STATISTIQUE COMPUTATIONNELLE – 5

ESTIMATION PONCTUELLE

L'estimation ponctuelle d'une grandeur (caractéristique) statistique d'une population est le calcul d'une estimée de la valeur de cette grandeur. Cette estimée est associée à un estimateur. En règle générale, on cherche à calculer l'estimée la plus précise, c-à-d. l'estimateur le meilleur.

5.1 Estimation et estimateur

On s'intéresse à la caractéristique X d'une population (éventuellement à un vecteur de caractéristiques), dont la loi dépend d'un paramètre inconnu $\Theta \in \mathcal{G} \subset \mathbb{R}^m$.

On note par

- $P_\Theta(X = x)$ la loi de X au point x si la v.a. X est discrète ;
- $f_\Theta(x)$ la densité de la loi de X au point x si la v.a. X est continue.

On effectue un tirage de n articles d'une population de taille potentiellement infinie, noté $(x_1, \ldots, x_n)$. On note $(X_1, \ldots, X_n)$ l'échantillon aléatoire associé à ce tirage (il s'agit d'un vecteur aléatoire dont une réalisation particulière est $(x_1, \ldots, x_n)$).

Avoir une estimation de la valeur θ de la caractéristique Θ consiste à avoir une valeur approchée $\widehat{\theta}$ de cette caractéristique en utilisant le tirage.

DÉFINITION 5.0.7 *Soit $\Theta \in \mathcal{G} \subset \mathbb{R}^m$ la caractéristique de valeur inconnue de la loi qu'elle suit la v.a. X.*

Un estimateur $\widehat{\Theta}_n$ de Θ est une statistique de l'échantillon

$$\widehat{\Theta}_n = g(X_1, \ldots, X_n)\,; \quad g : \mathbb{R}^n \to \mathbb{R}^m$$

telle que pour chaque échantillon $(x_1, \ldots, x_n)$ la valeur $\widehat{\theta}_n = g(x_1, \ldots, x_n)$ de $\widehat{\Theta}_n$ est proche de la valeur θ de Θ.

$\widehat{\theta}_n$ est une estimation de θ.

À titre d'exemple considérons la moyenne μ_X d'une population. On a donc $\Theta = \mu_X$. Soit $(X_1, \ldots, X_n)$ un échantillon de la population. On prendra comme estimateur de la moyenne μ_X, la moyenne empirique de l'échantillon $\widehat{\Theta}_n = \overline{X}_n = \frac{1}{n} \sum_{k=1}^{n} X_k$. Si $(x_1, \ldots, x_n)$ est une réalisation de l'échantillon, alors l'estimation a comme valeur $\widehat{\theta}_n = \bar{x}_n = \frac{1}{n} \sum_{k=1}^{n} x_k$.

5.2 Critères de qualité d'un estimateur

Pour choisir le meilleur estimateur il faut se doter de critères de qualité pour les estimateurs.

Estimateur non biaisé

Lorsqu'on prélève plusieurs échantillons de taille n de la population, on obtient plusieurs valeurs différentes de l'estimation $\widehat{\theta}_n$ de la caractéristique statistique que nous voulons estimer. Ces valeurs sont des valeurs aléatoires, dans la mesure où l'échantillonnage se fait de manière aléatoire. On peut donc calculer son espérance. Si on souhaite que les estimations nous fournissent la valeur exacte θ de cette caractéristique, on peut imposer que l'estimateur soit *sans biais*, c-à-d. que $E\left(\widehat{\Theta}_n\right) = \Theta$.

Estimateur efficace

Une estimation $\widehat{\theta}_n$ est d'autant plus fiable que sa variance $V\left(\widehat{\theta}_n\right)$ (c-à-d. la dispersion de ces valeurs autour de $E\left(\widehat{\theta}_n\right)$) est petite. Un estimateur est *efficace* si la variance de ses estimations est minimale.

Estimateur convergent

Un estimateur est *convergent* si sa variance tend vers zéro lorsque le nombre d'échantillons tend vers l'infini : $\lim_{n\to\infty} V\left(\widehat{\Theta}_n\right) = 0$.

