remarque d'abord que

$$P(X - a \geq 0) = P(X \geq a) = P(X > a) + P(X = a) = \mathrm{e}^{-ab} + (1 - \mathrm{e}^{-ab}) = 1,$$

donc $X - a$ est presque-sûrement positive et on peut la traiter comme une v.a. positive pour le calcul de son espérance qui s'écrit ainsi :

$$\begin{aligned}
\mathbf{E}(X - a) = \int_0^{+\infty} P(X - a > t)\,\mathrm{d}t &= \int_0^{+\infty} P(X > a + t)\,\mathrm{d}t \\
&= \int_0^{+\infty} \exp(-b(a + t))\,\mathrm{d}t \\
&= \mathrm{e}^{-ab} \int_0^{+\infty} \exp(-bt)\,\mathrm{d}t \\
&= \mathrm{e}^{-ab} \left[\frac{\exp(-bt)}{-b} \right]_0^{+\infty},
\end{aligned}$$

d'où finalement

$$\mathbf{E}(X - a) = \frac{1}{b \exp(ab)}. \tag{7}$$

On utilise la même méthode pour calculer $\mathbf{E}(X - a)^2$ en commençant par remarquer que pour tout $t \geq 0$, $P((X - a)^2 > t) = P(X - a > \sqrt{t})$ puisque $P(X - a < 0) = 0$ implique $P(X - a < -\sqrt{t}) = 0$. Par conséquent,

$$\begin{aligned}
\mathbf{E}(X - a)^2 = \int_0^{+\infty} P\big((X - a)^2 > t\big)\,\mathrm{d}t &= \int_0^{+\infty} P(X - a > \sqrt{t})\,\mathrm{d}t \\
&= \int_0^{+\infty} P(X > a + \sqrt{t})\,\mathrm{d}t \\
&= \int_0^{+\infty} \exp(-b(a + \sqrt{t}))\,\mathrm{d}t \\
&= \mathrm{e}^{-ab} \int_0^{+\infty} \exp(-b\sqrt{t})\,\mathrm{d}t \\
&= 2\mathrm{e}^{-ab} \int_0^{+\infty} \exp(-bs)s\,\mathrm{d}s.
\end{aligned}$$

Cette dernière intégrale se calcule par parties :

$$\int_0^{+\infty} \exp(-bs)s\,\mathrm{d}s = \left[\frac{s\exp(-bs)}{-b}\right]_0^{+\infty} + \frac{1}{b}\int_0^{+\infty} \exp(-bs)\,\mathrm{d}s$$
$$= 0 + \frac{1}{b}\left[\frac{\exp(-bt)}{-b}\right]_0^{+\infty} = \frac{1}{b^2}.$$

Finalement,

$$\mathbf{E}(X-a)^2 = \frac{2}{b^2\exp(ab)}. \tag{8}$$

Notons au passage que ce calcul montre que la v.a. positive $(X-a)^2$ est intégrable. Comme $X^2 \leq 2(X-a)^2 + 2a^2$, on en déduit que X est de carré intégrable.

De (7) et (8) on tire :

$$\frac{2\mathbf{E}(X-a)}{\mathbf{E}(X-a)^2} = b. \tag{9}$$

3) En fait les paramètres a et b sont tous deux inconnus et on cherche à les estimer. Pour cela on note $\theta = (a,b)$, $\Theta =]0,+\infty[^2$ et on considère le modèle statistique $(\Omega, \mathfrak{F}, (P_\theta)_{\theta\in\Theta})$, avec un échantillon associé $X_1,\dots,X_n$, où pour tout $\theta \in \Theta$, les X_i sont P_θ indépendantes et de même loi sous P_θ, donnée par la fonction de survie $G_\theta = G_{a,b}$ définie par (6). On commence par estimer a. Pour cela on introduit la statistique :

$$M_n := \min_{1\leq i\leq n} X_i.$$

Il est facile d'exprimer $P_\theta(M_n > t)$ en fonction de $G_\theta(t)$. En effet pour tout $t > 0$,

$$\{M_n > t\} = \{\forall k = 1,\dots,n,\ X_k > t\} = \bigcap_{k=1}^n \{X_k > t\}.$$

Or les X_k sont P_θ-indépendantes et de même loi sous P_θ, par conséquent

$$P_\theta(M_n > t) = \prod_{k=1}^n P_\theta(X_k > t) = P_\theta(X_1 > t)^n = G_\theta(t)^n$$

et ces égalités sont vérifiées pour tout $t > 0$ et tout $\theta \in \Theta$.

