Statistique inférentielle et estimation de paramètre

- L'estimation

En stat descriptive le point de départ, ce sont des données $x_1, \ldots, x_n$

On en fait des graphiques, ou des indicateurs statistiques

Mais on ne se préoccupe pas du modèle probabiliste sous-jacent à ces données

En stat inférentielle, on va au contraire chercher à caractériser le modèle probabiliste associé

On va plus exactement chercher la loi de probabilité P telle que les données $x_1, \ldots, x_n$ sont des réalisations indépendantes d'une v.a de loi P

$$x_1, \ldots, x_n \to P$$

1° idée : on cherche P parmi toutes les lois de probabilités qui existent

➔ Pb : la famille est infinie : le pb est très difficile, le cadre non paramétrique

2° idée : on cherche P au sein d'une famille restreinte de la loi de proba , et cette famille est caractérisée par un paramètre

$$P \in \{ P_\theta, \theta \in \Theta \}$$

Ex :

- on peut se restreindre à la famille des lois exponentielles :

$$\{ P_\theta, \theta \in \Theta \} = \{ \mathcal{E}(\theta), \theta > 0 \}$$

- on peut se restreindre au cadre des lois normales

$$\{ P_\theta, \theta \in \Theta \} = \{ \mathcal{N}(n, \sigma^2), (n, \sigma^2) \epsilon \, \mathbb{R} * \, \mathbb{R}^2{}_x \}$$

➔ cadre paramétrique (beaucoup plus simple)

$$(x_1, \ldots, x_n) \to \theta$$

Remarque : en général, $\theta \in \mathbb{R}^p$

Déterminer θ revient à faire de l'estimation

2 manières :

- estimation partielle : à partir d'un jeu de données, on va fournir une valeur approchée de θ. Cette valeur sera notée θ
- estimation par intervalle de confiance : à partir d'un jeu de données, on fournit un intervalle dans lequel θ a α % de chance d'y être réellement

I. Estimation ponctuelle

Qu'est-ce qu'un estimateur et qu'est-ce qu'une estimation ?

- Un estimateur est une v.a, autrement dit une fonction
- Une estimation est la valeur de l'estimation sur un point

Ex : on souhaite connaitre l'opinion politique quand au président

➔ Un sondage

On note A l'ensemble des personnes favorables au président

$X(w) \begin{vmatrix} 0 \; si \; w \; \notin A \\ 1 \; si \; w \; \in A \end{vmatrix}$ avec w une personne

➔ La loi de X, notée P, appartient à la famille des lois de Bernouilli

$$P \; \in \{ \mathcal{B}(p), p \; \in [0, 1]\}$$

Pb : estimer le paramètre p de la loi

On dispose pour cela d'observation $x_1, \dots, x_n$ qui sont des réalisation des v.a $X_1, \dots, X_n$ où $X_1, \dots, X_n$ sont des v.a indépendantes et de même type que X

Remarque : P = P(X = 1) = P(A)

P peut être approché par $\underbrace{\frac{1}{n} * \sum_{i=1}^{n} x_i}_{une \; observation}$

➔ Un estimateur $\hat{p}$ de p est donnée par

$\hat{p} = \frac{1}{n} * \sum_{i=1}^{n} X$ -> estimateur

Un estimateur de θ est une fonction Tn définie sur E^n à calculer dans F qui à un échantillon (X_1, …, X_n) de la loi P_θ associe Tn(X_1, …, X_n) une v.a dont on peut connaître la loi

Remarque : E = X(Ω) F = Θ

Remarque : Si θ = E[X]

Un estimateur évident de θ, obtenu par la loi des grands nombres, est

$$\theta^n = \sum_{i=1}^{n} X_i = \overline{X_n}$$

Si θ = V[X]

Un estimateur évident est $\frac{1}{n} * \sum_{i=1}^{n} (X_i - \overline{X_n})^2$

Mais, on n'a pas toujours θ = E[X] ou $\theta = V[X]$

Comment faire alors ?