5.3 Précision d'un estimateur

On mesure l'erreur quadratique moyenne (EQM) ou risque d'une estimation $\widehat{\theta}_n$ à l'aide de la formule

$$R_\Theta\left(\widehat{\theta}_n\right) = E\left[\left(\widehat{\theta}_n - \theta\right)^2\right] = V\left(\widehat{\theta}_n\right) + \left(E\left(\widehat{\theta}_n\right) - \theta\right)^2$$

Nous savons que

$$B_\Theta\left(\widehat{\theta}_n\right) = E\left(\widehat{\theta}_n\right) - \theta$$

représente le biais de l'estimation. Donc l'EQM de l'estimateur est la variance de l'estimation plus son biais au carré :

$$R_\Theta\left(\widehat{\theta}_n\right) = V\left(\widehat{\theta}_n\right) + B_\Theta^2\left(\widehat{\theta}_n\right)$$

Par conséquent un estimateur est d'autant plus précis qu'il est de variance minimale est sans biais. On appellera cet estimateur *estimateur sans biais de variance minimale*. Si cet estimateur existe, il est unique presque sûrement.

En règle générale on cherche un estimateur d'EQM minimale. Dans la plupart de cas il faut faire un compromis entre la valeur la variance de l'estimateur et son biais.

5.4 Exemples d'estimateurs

Nous donnons dans la suite une liste d'estimateurs des principales caractéristiques (grandeurs) statistiques.

Espérance $E(X) = \mu_X$ de la v.a. X C'est la moyenne empirique $\overline{X}_n = \frac{1}{n}\sum_{k=1}^{n} X_k$, c-à-d. nous avons pour l'estimateur $\widehat{\Theta}_n = \overline{X}_n$. Pour un échantillon donné $(x_1, \ldots, x_n)$ on a $\overline{x}_n = \frac{1}{n}\sum_{k=1}^{n} x_k$ et donc l'estimation est $\widehat{\theta}_n = \overline{x}_n$.

Nous avons $E\left(\overline{X}_n\right) = E(X) = \mu_X$ et $V\left(\overline{X}_n\right) = \frac{V(X)}{n} = \frac{\sigma_X^2}{n}$.

Variance $V(X)$ de la v.a. X lorsque $E(X) = \mu_X$ est connue C'est la quantité $T_n^2 = \frac{1}{n}\sum_{k=1}^{n}(X_k - \mu_X)^2$, c-à-d. nous avons pour l'estimateur $\widehat{\Theta}_n = T_n^2$. Pour un échantillon donné $(x_1,\ldots,x_n)$ on a $t_n^2 = \frac{1}{n}\sum_{k=1}^{n}(x_k-\mu_X)^2$et donc l'estimation de la variance est $\widehat{\theta}_n = t_n^2$.

Nous avons $E\left(T_n^2\right) = \sigma_X^2$ et $V\left(T_n^2\right) = \frac{\mu_X^4-\sigma_X^4}{n}$.

Variance $V(X)$ de la v.a. X lorsque $E(X)=\mu_X$ est inconnue C'est la variance empirique $S_n^2 = \frac{1}{n}\sum_{k=1}^{n}\left(X_k-\overline{X}_n\right)^2$, c-à-d. nous avons pour l'estimateur $\widehat{\Theta}_n = S_n^2$.

Pour un échantillon donné $(x_1,\ldots,x_n)$ on a $s_n^2 = \frac{1}{n}\sum_{k=1}^{n}(x_k-\overline{x}_n)^2$et donc l'estimation de la variance est $\widehat{\theta}_n = s_n^2$.

Nous avons $E\left(S_n^2\right) = \frac{n-1}{n}\sigma_X^2$ et $V\left(T_n^2\right) = \frac{n-1}{n^3}\left((n-1)\mu_X^4-(n-3)\sigma_X^4\right)$.

