FEUILLE TD2

Rappels

- $x_1, \dots, x_n \in \mho$ résultats de la répétition de n exp. indep.
  $\hookrightarrow$ réalisations indépendantes d'une $\mathfrak{m}$ loi $\mathcal{L}_{\theta^*}$ (θ^* inconnu), avec $\theta^* \in \Theta$
- Un échantillon est une suite $X_1, \dots, X_n$ de v.a.i.i.d. de loi $\mathcal{L}_\theta$, $\theta \in \Theta$
  $\hookrightarrow$ (alors $\exists \omega \in \Omega$ tq $(x_1, \dots, x_n) = (X_1(\omega), \dots, X_n(\omega))$, où $X_1, \dots, X_n$ iid de loi $\mathcal{L}_{\theta^*}$).
- Le modèle stat est décrit par tous les échantillons $X_1, \dots, X_n$ iid de loi $\mathcal{L}_\theta$, $\forall \theta \in \Theta$.
- Paramètre d'intérêt : $g(\theta)$ – Estimateur : $\hat{g}$ = fonction de $X_1, \dots, X_n$, où $X_1, \dots, X_n$ échantillon iid de loi $\mathcal{L}_\theta$, $\theta \in \Theta$ – Si "$\hat{g} \simeq g(\theta)$" $\forall \theta \in \Theta$ alors, comme $x_1, \dots, x_n$ issu de la loi $\mathcal{L}_{\theta^*}$:

$$\hat{g}(x_1, \dots, x_n) \simeq g(\theta^*)$$

Exercice 1

① On est dans le cadre de la répétition de n expériences indépendantes, donc un échantillon est une suite de v.a.i.i.d $X_1, \dots, X_n$, chaque X_i modélise le résultat de l'exp. i : X_i = nb de tirages effectués jqà obtenir une vis défectueuse.

Supposons que $\theta \in]0,1[$ $\overset{?!!}{=}$ est la proportion de vis défectueuses. On cherche la loi de X_i : X_i est à valeurs dans $\mathbb{N}^*$ et

$$\forall k \in \mathbb{N}^* \quad \mathbb{P}(X_i = k) = (1-\theta)^{k-1}\theta \implies X_i \sim \overset{\mathcal{L}_\theta}{\overset{\shortparallel}{\mathcal{G}(\theta)}}.$$

$$\boxed{\text{Alors, le modèle stat est décrit par tous les échantillons } X_1, \dots, X_n \text{ de v.a.i.i.d de loi } \mathcal{G}(\theta), \text{ et ceci } \forall \theta \in]0,1[.}$$

② Soit $X_1, \dots, X_n$ un échantillon i.i.d de la loi $\mathcal{G}(\theta)$, $\theta \in]0,1[$.

Par la LGN : $\overline{X_n} = \frac{1}{n}\sum_{i=1}^{n} X_i \approx \mathbb{E}(X_1)$ si n grand. Comme $X_1 \sim \mathcal{G}(\theta)$, on a $\mathbb{E}(X_1) = \frac{1}{\theta}$ (ex 3, TD1) d'où $\overline{X_n} \approx \frac{1}{\theta} \implies$ " $\frac{1}{\overline{X_n}} \approx \theta$ "

Comme θ représente la proportion de vis défectueuses, on a un estimateur "naturel" de θ est donné par $\frac{1}{\overline{X}_n}$

"AN" $x_1, \dots, x_n$ résultats des n exp. alors si $x_1, \dots, x_n$ réalisations indép. de $g(\theta^*)$,

on a $\frac{1}{\overline{x}_n} \underset{\sim}{\approx} \theta^*$ $\left(\overline{x}_n = \frac{1}{n}\sum_{i=1}^{n} x_i\right)$.

Exercice 3

① ⓐ Répétition de n exp. indép $\Rightarrow$ un échantillon est une suite de v.a. iid

$X_1, \dots, X_n$, chaque X_i modélisant le résultat de l'exp. i :

$$X_i = \begin{cases} 1 \text{ si un poisson rouge est tiré} \\ 0 \text{ sinon} \end{cases}$$

Si θ est la proportion de poissons rouges, $X_i \sim \mathcal{B}(\theta)$

Modèle stat = ens. des échantillons $X_1, \dots, X_n$ iid de loi $\mathcal{B}(\theta)$, $\forall \theta \in]0,1[$.

ⓑ Soit $\theta \in]0,1[$ et $X_1, \dots, X_n$ échantillon iid de loi $\mathcal{B}(\theta)$. Par la LGN :

$\overline{X}_n \approx \mathbb{E}(X_1)$ pour n grand, or $\mathbb{E}(X_1) = \theta$ car $X_1 \sim \mathcal{B}(\theta) \Rightarrow \overline{X}_n \approx \theta$ (n gd).

