# Colles semaine 16 - Estimation paramétrique

## 1 Principe de l'estimation paramétrique

#### Notion d'échantillon de la loi étudiée

Un échantillon  $X_1, ..., X_n$  de la loi étudiée, est une famille de variables aléatoires qui sont :

- mutuellement indépendantes,
- ▶ toutes de même loi, dépendant d'un paramètre inconnu *a*.

#### Estimateur

Un **estimateur** est une **statistique** : une variable aléatoire  $\varphi(X_1,...,X_n)$  définie en termes de l'échantillon, qui vise à **estimer** le paramètre inconnu a.

#### Biais, risque quadratique

Notion	(Interprétation)	Formule
Biais	(erreur moyenne)	$b_a(A_n) = \mathbb{E}_a[A_n - a] = \mathbb{E}_a[A_n] - a$
Risque quadratique	(erreur <b>quadratique</b> )	$r_a(A_n) = \mathbb{E}_a\big[(A_n - a)^2\big]$
Décomposition biais-variance (Kænig-Huygens)		$r_a(A_n) = (b_a(A_n))^2 + \operatorname{Var}_a(A_n)$

### 2 Estimateurs usuels

#### > Estimateur de moyenne empirique

La statistique :  $\overline{X_n} = \frac{1}{n} \cdot \sum_{k=1}^n X_k$  est un **estimateur sans biais** de  $\mu = \mathbb{E}[X]$ .

Son risque quadratique est sa variance :  $\frac{1}{n} \cdot \text{Var}(X) = \frac{\sigma^2}{n} \xrightarrow[n \to \infty]{} 0.$ 

- ▶ Estimateurs par min/max On trouve la loi de :
  - ►  $S_n = \max(X_1, ..., X_n)$  par la fonction de répartition :  $\mathbb{P}(S_n \le x) = (\mathbb{P}(X \le x))^n$ .
  - ▶  $I_n = \min(X_1, ..., X_n)$  par la fonction d'anti-répartition :  $\mathbb{P}(I_n > x) = (\mathbb{P}(X > x))^n$ .

(→ calcul de l'espérance/variance et biais/risque)

### 3 Convergence en probabilités

▶ **Inégalité de Markov** Soit X une variable aléatoire  $\ge 0$  admettant une espérance.

Pour A > 0, on a:  $\mathbb{P}(X \ge A) \le \frac{\mathbb{E}[X]}{A}$ .

▶ **Inégalité de Bienaymé-Tchebychev** Soit *X* une *v.a.* ayant un moment d'ordre 2.

Pour  $\lambda > 0$ , on a:  $\mathbb{P}(|X - \mathbb{E}[X]| \ge \lambda) \le \frac{\text{Var}(X)}{\lambda^2}$ .

### ▶ Loi faible des grands nombres

pour un échantillon de loi de X, la moyenne empirique  $\overline{X_n}$  vérifie :  $\mathbb{P}(|\overline{X_n} - \mathbb{E}[X]| \ge \epsilon) \xrightarrow[n \to \infty]{} 0$ .

#### **Estimateur convergent**

la suite d'estimateurs  $(T_n)$  pour  $\theta$  est **convergente** si :  $\mathbb{P}_{\theta}(|T_n - \theta| \ge \epsilon) \xrightarrow[n \to \infty]{} 0$ .

(« pas convergent » ←→ « mauvais » : même avec un échantillon « infini », l'estimation fluctue)

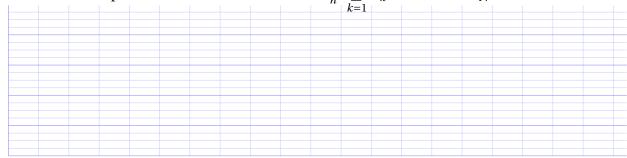
#### Condition suffisante de convergence

Pour que  $(T_n)$  soit convergent, il suffit que le risque quadratique :  $r_{\theta}(T_n) \xrightarrow[n \to \infty]{} 0$ .

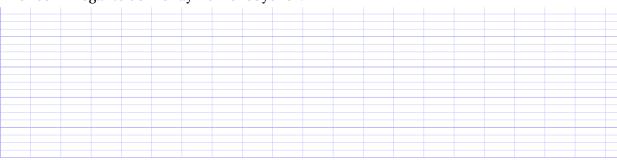
## 4 Questions de cours

1. Déterminer l'espérance et la variance de :  $\overline{X} = \frac{1}{n} \cdot \sum_{k=1}^{n} X_k$ .

(+ hypothèses sur l'échantillon)



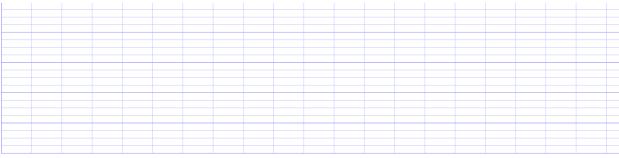
2. Énoncer l'inégalité de Bienaymé-Tchebychev.



**3.** Déterminer la fonction de répartition du max indépendant  $S_n = \max(X_1, ..., X_n)$ .



**4.** Définition et condition suffisante de convergence d'un estimateur.



5. La loi faible des grands nombres.