Si la variance empirique est corrigée (c-à-d. elle est sans biais), alors $S_n^2 = \frac{1}{n-1}\sum_{k=1}^{n}\left(X_k-\overline{X}_n\right)^2$ et dans ce cas $E\left(S_n^2\right) = \sigma_X^2$ et $V\left(T_n^2\right) = \frac{1}{n}\left(\mu_X^4 - \frac{n-3}{n-1}\sigma_X^4\right)$.

Covariance $C(X,Y)$ entre deux v.a. X,Y lorsque $E(X)=\mu_X$ et $E(Y)=\mu_Y$ sont connues $\widehat{\Theta}_n = cov_n(X,Y) = \frac{1}{n}\sum_{k=1}^{n}(X_k-\mu_X)(Y_k-\mu_Y)$ est un estimateur de la covariance.

Pour deux échantillons donnés $(x_1,\ldots,x_n)$ et $(y_1,\ldots,y_n)$ on a que $cov_n(x,y) = \frac{1}{n}\sum_{k=1}^{n}(x_k-\mu_X)(y_k-\mu_Y)$ est la valeur de l'estimation et donc $\widehat{\theta}_n = cov_n(x,y)$.

D'après ce que nous venons de voir les estimateurs de l'espérance et celui de la variance corrigée sont sans biais. Par contre la variance de l'estimateur de la variance est plus petite que la variance de l'estimateur de la variance corrigée.

Covariance $C(X,Y)$ entre deux v.a. X,Y lorsque $E(X)=\mu_X$ et $E(Y)=\mu_Y$ sont inconnues $\widehat{\Theta}_n = \widehat{cov}_n(X,Y) = \frac{1}{n}\sum_{k=1}^{n}\left(X_k-\overline{X}_n\right)\left(Y_k-\overline{Y}_n\right)$. Pour deux échantillons donnés $(x_1,\ldots,x_n)$ et $(y_1,\ldots,y_n)$ on a que $\widehat{cov}_n(x,y) = \frac{1}{n}\sum_{k=1}^{n}(x_k-\overline{x})(y_k-\overline{y})$ est l'estimation et donc $\widehat{\theta}_n = \widehat{cov}_n(x,y)$.

Corrélation $\rho(X,Y)$ entre deux v.a. X,Y $\widehat{\Theta}_n = r_n(X,Y) = \frac{\widehat{cov}(X,Y)}{S_n(X)S_n(Y)}$. Pour deux échantillons donnés $(x_1,\ldots,x_n)$ et $(y_1,\ldots,y_n)$ on a que $r_n(x,y) = \frac{\sum_{k=1}^{n}(x_k-\overline{x})(y_k-\overline{y})}{\sum_{k=1}^{n}(x_k-\overline{x}_n)^2\sum_{k=1}^{n}(y_k-\overline{y}_n)^2}$ est la valeur de l'estimation et donc $\widehat{\theta}_n = r_n(x,y)$. Notons que $r_n(x,y) \in [-1,1]$.

5.5
Estimateur des moindres carrés

On cherche à calculer l'estimation $\widehat{\theta}_n$ sur la base d'une réalisation $(x_1,\ldots,x_n)$ de v.a. $(X_1,\ldots,X_n)$. La méthode des moindres carrés consiste à minimiser la somme des carrés des écarts $x_k - \widehat{\theta}_n$ entre les valeurs observées x_k et l'estimation $\widehat{\theta}_n$ de l'estimateur, c-à-d. on cherche à minimiser la quantité

$$Q\left(\widehat{\theta}_n\right) = \sum_{k=1}^{n} \left(x_k - \widehat{\theta}_n\right)^2$$

Au minimum nous avons

$$\frac{dQ}{d\widehat{\theta}_n} \equiv -2\sum_{k=1}^{n} \left(x_k - \widehat{\theta}_n\right) = 0$$

d'où on peut calculer $\widehat{\theta}_n$. De plus la valeur de $\widehat{\theta}_n$ doit vérifier la relation

$$\frac{d^2Q}{\left(d\widehat{\theta}_n\right)^2} > 0$$

Cette méthode ne donne pas de bons résultats si n est modéré. De plus ces estimateurs ne sont pas convergents.