Comme θ représente la proportion de poissons rouge, un estimateur de θ est donc $\overline{X}_n$.

② → strictement le même exo. que l'exo 1

Exercice 2

① Répétition de n exp. indép., donc un échantillon est une $X_1, \ldots, X_n$ de var. i.i.d., chaque X_i modélisant le résultat de l'exp i :

$$X_i = \begin{cases} 1 \text{ si il y a un vainqueur} \\ 0 \text{ sinon} \end{cases}$$

Notons $\begin{cases} \alpha = \text{proba. que la pièce de A donne pile} \\ \beta = \quad " \quad B \quad " \end{cases}$ $\alpha, \beta \in]0,1[$.

Alors, $\mathbb{P}(X_i = 1) =$ proba("A donne P <u>et</u> B donne F" <u>ou</u> "A donne F <u>et</u> B donne P")

évts disjoints

$=$ proba(A donne P <u>et</u> B donne F) + proba(A donne F <u>et</u> B donne P)

Comme A et B jouent indépendamment, on trouve :

$$\mathbb{P}(X_i=1) = \text{proba}(A \text{ donne } P)\,\text{proba}(B \text{ donne } F) + \text{proba}(A \text{ donne } F)\,\text{proba}(B \text{ donne } P)$$
$$= \alpha(1-\beta)+(1-\alpha)\beta = \underbrace{\alpha+\beta-2\alpha\beta}.$$

↳ représente la proba qu'il y ait un vainqueur.

Alors $X_i \sim \mathcal{B}(\alpha+\beta-2\alpha\beta)$.

Modèle stat = ensemble de tous les échantillons iid $X_1,\dots,X_n$ de loi $\mathcal{B}(\alpha+\beta-2\alpha\beta)$, $\forall \alpha,\beta \in]0,1[$.

(2) Soient $\alpha,\beta\in]0,1[$ et $X_1,\dots,X_n$ un échantillon iid de loi $\mathcal{B}(\alpha+\beta-2\alpha\beta)$.

LGN : $\overline{X}_n \simeq E(X_1)=\alpha+\beta-2\alpha\beta$ car $X_1 \sim \mathcal{B}(\alpha+\beta+2\alpha\beta)$.

Comme $\alpha+\beta-2\alpha\beta$ représente la proba qu'un joueur soit déclaré vainqueur à l'issue d'une partie, on en déduit que $\overline{X}_n$ est un estimateur naturel de cette probabilité.

Exercice 4

① Répétition de n expériences aléatoires indépendantes $\Rightarrow$ un échantillon est donc une suite de va iid $X_1, \dots, X_n$, avec

$$X_i = \text{numéro de la boule tirée à la } i\text{-ème exp.}$$

Si K est nombre de boules présentes dans le sac, $K \in \mathbb{N}^*$, on a alors $X_i \in \{1, \dots, K\}$ et $\forall k \in \{1, \dots, K\}$:

$$\mathbb{P}(X_i = k) = \frac{1}{K} \quad \Rightarrow \quad X_i \sim \mathcal{U}(\{1, \dots, K\}).$$

$\boxed{\text{Donc le modèle stat est l'ensemble des échantillons } X_1, \dots, X_n \text{ iid de loi } \mathcal{U}(\{1, \dots, K\}),\ \forall K \in \mathbb{N}^*.}$

② Soit $K \in \mathbb{N}^*$ et $X_1, \dots, X_n$ iid de loi $\mathcal{U}(\{1, \dots, K\})$.

1er estimateur

On a $\overline{X}_n \approx E(X_1)$ pour n grand et $\mathbb{E}(X_1) = \sum_{k=1}^{K} k \underbrace{\mathbb{P}(X_1 = k)}_{= \frac{1}{K} \text{ si } k \in \{1, \dots, K\}} = \sum_{k=1}^{K} \frac{k}{K} = \frac{1}{K} \sum_{k=1}^{K} k$

(LGN)

$$= \frac{1}{K} \frac{K(K+1)}{2} = \frac{K+1}{2}$$

Donc $\overline{X}_n \approx \frac{K+1}{2}$ si n grand $\Rightarrow 2\overline{X}_n \approx K+1 \Rightarrow K \approx 2\overline{X}_n - 1$

Si on pose $\boxed{\hat{K}_1 = 2\overline{X}_n - 1}$, on a bien $\hat{K}_1 \approx K$ pour n grand, et $\hat{K}_1$ est un estimateur de K.

2^{e} estimateur

$$\boxed{\hat{K}_2 = \max_{i=1,\dots,n} X_i}$$: bcp plus astucieux !

L'idée est de dire que + n est grand, plus $\hat{K}_2$ se rapproche de K.