FEUILLE TD2

Rappels

- Trouver le modèle stat, c'est trouver tous les échantillons
- Dans le cadre de la répétition de n expériences indépendantes, un échantillon est une suite de vaiid $X_1, \ldots, X_n$ de même loi $\mathcal{L}_\theta$.
- On a tous les échantillons si on a les échantillons pour toutes les valeurs possibles de θ : il faut donc trouver l'ensemble des valeurs possibles pour θ, noté Θ.

$\longrightarrow$ Au final, il faut trouver $\mathcal{L}_\theta$ et Θ !

Exercice 1

1. Comme les n expériences sont indépendantes, un échantillon est une suite $X_1, \ldots, X_n$ de vaiid, chaque X_i modélise le résultat de l'expérience aléatoire i.

Supposons que $\boxed{\theta \in]0,1[\text{ soit la proportion de vis défectueuses}}$. A l'exp. i, X_i modélise le nb de tirages effectués jusqu'à tirer une vis défectueuse. Donc X_i prend ses valeurs dans $\mathbb{N}^*$. Alors $\forall k \in \mathbb{N}^*$:

$$\mathbb{P}(X_i = k) = (1-\theta)^{k-1}\theta \quad \text{i.e. } X_i \sim \mathcal{G}(\theta)$$

Ainsi, le modèle statistique est décrit par les échantillons $X_1, \dots, X_n$ iid de loi $\underset{\mathcal{L}_\theta}{\underset{\|}{\mathcal{G}(\theta)}}$, avec $\theta \in \underset{\mathcal{H}}{\underset{\|}{]0,1[}}$

Rappel LGN

Si $Z_1, \dots, Z_n$ iid alors

$$\lim_{n\to+\infty} \mathbb{E}\left((\bar{Z}_n - \mathbb{E}(Z_1))^2\right) = 0$$

avec $\bar{Z}_n = \frac{1}{n}\sum_{i=1}^{n} Z_i$

donc

"$\bar{Z}_n \approx \mathbb{E}(Z_1)$" si n grand.

(2) Soit $\theta \in]0,1[$ et $X_1, \dots, X_n$ iid de loi $\mathcal{G}(\theta)$ (un échantillon). Par la LGN, on sait que

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \approx \mathbb{E}(X_1)$$

On a vu à l'ex. 3 feuille 1 que $\mathbb{E}(X_1) = \frac{1}{\theta}$ car $X_1 \sim \mathcal{G}(\theta)$.

Donc $\boxed{\frac{1}{\bar{X}_n} \text{ est un estimateur de } \theta.}$

Exercice 2

① On est dans le cadre de la répétition de n expériences indépendantes donc un échantillon est une suite $X_1, \dots, X_n$ de va iid.

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

On cherche maintenant la loi de X_i, qui modélise le résultat de la i-ème expérience. On a

$$X_i = \begin{cases} 1 & \text{si il y a un vainqueur à la i-ème exp.} \\ 0 & \text{sinon.} \end{cases}$$

Or, $\mathbb{P}(X_i = 1)$ = proba (A donne pile et B donne face) ou (A donne face et B donne pile)

$$= \alpha(1-\beta) + (1-\alpha)\beta \longrightarrow \boxed{X_i \sim \mathcal{B}(\alpha(1-\beta)+(1-\alpha)\beta)}$$

Le modèle statistique est donc décrit par les échantillons $X_1, \dots, X_n$ iid de loi $\mathcal{B}(\alpha(1-\beta)+(1-\alpha)\beta)$, pour toutes les valeurs $\alpha, \beta \in]0,1[$.

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

Par la LGN, on sait que $\overline{X}_n \overset{p.s.}{\to} \mathbb{E}(X_1) = \underbrace{\alpha(1-\beta)+\beta(1-\alpha)}$

= proba qu'un joueur soit déclaré vainqueur à l'issue d'un lancer.

Donc $\overline{X}_n$ est un estimateur de la proba qu'un joueur soit déclaré vainqueur à l'issue d'un lancer.

Exercice 4

① Il s'agit de n répétitions indépendantes de la même expérience, donc un n-échantillon est une suite $X_1, \dots, X_n$ de v.a.i.i.d.

Supposons que le nombre de boules est $K \in \mathbb{N}^*$. Dans l'exp., X_i modélise le résultat de l'expérience i :

$$X_i = n^o \text{ de la boule tirée à l'exp. } i$$

Donc X_i prend ses valeurs dans $\{1, 2, \dots, K\}$ et $\forall x \in \{1, 2, \dots, K\}$:

$$\mathbb{P}(X_i = x) = \frac{1}{K} \quad \text{ie} \quad \boxed{X_i \sim \mathcal{U}(\{1, 2, \dots, K\})}.$$

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

$\overset{\shortparallel}{\mathcal{L}_\theta}$ $\qquad$ $\overset{\shortparallel}{(H)}$

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

D'après la LGN, $\overline{X}_n \simeq \mathbb{E}(X_1)$ or

$$\mathbb{E}(X_1) = \sum_{x=1}^{K} x \, \mathbb{P}(X_1 = x)$$ car X_1 prend ses valeurs dans $\{1, 2, \dots, K\}$

$$= \sum_{x=1}^{K} x \times \frac{1}{K}$$ car $X_1 \sim \mathcal{U}(\{1, 2, \dots, K\})$

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

Donc $\overline{X}_n \simeq \frac{K+1}{2}$: un estimateur de K est donc

$$\boxed{\hat{K} = 2\overline{X}_n - 1}$$

Autre estimateur

$$\widehat{K'_n} = \max_{i=1,\dots,n} X_i$$: plus n augmente, + $\widehat{K'_n}$ doit s'approcher de K !

$(\widehat{K'_n})_n$ est croissante et bornée par K donc elle converge. Notons L la limite (L est une va!). Est-il possible que $\mathbb{P}(L=K)<1$?