5.6 Exercices

EXERCICE 5.1 *À l'occasion d'un test qualité, on a examiner 600 articles et on a trouvé 78 défaillants. On cherche à évaluer la probabilité p qu'un article soit défaillant.*

(1) *Proposez un estimateur pour p et justifier votre réponse. Calculer l'estimation qui correspond au résultat du test.*

(2) *Trouver la loi suivie par cet estimateur. Est-il possible d'utiliser l'approximation de la loi normale pour effectuer les calculs numériques ?*

SOL.- *X_i suit la loi de Bernoulli $B(1, 0.13)$. Les X_i sont des v.a.i.i.d. L'estimateur est la somme $\overline{X} = \sum_{i=1}^{600} X_i$. Il suit donc la loi binomiale $\mathcal{B}\ (600, 0.13)$.*

Attention : Il n'est pas dans le poly. Si $n > 30$ et $np < 15$, alors la loi $\mathcal{B}(n,p)$ converge vers la loi $\mathcal{P}(np)$. Si $np(1-p) > 15$, alors la loi $\mathcal{B}(n,p)$ converge vers la loi $\mathcal{N}(np, np(1-p))$.

Ici nous pouvons utiliser l'approximation normale car $np(1-p) = 67.4 > 15$.

EXERCICE 5.2 *Lors d'un test qualité on trouve 40% d'articles hors normes. Calculer la probabilité que dans un nouveau test portant sur 400 articles, au moins 150 articles soient hors normes.*

SOL.- Modélisation : A tout individu k de l'échantillon on associe la variable aléatoire X_k prenant la valeur 1 si l'article k est hors normes et 0 sinon. On adopte le modèle statistique suivant : les X_k sont indépendantes et identiquement distribuées selon une loi de Bernouilli de paramètre $p = 0.4$. Donc $X_k \sim B(1,p)\,; k = 1,\ldots,n$. Par conséquent $\sum\limits_{k=1}^{n} X_k \sim \mathcal{B}(n,p)$ loi binomiale car les X_k sont des v.a.i.i.d.

Pour un échantillon donné, soit $\overline{X}_n = \frac{1}{n}\sum\limits_{k=1}^{n} X_k$ la moyenne empirique d'articles hors normes dans l'échantillon. On a donc $n\overline{X}_n \sim \mathcal{B}(n,p)$. Comme $n = 400$, les tables de la loi binomiale n'ont pas la probabilité correspondante. On doit utiliser l'approximation par la loi normale car $np(1-p) = 400 \times 0.4 \times 0.6 = 96 > 15$. Par conséquent approximativement on a $n\overline{X}_n \sim \mathcal{N}(np, np(1-p))$ d'où $\overline{X}_n \sim \mathcal{N}\left(p, \frac{p(1-p)}{n}\right)$. On cherche donc la probabilité $P\left(\overline{X}_n \geq \frac{150}{400} = 0.375\right)$.

Pour lire les tables de la loi normale, il faut passer à la forme standarisée de la v.a. $\overline{X}_n$. En standarisant on a que la v.a. standarisée $Z_n = \frac{\overline{X}_n - p}{\sqrt{\frac{p(1-p)}{n}}}$ de la v.a. $\overline{X}_n$ suit la loi normale $\mathcal{N}(0,1)$. On cherche donc à calculer la probabilité $P\left(\frac{\overline{X}_n - 0.4}{\sqrt{\frac{0.4(1-0.4)}{400}}} \geq \frac{0.375-0.4}{\sqrt{\frac{0.4(1-0.4)}{400}}}\right)$. On a que $\frac{0.375-0.4}{\sqrt{\frac{0.4(1-0.4)}{400}}} = -1.02$. Par conséquent $P\left(\overline{X}_n \geq 0.375\right) = P(Z_n \geq 1.02) = 1 - F(-1.02) = 0.85$ où F désigne la fonction de réaprtition de $\mathcal{N}(0,1)$.

