# Stats

**Convergence**

- Loi : $X_n \xrightarrow{L} X \iff F_n(x)$ CVS $F(x)$
- Probabilité : $X_n \xrightarrow{P} X \iff P(|X_n - X| > \varepsilon) \longrightarrow 0$
- Moyenne Quadratique : $X_n \xrightarrow{MQ} X \iff E((X_n - X)^2) \longrightarrow 0$
- Presque sûrement : $X_n \xrightarrow{PS} X \iff X_n(\omega) \longrightarrow X(\omega)$.

**Loi faible des gd nombres**

- $X_1, \dots, X_n$ II, même loi
- $E(X_1) = m < +\infty$

$\Rightarrow \overline{X_n} = \frac{1}{n} \sum_{k=1}^{n} X_k \xrightarrow{P} m$

**Loi forte des gd nombres**

- $X_1, \dots, X_n$ II, même loi
- $E(X_1) = m < +\infty$ $V(X_1) = \sigma^2 < +\infty$

$\Rightarrow \overline{X_n} \xrightarrow{MQ} m$

**Thm central limites**

- $X_1, \dots, X_n$ II, même loi
- $E(X_1) = m < +\infty$, $V(X_1) = \sigma^2 < +\infty$

$\Rightarrow Y_n = \frac{1}{\sqrt{n\sigma^2}} \sum_{k=1}^{n} X_k - nm \xrightarrow{L} N(0,1)$

**Estimateur :** $\hat{\theta}$

- Biais : $b_n(\theta) = E(\hat{\theta}_n) - \theta$
- Variance : $v_n(\theta) = E((\hat{\theta}_n - E(\hat{\theta}_n))^2) = E(\hat{\theta}_n^2) - E(\hat{\theta}_n)^2$.
- Erreur quad moyenne : $e_n(\theta) = E((\hat{\theta}_n - \theta)^2) = v_n(\theta) + b_n(\theta)^2$.
- Matrice de covariance : $E((\hat{\theta}_n - E(\hat{\theta}_n))(\hat{\theta}_n - E(\hat{\theta}_n))^T)$.
- CS de convergence : $b_n(\theta) \to 0$, $v_n(\theta) \to 0 \Rightarrow \hat{\theta}_n$ convergent.

**Vraisemblance :** $L(x_1, \dots, x_n / \theta) = \begin{cases} P(X_1 = x_1, \dots, X_n = x_n / \theta) \\ p(x_1, \dots, x_n / \theta) \end{cases}$.

**Borne de Cramer Rao**

$$BCR(\theta) = \frac{(1 + b_n'(\theta))}{-E\left(\frac{\partial^2 \ln L(x_1, \dots, x_n / \theta)}{\partial \theta^2}\right)}$$

- $BCR(\theta) \leq v_n(\theta)$.
- Estimateur efficace : $BCR(\theta) = v_n(\theta)$ et $b_n(\theta) = 0$.

Max de Vraisemblance $\hat{\theta}_{MV} = \arg\max_{\theta} L(x_1, \dots, x_n / \theta)$.

↳ - $\theta / \frac{\partial L}{\partial \theta} = 0$ ou $\frac{\partial \ln(L)}{\partial \theta} = 0$

- vérifier : $\left.\frac{\partial^2 \ln(L)}{\partial \theta^2}\right|_{\theta = \hat{\theta}_{MV}} \leq 0$.

Moments : $\hat{\theta}_{MO} = h(\hat{m}_1, \dots, \hat{m}_q)$ avec $\hat{m}_i = \frac{1}{n} \sum_{k=1}^{n} X_k^i$

Estimateur MMSE $\hat{\theta}_{MMSE} = E(\theta / X_1, \dots, X_n)$. ↳ Bayes

Max à postériori $\hat{\theta}_{MAP} = \arg\max_{\theta} P(\theta / X_1, \dots, X_n)$ ↳ Bayes + $\frac{\partial \ln(P\dots)}{\partial \theta}$

Intervalle de confiance $I = [a, b]$ / $P(\theta \in I) = \alpha$.

↳ - trouver un estimateur $\hat{\theta}$

- trouver un test stat $T(X_1, \dots, X_n)$ en Fct de $\hat{\theta}$ qui suit une loi connue.
- trouver les bornes avec les tables.

Risque stat : $H_0$ rejeté si $T(X_1, \dots, X_n) \in \Delta$.

- Risque de 1° espèce : $PFA = \alpha = P(\text{rejet } H_0 / H_0 \text{ vraie})$
- Risque de 2° espèce : $PND = \beta = P(\text{rejet } H_1 / H_1 \text{ vraie})$.
- Puissance du test : $\pi = 1 - \beta$.
- Région critique : $\{(x_1, \dots, x_n) / T(x_1, \dots, x_n) \in \Delta\}$.

P-valeur : Plus petite valeur de $\alpha$ pour laquelle $H_0$ est rejetée.

THM de Neymann - Pearson : Rejet de $H_0$ si $\frac{L(X_1, \dots, X_n / H_1)}{L(X_1, \dots, X_n / H_0)} > S_\alpha$.

$\alpha$ fixé.

Test GLR : Rejet de $H_0$ si $\frac{L(X_1, \dots, X_n / \hat{\theta}_1^{MV})}{L(X_1, \dots, X_n / \hat{\theta}_0^{MV})} > S_\alpha$

**Test du $\chi^2$**

$$\phi_n = \sum_{k=1}^{K} \frac{(Z_k - n p_k)^2}{n p_k} . \qquad \phi_n \xrightarrow{\mathcal{L}} \chi_{K-1} .$$

- Découper l'espace en K classes.
- $Z_k$ : nb d'observations / classes
- $p_k$ : proba d'appartenir à 1 classe : $\int_{C_k} p(x)dx$ suivant $H_0$.

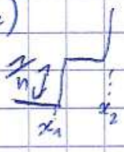
Rejet de $H_0$ si $\phi_n > S_\alpha$.

$\phi_n \xrightarrow{\mathcal{L}} \chi_{K-1-np}$ avec np : paramètres inconnus.

**Test de Kolmogorov**

$$D_n = \sup \left| \hat{F}(x) - F_0(x) \right|$$

- $\hat{F}(x)$ : fct de répartition théorique $P(X < x)$
- $F_0(x)$ : fct de répartition empirique :




$\cancel{K^2}$ $X \sim \chi^2$ : $X = \sum_{i=1}^{K} X_i^2$ $\quad X_i \sim N(0,1)$.

$$p_Y(y) = p_X(g^{-1}(y)) \left| \frac{dx}{dy} \right| .$$

$$\epsilon_i^+ = \left| \frac{i}{n} - F_0(x_i) \right| .$$