Calculons pour tout $\theta \in \Theta$ et tout entier $n \geq 1$, $P_\theta(M_n \geq a)$. Cette probabilité est la limite à gauche au point a de la fonction de survie de M_n. Or pour $0 < t < a$, $P_\theta(M_n > t) = G_\theta(t)^n = 1^n = 1$, donc cette limite à gauche vaut 1. Autrement dit, pour tout $n \geq 1$ et tout $\theta \in \Theta$, M_n est P_θ-presque-sûrement supérieur ou égal à a, d'où $P_\theta(|M_n - a| = M_n - a) = 1$. Cette remarque nous permet d'exprimer simplement $P_\theta(|M_n - a| > \varepsilon)$ pour $\varepsilon > 0$ quelconque :

$$P_\theta(|M_n - a| > \varepsilon) = P_\theta(M_n - a > \varepsilon) = P_\theta(M_n > a+\varepsilon) = G_\theta(a+\varepsilon)^n = \exp\big(-nb(a+\varepsilon)\big).$$

On voit immédiatement que cette dernière expression tend vers 0 quand n tend vers $+\infty$, et ce pour tout $\varepsilon > 0$. On en déduit que M_n converge vers a en P_θ-probabilité.

En fait, on peut faire bien mieux puisque

$$\sum_{n=1}^{+\infty} P_\theta(|M_n - a| > \varepsilon) = \sum_{n=1}^{+\infty} \exp\big(-nb(a+\varepsilon)\big)$$

est la somme d'une série géométrique de raison $q = \exp(-b(a+\varepsilon))$ positive et strictement inférieure à 1 car $b(a+\varepsilon) > 0$. Donc cette série converge pour tout $\varepsilon > 0$ et tous $a, b > 0$ donc pour tout $\theta \in \Theta$, ce qui est une condition *suffisante* de convergence P_θ-p.s. de M_n vers a :

$$\forall \theta \in \Theta, \quad M_n \xrightarrow[n\to+\infty]{P_\theta-\text{p.s.}} a. \tag{10}$$

Ainsi M_n est un estimateur fortement consistant de a.

4) Pour estimer b, on propose compte-tenu de (9), la statistique :

$$T_n := \frac{2(S_n - nM_n)}{1 + \sum_{k=1}^n (X_k - M_n)^2}, \quad \text{où} \quad S_n := \sum_{k=1}^n X_k.$$

Si on veut exploiter (9), la première idée qui vienne à l'esprit est d'estimer b par

$$T'_n := \frac{2(S_n - na)}{\sum_{k=1}^n (X_k - a)^2} = \frac{2\left(\frac{S_n}{n} - a\right)}{\frac{1}{n}\sum_{k=1}^n (X_k - a)^2}.$$

En effet, les deux suites $(X_k)_{k\geq 1}$ et $\big((X_k - a)^2\big)_{k\geq 1}$ étant i.i.d. et intégrables vérifient la loi forte des grands nombres. Ainsi, le numérateur dans la deuxième expression de T'_n ci-dessus converge p.s. vers $2(\mathbf{E}X - a) = 2\mathbf{E}(X - a)$ et le dénominateur converge p.s. vers $\mathbf{E}(X - a)^2$. Par conséquent T'_n converge p.s. vers le quotient de ces deux limites, c'est-à-dire b d'après (9).

On peut formuler ici deux objections. La première est que l'on ne connaît pas a, donc T'_n n'est pas calculable à partir des seules valeurs observées de l'échantillon : ce n'est pas une *statistique*. Un autre problème est que T'_n peut s'écrire sous la forme indéfinie « $\frac{0}{0}$ », lorsque toutes les valeurs de l'échantillon sont égales à a, cet évènement étant de probabilité non nulle[3]. Il est facile de résoudre cette difficulté en rajoutant au dénominateur une quantité déterministe strictement positive, tendant vers 0 quand n tend vers l'infini, disons $1/n$. Pour répondre à la première objection, on va remplacer a dans T'_n par M_n qui est un estimateur fortement consistant de a d'après (10). Tout ceci nous conduit à estimer b par

$$\frac{2\left(\frac{S_n}{n} - M_n\right)}{\frac{1}{n} + \frac{1}{n}\sum_{k=1}^n (X_k - M_n)^2} = \frac{2(S_n - nM_n)}{1 + \sum_{k=1}^n (X_k - M_n)^2} = T_n.$$