EXERCICE 5.3 *On considère un échantillon X_1, X_3, X_n tiré d'une population de moyenne μ_X et de variance σ_X^2. On propose pour la moyenne les deux estimateurs suivants :*

- $\overline{x}_1 = \frac{x_1 + 2x_2 + 3x_3}{6}$
- $\overline{x}_2 = \frac{x_1 + x_2 + x_3}{3}$

Comparer ces deux estimateurs du point de vue de leur biais et de leur efficacité.

SOL.- $E\left(\overline{X}_1\right) = E\left(\frac{x_1+2x_2+3x_3}{6}\right) = \frac{1}{6}E(x_1 + 2x_2 + 3x_3) = \frac{6}{6}\mu_X = \mu_X$. *Donc biais* $b_1 = 0$.

$E\left(\overline{X}_2\right) = E\left(\frac{x_1+x_2+x_3}{3}\right) = \mu_X$. *Donc biais $b_2 = 0$, c-à-d. les deux estimateurs sont non biaisés.*

Variance : $V\left(\overline{X}_1\right) = V\left(\frac{x_1+2x_2+3x_3}{6}\right) = \frac{1}{36}V(x_1 + 2x_2 + 3x_3) = \frac{14}{36}\sigma_X^2$, $V\left(\overline{X}_2\right) = V\left(\frac{x_1+x_2+x_3}{3}\right) = \frac{3}{9}\sigma_X^2$. *Donc $V\left(\overline{X}_2\right) < V\left(\overline{X}_1\right)$, c-à-d. l'estimateur $\overline{X}_2$ est plus efficace que l'estimateur $\overline{X}_2$.*

EXERCICE 5.4 *Soit $X \sim \mathcal{N}(\mu,\sigma)$; $\mu = \sigma = \theta$ inconnu. On veut estimer* θ.

Soient $T = X$ et $U = \frac{X}{2}$ deux estimateurs.

Comparer ces deux estimateurs.

SOL.- $E(T) = E(X) = \theta \Rightarrow$ sans biais, c-à-d. $B_\Theta\left(\widehat{\theta}_n\right) = 0$. $V(T) = V(X) = \sigma^2 = \theta^2$. Risque $R_\Theta\left(\widehat{\theta}_n\right) = V\left(\widehat{\theta}_n\right) + B_\Theta^2\left(\widehat{\theta}_n\right) = \theta^2$.

$E(U) = \frac{1}{2}\theta \Rightarrow$ avec biais, $B_\Theta\left(\widehat{\theta}_n\right) = -\frac{1}{2}\theta$. $V(U) = \frac{1}{4}V(X) = \frac{1}{4}\theta^2$. Risque $R_\Theta\left(\widehat{\theta}_n\right) = V\left(\widehat{\theta}_n\right) + B_\Theta^2\left(\widehat{\theta}_n\right) = \frac{1}{2}\theta^2$.

On ne peut pas comparer deux estimateurs dont l'un est biaisé et l'autre non. Le choix se fera en fonction de l'application. Si on veut privilegier le caractère non biaisé, on préferera le prmier. Si le critère du risque est important, on utilisera le second qui assure un rique plus petit que le premier.

EXERCICE 5.5 *Soit X une v.a. de densité*

$$f(x;\theta) = \begin{cases} \frac{2}{\theta}(1 - \frac{x}{\theta}) & si \quad 0 \le x \le \theta, \\ 0 & sinon \end{cases}$$

où θ est un paramètre inconnu strictement positif.

(1) *Calculer $E(X)$ et $V(X)$*

SOL.- $E(X) = \int_0^\theta x f(x;\theta)dx = \int_0^\theta x\left(\frac{2}{\theta}\left(1 - \frac{x}{\theta}\right)\right)dx = \frac{\theta}{3}$. $V(X) = \int_0^\theta x^2\left(\frac{2}{\theta}\left(1 - \frac{x}{\theta}\right)\right)dx - \left(\frac{\theta}{3}\right)^2 = \frac{\theta^2}{18}$

(2) *On cherche à estimer θ à partir d'un échantillon $(X_1,...,X_n)$ de v.a. indépendantes, de même loi que X. On utilise $T_n = \overline{X}$ comme estimateur.*