2 méthodes :

- Méthode des moments
- Méthode du maximum de vraisemblance

a) Méthode des « moments »

Rappel :

Soit k ∈ $\mathbb{N}^*$ et X une v.a

Le moment d'ordre k de X est $m_k(X) = E[X^k]$

Le moment centré d'ordre k de X est $\mu_k(X) = E[(X - E[X]^k]$

La version empirique de ces moments est respectivement

$\widehat{m}_k(X) = \frac{1}{n} * \sum_{i=1}^{n} X_i{}^k$ et $\hat{\mu}_k(X) = \frac{1}{n} * \sum_{i=1}^{n} (X_i - \overline{X_n})^k$

Sous l'hypothèse que $X \sim P_\theta$, alors $m_k(X)$ est une fonction de θ, de mpeme que $\mu_k(X)$

➔ $m_k(X) = m_k(\theta)$ et $\mu_k(X) = \mu_k(\theta)$

Principe de la méthode

On cherche le paramètre valeur de k, notre k, pour laquelle $m_k(\theta)$ (ou $\mu_k(\theta)$) dépend réellement de θ

On résout l'équation en θ

$m_k(\theta) = \frac{1}{n} * \sum_{i=1}^{n} X^{k'}$ ou $\mu_k(\theta) = \frac{1}{n} * \sum_{i=1}^{n} (X_i - \overline{X_n})^{k'}$

Remarque : généralisation

S'il existe 2 fonction f et g tel que

$g(\theta) = E[f(X)]$

$\hat{\theta}$ estimateur de θ est solution de $g(\hat{\theta}) = \frac{1}{n} * \sum_{i=1}^{n} f(X_i)$

Ex : $X_1, \dots, X_n$ des v.a indépendante et de même loi que X avec X$\sim\mathcal{E}(\theta)$ avec θ inconnue ($\theta > 0$)

On sait que E[X] = $\frac{1}{\theta}$

Par la méthode des moments, on sait qu'un estimateur $\hat{\theta}$ de θ est solution de

$$\frac{1}{\hat{\theta}} = \frac{1}{n} * \sum_{i=1}^{n} x_i \rightarrow \quad \hat{\theta} = \frac{n}{\sum_{i=1}^{n} X_n} = \frac{1}{\overline{X_n}}$$

Remarque : On sait que V[X] = $1/\theta^2$

Un estimateur de $\hat{\theta}_k$ de θ est donc solution de

$$\frac{1}{\hat{\theta}_2{}^2} = \frac{1}{n} * \sum_{i=1}^{n} (X_i - \overline{X_n})^2 \text{ -> } \hat{\theta}_2 = \sqrt{\frac{n}{\sum(X_i - \overline{X_n})^2}}$$

b) la méthode des maximums de vraisemblance

idée : la vraisemblance L($\theta, X_1, \dots, X_n$) est la probabilité d'observé, à θ fini, le n couplet ($X_1, \dots, X_n$)

Dans notre situation, on connait (X_1, …, X_n) et on cherche la valeur de θ la plus probable

La valeur de θ cherché est approché par la valeur de $\hat{\theta}$ qui maximise, à (X_1, …, X_n) fixé, la vraisemblance

Principe des maux de vraisemblance

Soit X une v.a de la loi P_θ

Soit (X_1, …, X_n) un n échantillon de même loi que X

Un estimateur $\hat{\theta}$, par la méthode des maux de vraissemblance, est tel que

$L(\hat{\theta}, X_1, \dots, X_n)$ = max L($\theta, X_1, \dots, X_n$)

Def : soit (X_1, …, X_n) un échantillon de loi P

$$L(\theta, X_1, \dots, X_n) = \left| \begin{array}{c} P(X_i = x_i) \; et \; P(X_n = x_n) \;\; si \; P \; est \; une \; loi \; droite \\ f(x1) x f(xn) si \; P \; est \; une \; loi \; continue \; (f \; fonction \; de \; droite \; associé) \end{array} \right.$$