Il nous reste maintenant à montrer que T_n converge P_θ-presque sûrement vers b pour tout $\theta \in \Theta$. Fixons θ quelconque dans Θ. La loi forte des grands nombres appliquée à $(X_k)_{k\geq 1}$ et (10) nous fournissent deux évènements Ω_1 et Ω_2 (dépendants de θ) tels que :

$$P_\theta(\Omega_1) = 1 \quad \text{et} \quad \forall \omega \in \Omega_1, \quad \frac{S_n(\omega)}{n} \xrightarrow[n\to+\infty]{} \mathbf{E}X_1 = \mathbf{E}X \tag{11}$$

3. D'après le résultat de la question 1), cette probabilité vaut $(1 - \exp(-ab))^n$.

et

$$P_\theta(\Omega_2) = 1 \quad \text{et} \quad \forall \omega \in \Omega_2, \quad M_n(\omega) \xrightarrow[n\to+\infty]{} a. \tag{12}$$

Pour le dénominateur de T_n, les choses sont un peu plus compliquées : on ne peut pas appliquer directement la loi forte des grands nombres aux variables aléatoires $(X_k - M_n)^2$. En effet, même si ces variables sont de même loi (pourquoi ?) et intégrables, elles ne sont en général pas indépendantes puisque M_n dépend de tous les X_i pour $1 \leq i \leq n$. Une idée naturelle est alors de *séparer* X_k et M_n en écrivant :

$$(X_k - M_n)^2 = \big((X_k - a) - (M_n - a)\big)^2 = (X_k - a)^2 - 2(M_n - a)(X_k - a) + (M_n - a)^2,$$

d'où

$$\begin{aligned}\frac{1}{n}\sum_{k=1}^{n}(X_k - M_n)^2 &= \frac{1}{n}\sum_{k=1}^{n}(X_k - a)^2 - \frac{2}{n}\sum_{k=1}^{n}(X_k - a)(M_n - a) + \frac{1}{n}\sum_{k=1}^{n}(M_n - a)^2 \\ &= \frac{1}{n}\sum_{k=1}^{n}(X_k - a)^2 - 2(M_n - a)\frac{1}{n}\sum_{k=1}^{n}(X_k - a) + (M_n - a)^2 \qquad (13)\end{aligned}$$

La loi forte des grands nombres appliquée à la suite i.i.d. $(Y_k)_{k\geq 1} = \big((X_k - a)^2\big)_{k\geq 1}$, intégrable puisque $\mathbf{E}|Y_1| = \mathbf{E}(X-a)^2 < +\infty$, nous donne un évènement Ω_3 dépendant de θ , tel que :

$$P_\theta(\Omega_3) = 1 \quad \text{et} \quad \forall \omega \in \Omega_2, \quad \frac{1}{n}\sum_{k=1}^{n}(X_k(\omega) - a)^2 \xrightarrow[n\to+\infty]{} \mathbf{E}(X_1 - a)^2 = \mathbf{E}(X-a)^2. \tag{14}$$

Posons $\Omega' = \Omega_1 \cap \Omega_2 \cap \Omega_3$. Alors $P_\theta(\Omega') = 1$ et en reportant dans (13) les convergences (11), (12) et (14), on obtient :

$$\forall \omega \in \Omega', \quad \frac{1}{n}\sum_{k=1}^{n}\left(X_k(\omega) - M_n(\omega)^2\right) \xrightarrow[n\to+\infty]{} \mathbf{E}(X_1 - a)^2 = \mathbf{E}(X-a)^2 - 2\times 0 \times 0 + 0^2,$$

d'où

$$\forall \omega \in \Omega', \quad \frac{1}{n} + \frac{1}{n}\sum_{k=1}^{n}\left(X_k(\omega) - M_n(\omega)^2\right) \xrightarrow[n\to+\infty]{} \mathbf{E}(X-a)^2. \tag{15}$$

Compte-tenu de (11), (12) et (15) et en notant que la limite $\mathbf{E}(X-a)^2$ dans (15) est non nulle, cf. (8), on aboutit à :

$$\forall \omega \in \Omega', \quad T_n(\omega) = \frac{2\left(\frac{S_n(\omega)}{n} - M_n(\omega)\right)}{\frac{1}{n} + \frac{1}{n}\sum_{k=1}^{n}(X_k(\omega) - M_n(\omega))^2} = \frac{2(\mathbf{E}X - a)}{\mathbf{E}(X-a)^2} = b.$$

Puisque $P_\theta(\Omega') = 1$, nous avons montré que

$$T_n \xrightarrow[n\to+\infty]{P_\theta-\text{p.s.}} b.$$

Comme θ était quelconque, ceci est vrai pour tout $\theta \in \Theta$ et T_n est bien un estimateur fortement consistant de b.