(a) *Calculer le biais de cet estimateur. Pouvez vous lui éliminer son biais ?*

SOL.- $E(T_n) = E\left(\overline{X}\right) = \frac{1}{n}E\left(\sum_{k=1}^{n} X_k\right) = \frac{\theta}{3} \Rightarrow$ *avec biais,* $B_\Theta\left(\widehat{\theta}_n\right) = E(T_n) - \theta = \frac{2\theta}{3}$. *Pour éliminer le biais on doit le multiplier par 3, c-à-d. prendre* $T_n = 3\overline{X}$.

(b) *Établir que T_n modifié sans biais, est un estimateur de θ convergent.*

SOL.- $V(T_n) = V\left(3\overline{X}\right) = 9V\left(\frac{1}{n}\sum_{k=1}^{n} X_k\right) = \frac{9}{n}\cdot\frac{\theta^2}{18} = \frac{\theta^2}{2n} \to_{n\to\infty} 0.$

(c) *Calculer $V(T_n)$ et en conclure ue T_{qn} converge en moyenne quadratique vers θ.*

SOL.- *Convergence en moyenne quadratique :* $E(T_n - E(T_n))^2 \to_{n\to\infty} 0$ *ce qui donne* $E(T_n - \theta)^2 = V(T_n) = \frac{\theta^2}{2n} \to_{mq} 0.$

(d) *Déterminer la loi asymptotique de T_n.*

***SOL.**- Loi asymptotique :* $\overline{X} = \frac{1}{n}\sum_{k=1}^{n} X_k$. *Donc* $\overline{X} \sim \mathcal{N}\left(\mu_X, \sigma_X^2\right) \Rightarrow T_n = 3\,\overline{X}$ *avec* $E(T_n) = \theta$, $V(T_n) = \frac{\theta^2}{2n}$. *Pour* $n \to \infty$ *on a* $T_n \sim \mathcal{N}\left(\mu_{T_n}, \sigma_{T_n}^2\right) \Rightarrow \frac{T_n - \mu_{T_n}}{\sigma_{T_n}} \sim \mathcal{N}(0,1)$.

EXERCICE 5.6 *Nous avons un échantillon de* n *tirages* $X_1, \ldots, X_n$ *de v.a.i.i.d. qui suivent la loi normale* $\mathcal{N}\left(\mu_X, \sigma_X^2\right)$. *On voudrait estimer* μ_X *et prévoir le prochain tirage* X_{n+1}. *La variance est connue et égale à 4.*

(1) *Montrer que l'estimateur sans biais le plus précis de* μ_X *est la moyenne empirique* $\overline{X}_n = \frac{1}{n}\sum_{k=1}^{n} X_k$ *et préciser sa loi.*

***SOL.**- On considère l'estimateur* $\widehat{\Theta}_n = \overline{X}_n$. *Cet estimateur est sans biais :* $E\left(\overline{X}_n\right) = \frac{1}{n}E\left(\sum_{k=1}^{n} X_k\right) = \frac{1}{n}nE(X) = \mu_X$. *La variance de l'estimateur est* $V\left(\overline{X}_n\right) = V\left(\frac{1}{n}\sum_{k=1}^{n} X_k\right) = \frac{1}{n^2}nV(X) = \frac{V(X)}{n} = \frac{\sigma_X^2}{n} = \frac{4}{n}$.

Pour que l'esimateur soit le plus précis il faut que son risque $R_\Theta\left(\widehat{\theta}_n\right) = V\left(\widehat{\theta}_n\right) + B_\Theta^2\left(\widehat{\theta}_n\right)$ *soit minimal. Puisque* $B_\Theta^2\left(\widehat{\theta}_n\right) = 0$, *il faut que sa variance* $V\left(\widehat{\theta}_n\right)$ *soit minimale. Il faut maintenant montrer que la variance de tout autre estimateur sans biais soit plus grande.*

