

# Informatique

## TP 9

### Estimation

#### PARTIE 1 : COMPARAISON DE PLUSIEURS ESTIMATEURS

##### Exercice 1

Étant donné un  $n$ -échantillon  $(X_1, \dots, X_n)$  i.i.d. de loi uniforme sur  $[0, \theta]$ , où  $\theta > 0$  est un paramètre inconnu, on a montré en TD que

$$S_n = \frac{2}{n} \sum_{k=1}^n X_k \quad \text{et} \quad T_n = \frac{n+1}{n} \max(X_1, \dots, X_n)$$

sont deux estimateurs sans biais de  $\theta$ .

Représenter la loi de ces estimateurs pour  $\theta = 1$  et différentes valeurs de  $n$  afin d'en comparer les performances.

#### PARTIE 2 : NIVEAU RÉEL D'UN INTERVALLE DE CONFIANCE

Le niveau réel d'un intervalle de confiance  $[U_n, V_n]$  pour  $g(\theta)$  est le réel

$$\inf_{\theta \in \Theta} \mathbb{P}_\theta (U_n \leq g(\theta) \leq V_n).$$

On peut estimer ce niveau par la méthode de Monte-Carlo.

##### Exercice 2

Étant donnée une suite  $(X_n)_{n \in \mathbb{N}^*}$  de variables aléatoires i.i.d. de loi  $\mathcal{B}(p)$  où  $p \in ]0, 1[$  est un paramètre inconnu, on considère l'intervalle de confiance asymptotique

$$J_n = \left[ \bar{X}_n - \frac{1,96}{2\sqrt{n}}, \bar{X}_n + \frac{1,96}{2\sqrt{n}} \right]$$

de  $p$  au niveau de confiance 95%.

1. Simuler 100 réalisations de l'intervalle  $J_{500}$  et les représenter verticalement au dessus des abscisses entières de 1 à 100 en faisant apparaître en bleu ceux qui contiennent  $p$  et en rouge ceux qui ne contiennent pas  $p$ .

Pour représenter le segment d'extrémités  $A = (x_A, y_A)$  et  $B = (x_B, y_B)$ , on pourra utiliser l'instruction `plot([xA;xB], [yA;yB], opt)` où `opt='b-+'` (resp. `opt='r-+'`) pour un segment bleu (resp. rouge).

2. Déterminer le niveau réel de l'intervalle  $J_{500}$  par la méthode de Monte-Carlo.

#### PARTIE 3 : COMPARAISON DE DIFFÉRENTS INTERVALLES DE CONFIANCE

Étant donnée une suite  $(X_n)_{n \in \mathbb{N}^*}$  de variables aléatoires i.i.d. de loi  $\mathcal{B}(p)$  où  $p \in ]0, 1[$  est un paramètre inconnu, on a construit en classe les trois intervalles de confiance (asymptotiques pour les deux derniers) pour  $p$  au niveau de confiance  $1 - \alpha$  :

$$I_n = \left[ \bar{X}_n - \frac{1}{2\sqrt{n\alpha}}, \bar{X}_n + \frac{1}{2\sqrt{n\alpha}} \right], \quad J_n = \left[ \bar{X}_n - \frac{z_{1-\alpha/2}}{2\sqrt{n}}, \bar{X}_n + \frac{z_{1-\alpha/2}}{2\sqrt{n}} \right]$$

$$\text{et } K_n = \left[ \bar{X}_n - z_{1-\alpha/2} \sqrt{\frac{\bar{X}_n(1-\bar{X}_n)}{n}}, \bar{X}_n + z_{1-\alpha/2} \sqrt{\frac{\bar{X}_n(1-\bar{X}_n)}{n}} \right]$$

où  $z_{1-\alpha/2}$  désigne le quantile d'ordre  $1 - \frac{\alpha}{2}$  de la loi  $\mathcal{N}(0, 1)$  et

$$\bar{X}_n = \frac{1}{n} \sum_{k=1}^n X_k,$$

dont on souhaite comparer les performances.

### Exercice 3

Pour différentes valeurs de  $p \in ]0, \frac{1}{2}]$ , de  $\alpha \in [\frac{1}{2}, 1[$  et  $n = 500$ , simuler  $N = 10^3$  intervalles de confiance des trois types ci-dessus et comparer :

- > le niveau de confiance réel de chaque intervalle ;
- > la demi-largeur moyenne de chaque intervalle.

Commenter.