Colles semaine 16 - Estimation paramétrique

1 Principe de l'estimation paramétrique

- **Notion d'échantillon de la loi étudiée**
  Un échantillon $X_1, \dots, X_n$ 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 $\varphi(X_1, \dots, X_n)$ 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 **quadratique**)*	$r_a(A_n) = \mathbb{E}_a\left[(A_n - a)^2\right]$
Décomposition biais-variance	*(Kœnig-Huygens)*	$r_a(A_n) = \big(b_a(A_n)\big)^2 + \mathrm{Var}_a(A_n)$

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 : $\frac{1}{n} \cdot \mathrm{Var}(X) = \frac{\sigma^2}{n} \xrightarrow[n\to\infty]{} 0$.
- **Estimateurs par min/max** On trouve la loi de :
  - $S_n = \max(X_1, \dots, X_n)$ par la fonction de répartition : $\mathbb{P}(S_n \leqslant x) = \big(\mathbb{P}(X \leqslant x)\big)^n$.
  - $I_n = \min(X_1, \dots, X_n)$ par la fonction d'anti-répartition : $\mathbb{P}(I_n > x) = \big(\mathbb{P}(X > x)\big)^n$.

($\rightsquigarrow$ calcul de l'espérance/variance et biais/risque)

3 Convergence en probabilités

- **Inégalité de Markov** Soit X une variable aléatoire $\geqslant 0$ admettant une espérance.
  Pour $A > 0$, on a : $\mathbb{P}(X \geqslant A) \leqslant \frac{\mathbb{E}[X]}{A}$.
- **Inégalité de Bienaymé-Tchebychev** Soit X une *v.a.* ayant un moment d'ordre 2.
  Pour $\lambda > 0$, on a : $\mathbb{P}\big(|X - \mathbb{E}[X]| \geqslant \lambda\big) \leqslant \frac{\mathrm{Var}(X)}{\lambda^2}$.
- **Loi faible des grands nombres**
  pour un échantillon de loi de X, la moyenne empirique $\overline{X_n}$ vérifie : $\mathbb{P}\left(\left|\overline{X_n} - \mathbb{E}[X]\right| \geqslant \epsilon\right) \xrightarrow[n\to\infty]{} 0$.
- **Estimateur convergent**
  la suite d'estimateurs (T_n) pour θ est **convergente** si : $\mathbb{P}_\theta\big(|T_n - \theta| \geqslant \epsilon\big) \xrightarrow[n\to\infty]{} 0$.

(« pas convergent » $\longleftrightarrow$ « 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 $S_n = \max(X_1, \ldots, X_n)$.

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

5. La loi faible des grands nombres.