Remarquons d'abord qu'un estimateur sans biais de la moyenne ne peut être que combinaison linéaire de v.a. $X_1, \ldots, X_n$. *En effet si pour deux v.a.i.i.d.* X *et* Y *on prend comme estimateur* $\widehat{\Theta}_n = aXY$, *alors* $E\left(\widehat{\Theta}_n\right) = E(aXY) = aE(X)E(Y) = a\mu^2 \neq \mu$. *Considérons donc un estimateur sans biais du type* $\widehat{\Theta}_n = \frac{\sum_{k=1}^{n} a_k X_k}{\sum_{k=1}^{n} a_k}$. *Nous allons montrer que* $V\left(\widehat{\Theta}_n\right) = \frac{\sum_{k=1}^{n} a_k^2}{\left(\sum_{k=1}^{n} a_k\right)^2}\sigma_X^2 \geq \frac{\sigma_X^2}{n}$. *Pour* $n = 1$ *il y a l'égalité. Pour* $n = 2$ *on a* $\frac{a_1^2 + a_2^2}{(a_1 + a_2)^2}\sigma_X^2 \geq \frac{\sigma_X^2}{2} \Rightarrow 2\left(a_1^2 + a_2^2\right) \geq (a_1 + a_2)^2 \Rightarrow (a_1 - a_2)^2 \geq 0$. *On admet pour* $n = \ell$: $\frac{a_1^2 + \cdots + a_\ell^2}{(a_1 + \cdots + a_\ell)^2}\sigma_X^2 \geq \frac{\sigma_X^2}{\ell} \Rightarrow B_\ell \geq A_\ell^2$ *avec* $B_\ell = \ell\left(a_1^2 + \cdots + a_\ell^2\right)$ *et* $A_\ell = a_1 + \cdots + a_\ell$. *Il faut montrer que* $B_\ell + \sum_{k=1}^{\ell} a_k^2 + (\ell + 1)a_{\ell+1}^2 \geq (A_\ell + a_{\ell+1})^2$. *On a* $B_\ell + \sum_{k=1}^{\ell} a_k^2 + (\ell + 1)a_{\ell+1}^2 \geq A_\ell^2 + 2A_\ell a_{\ell+1} + a_{\ell+1}^2$ *d'où* $\sum_{k=1}^{n} a_k^2 + \ell a_{\ell+1}^2 \geq 2A_\ell a_{\ell+1} \Rightarrow \sum_{k=1}^{\ell}(a_k - a_{\ell+1})^2 \geq 0$. *Par conséquent la moyenne empirique est l'estimateur sans bias de variance minimale de la moyenne théorique.*

(2) *Le prochain tirage* X_{n+1} *est supposé indépendant des* $X_1, \ldots, X_n$ *et de même loi. Donner la meilleure prévision de* X_{n+1}.

SOL.- *La meilleure prévision de* X_{n+1} *est celle qui minimise l'EQM. Par conséquent la meilleure prévision pour* X_{n+1} *est la moyenne empirique* $\overline{X}_n$.

(3) *Trouver la loi de la v.a.* $Z_{n+1} = X_{n+1} - \overline{X}_n$ *qui représente l'erreur de la prévision.*

SOL.- Z_{n+1} *suit la loi normale en tant que combinaison linéaire de deux variables indépendantes suivant une loi normale (* X_{n+1} *et* $\overline{X}_n$ *sont indépendantes car* $\overline{X}_n$ *est la combinaison linéaire de variables indépendantes de* X_{n+1} *à savoir* $X_1, \ldots, Xn$*). L'espérance de* Z_{n+1} *est* $E(Z_{n+1}) = \mu_X - \mu_X = 0$ *et sa variance* $V(Z_{n+1}) = V(X_{n+1}) + V\left(\overline{X}_n\right)$ *car* X_{n+1} *et* $\overline{X}_n$ *sont indépendantes. D'où* $V(Z_{n+1}) = \sigma_X^2 + \frac{\sigma_X^2}{n} = 4 + \frac{4}{n} = 4\frac{n+1}{n}$. *Donc* $Z_{n+1} \sim \mathcal{N}\left(0, 4\frac{n+1}{n}\right)$.