FEUILLE TD2

Rappels

- $x_1, \dots, x_n \in \Theta$ résultats de la répétition de n expériences aléatoires indépendantes
  ↪ réalisations indépendantes d'une même loi $\mathcal{L}_{\theta^*}$ (θ^* inconnu), avec $\theta^* \in \Theta$.

- Un échantillon est une suite de vaiid de loi $\mathcal{L}_\theta$, $\theta \in \Theta$.
  ↪ alors $\exists \omega \in \Omega$, tq $(x_1, \dots, x_n) = (X_1(\omega), \dots, X_n(\omega))$, où $X_1, \dots, X_n$ vaiid de loi $\mathcal{L}_{\theta^*}$

- Le modèle statistique est décrit par l'ensemble de tous les échantillons $X_1, \dots, X_n$ vaiid de lois $\mathcal{L}_\theta$, $\theta \in \Theta$

- Paramètre d'intérêt : $g(\theta)$

- <u>Estimateur</u> : $\hat{g}$ = fonction de $X_1, \dots, X_n$, où $X_1, \dots, X_n$ échantillon de la loi $\mathcal{L}_\theta$, $\theta \in \Theta$.

- Alors, si on montre que "$\hat{g} \approx g(\theta)$" $\forall \theta \in \Theta$ alors, comme $x_1, \dots, x_n$ réalisat° de la loi $\mathcal{L}_{\theta^*}$, on obtient

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

<u>Exercice 1</u>

① On est ici dans le cadre de la répétition de n expériences aléatoires indépendantes, donc un échantillon est une suite de var iid $X_1, \dots, X_n$, où chaque X_i modélise le résultat de l'expérience i :

Θ $\quad X_i$ = nb de tirages effectués jq'à obtenir une vis défectueuse

Si $\theta \in]0,1[$ est la proportion de vis défectueuses, alors $\mathcal{L}_\theta$

$$\mathbb{P}(X_i = k) = (1-\theta)^{k-1}\theta \,, \quad \forall k \in \mathbb{N}^* \Rightarrow X_i \sim \mathcal{G}(\theta)$$

Donc le modèle stat est décrit par tous les échantillons $X_1, \dots, X_n$ de vaiid de loi $\mathcal{G}(\theta)$, $\theta \in]0,1[$

(2) Soit $X_1, \dots, X_n$ un échantillon de vaiid de 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 gd. Comme $X_1 \sim \mathcal{G}(\theta)$, on sait (ex. 3, TD1) que $\mathbb{E}(X_1) = \frac{1}{\theta}$. Ainsi :

$$\overline{X}_n \approx \frac{1}{\theta}, \text{ n gd.} \Rightarrow \theta \approx \frac{1}{\overline{X}_n}, \text{ n gd.}$$

Comme θ représente la proportion de vis défectueuses, θ est le paramètre d'intérêt, et un estimateur de θ est donc donné par $\frac{1}{\overline{X}_n}$

Exercice 3

1) a) On est ici dans le cadre de la répétition de n expériences aléatoires indépendantes, donc un échantillon est une suite de var iid $X_1, \dots, X_n$, où chaque X_i modélise le résultat de l'expérience i :

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

Si la proportion de poissons rouges est $\theta \in]0,1[$ alors

$$\mathbb{P}(X_i = 1) = \theta \implies X_i \sim \mathcal{B}(\theta)$$

Donc le modèle stat est décrit par l'ensemble de tous les échantillons $X_1, \dots, X_n$ de var iid de même loi $\mathcal{B}(\theta)$, $\theta \in]0,1[$

b) Soit $X_1, \dots, X_n$ un échantillon de var iid de loi $\mathcal{B}(\theta)$, $\theta \in]0,1[$

Par la LGN : $\overline{X}_n \simeq \mathbb{E}(X_1) = \theta$, si n gd, et car $X_1 \sim \mathcal{B}(\theta)$

Donc un estimateur de θ est donné par $\overline{X}_n$.

② Idem que pour ex. 1.

Exercice 2

① Répétition de n expériences aléatoires indépendantes et identiques, donc un échantillon est une suite $X_1, \dots, X_n$ de va i.i.d.

De plus, X_i modélise le résultat de l'exp i :

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

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

Alors, $\mathbb{P}(X_i = 1) = \text{proba}\Big[(\text{A donne pile } \underset{\cap}{\underline{\text{et}}} \text{ B donne face}) \underset{\cup}{\underline{\text{ou}}} (\text{A donne face } \underset{\cap}{\underline{\text{et}}} \text{ B donne pile})\Big]$

car la réunion est disjointe $= \text{proba}(\text{A donne pile } \underline{\text{et}} \text{ B donne face}) + \text{proba}(\text{A donne face } \underline{\text{et}} \text{ B donne pile})$

$= \text{proba}(A \text{ donne pile}) \times \text{proba}(B \text{ donne face}) + \text{proba}(A \text{ donne face}) \times \text{proba}(B \text{ donne pile})$

par indépendance

$= \alpha(1-\beta) + (1-\alpha)\beta = \alpha + \beta - 2\alpha\beta$

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

Ainsi, le modèle stat est décrit par tous les échantillons $X_1, \dots, X_n$ de vaiid de loi $\mathcal{B}(\alpha + \beta - 2\alpha\beta)$, $\alpha, \beta \in]0,1[$.

(2) Soit donc $X_1, \dots, X_n$ un échantillon de vaiid de loi $\mathcal{B}(\alpha+\beta-2\alpha\beta)$.

Par la LGN : $\bar{X}_n \approx \mathbb{E}(X_1) = \underbrace{\alpha + \beta - 2\alpha\beta}_{\text{proba qu'un joueur soit déclaré vainqueur.}}$, pour n gd.

Donc $\bar{X}_n$ est un estimateur de la proba qu'un joueur soit déclaré vainqueur.

Exercice 4

① Répétition de n expériences aléatoires indépendantes et identiques, donc un échantillon est une suite $X_1, \dots, X_n$ de vaiid.

Si K est le nombre de boules présentes dans le sac ($K \in \mathbb{N}^*$), on a alors

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

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

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

Paramètre d'intérêt : K

Estimateur 1

Par la LGN : $\bar{X}_n \approx \mathbb{E}(X_1)$, si n est grand.

$$\mathbb{E}(X_1) = \sum_{k=1}^{K} k\,\mathbb{P}(X_1 = k) \quad (\text{car } X_1 \text{ est à valeurs dans } \{1, \dots, K\})$$

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

Donc $\overline{X}_n \approx \frac{K+1}{2}$ si n gd. $\Rightarrow K \approx 2\overline{X}_n - 1$. Par suite, $2\overline{X}_n - 1$ est un estimateur de K.

Estimateur 2

Un autre estimateur, moins mauvais que le précédent : $\max_{i=1,\dots,n} X_i$