TD N°1 : Propriétés des estimateurs et estimateurs usuels

L'objectif de ce TD est d'étudier les propriétés d'un estimateur et de comparer deux estimateurs entre eux. Cette étude est faite d'un point de vue théorique mais aussi empirique à l'aide de simulations. Pour se faire, on utilisera le logiciel R.

Exercice 1

Soit X le temps d'attente avant d'être servi au RU. Une série d'observations journalières donne l'échantillon suivant (en min)

10	11	15	8	4	22	13	2	10	10	7	1	19	2	3
16	1	8	15	3	1	19	2	3	19	9	8	15	5	9

On calcule la moyenne $\overline{x}$ =9 min et l'écart-type empirique s^*=6.3 min.

1) Quelle loi suit X ?
2) Comment déterminer le paramètre de cette loi à partir de l'échantillon ?
3) En déduire la probabilité qu'un étudiant attende plus de 15 min avant d'être servi.

Exercice 2

Un expérimentateur utilise un thermomètre sophistiqué mais présentant une incertitude de mesure. Afin d'estimer cette incertitude, il effectue une série de mesures sur de l'eau en ébullition. Il souhaite en déduire la probabilité que la température soit mesurée avec une erreur de plus de 1° Celsius.

Soit T la température mesurée.

1) Quelle loi suit T ?
2) Comment déterminer le paramètre de cette loi à partir des mesures ?
3) Considérons que la valeur estimée sur un échantillon de 30 mesures soit 0.25. Quelle est alors la probabilité que la température mesurée avec une erreur supérieure à 1°.

Exercice 3

Soit X une variable aléatoire deb densité

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

On note que E(X)=θ/3 et V(X)=θ^2/18.

On cherche à estimer θ à partir d'un échantillon $X_1,\dots,X_n$ de même loi que X. On utilise l'estimateur

$$T_n = \overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \ .$$

1) Calculer le biais de cet estimateur. Pouvez-vous lui éliminer son biais ?
2) Etablir la convergence en probabilité de T_n (modifié).
3) Déterminer le risque quadratique puis la convergence en moyenne quadratique.

Exercice 4

Soit $X_1,...,X_n$ un échantillon de loi normale d'espérance θ et de variance $\theta(1-\theta)$ où $\theta \in]0,1[$ est un paramètre inconnu. On considère les estimateurs

$$T_1 = \overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \quad \text{et} \quad T_2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 \ .$$

1) Etude théorique des propriétés des estimateurs
   (a) Les estimateurs sont-ils sans biais ?
   (b) Sont-ils convergent ?
   (c) Quel est le meilleur des deux ?
   N.B. On note $V(X_i^2) = 2\theta^2(1-\theta^2)$
2) Illustration des propriétés à l'aide du logiciel R
   (a) Simuler un échantillon de taille n=50 avec θ=0,9 (vérifier la simulation en calculant la moyenne, la variance et en traçant la distribution de l'échantillon). Calculer la valeur de T_1 et T_2. Relancer la simulation et observer les valeurs obtenues pour les estimateurs (dans la fenêtre environment).
   (b) Simuler 30 échantillons de taille n=50 (matrix à 30 lignes et 50 colonnes de réalisation de $N(\theta,\theta(1-\theta))$). Calculer les 30 valeurs de T_1 et T_2. Comparer leur dispersion à l'aide d'un boxplot.
   (c) Que se passe-t-il si θ=0.05 ?
   (d) Que se passe-t-il si n=5000 ?