Devoir maison

Patrick Tardivel
Université de Bourgogne

Exercice 1. *Soient $X_1,\dots,X_n$ une suite de variables aléatoires de loi iid de loi de Bernoulli $\mathcal{B}(p)$, où $p\in]0,1[$ est un paramètre connu. On pose $\overline{X}_n=(\sum_{i=1}^n X_i)/n$ la moyenne empirique. Justifier qu'asymptotiquement et indépendamment de $p\in]0,1[$, la probabilité de l'événement $\overline{X}_n\in[p-1/\sqrt{n},p+1\sqrt{n}]$ est au moins égale à 0.95*

Solution : D'après le théorème de Moivre-Laplace, on a

$$\sqrt{n}\frac{\overline{X}_n-p}{\sqrt{p(1-p)}}\xrightarrow[n\to+\infty]{\mathcal{L}}\mathcal{N}(0,1).$$

Lorsque $Z\sim\mathcal{N}(0,1)$, on sait que $\mathbb{P}(-1,96\leq Z\leq 1,96)=0,95$. En approchant $1,96$ par 2 alors la probabilité que soit Z soit dans l'intervalle $[-2,2]$ est légèrement supérieure à $0,95$. Plus précisément $\mathbb{P}(-2\leq Z\leq 2)=0,954$. Le théorème de Moivre-Laplace nous donne le résultat suivant :

$$\lim_{n\to+\infty}\mathbb{P}\left(-2\leq\sqrt{n}\frac{\overline{X}_n-p}{\sqrt{p(1-p)}}\leq 2\right)=\lim_{n\to+\infty}\mathbb{P}\left(p-\frac{2\sqrt{p(1-p)}}{\sqrt{n}}\leq\overline{X}_n\leq p+\frac{2\sqrt{p(1-p)}}{\sqrt{n}}\right)=\mathbb{P}(-2\leq Z\leq 2).$$

Par ailleurs, pour tout $p\in]0,1[$ on a $2\sqrt{p(1-p)}\leq 2\sqrt{0,5(1-0,5)}=1$. Ainsi, pour tout $n\geq 1$ l'inégalité suivante est satisfaite

$$\mathbb{P}\left(p-\frac{2\sqrt{p(1-p)}}{\sqrt{n}}\leq\overline{X}_n\leq p+\frac{2\sqrt{p(1-p)}}{\sqrt{n}}\right)\leq\mathbb{P}\left(p-\frac{1}{\sqrt{n}}\leq\overline{X}_n\leq p+\frac{1}{\sqrt{n}}\right)$$

Comme le membre de gauche de l'inégalité ci-dessus converge vers $0,954$ alors, il existe $n_0\in\mathbb{N}$ tel que pour tout $n\geq n_0$ on a

$$0,95\leq\mathbb{P}\left(p-\frac{2\sqrt{p(1-p)}}{\sqrt{n}}\leq\overline{X}_n\leq p+\frac{2\sqrt{p(1-p)}}{\sqrt{n}}\right)\leq\mathbb{P}\left(p-\frac{1}{\sqrt{n}}\leq\overline{X}_n\leq p+\frac{1}{\sqrt{n}}\right).$$

Ainsi, asymptotiquement et indépendamment de $p\in]0,1[$, la probabilité de l'événement $\overline{X}_n\in[p-1/\sqrt{n},p+1\sqrt{n}]$ est au moins égale à 0.95.

Exercice 2. *Le nombre de décès en France dus au Covid-19 durant les quatre dernières semaines est donné par le tableau suivant :*

Semaines	1 (10/01 – 16/01)	2 (17/01 – 23/01)	3 (24/01 – 30/01)	4 (31/02 – 06/02)
Morts	1484	1617	1890	1848

1. *Représenter graphiquement le nombre de morts en fonction du temps.*
2. **(Modèle 1 : poids uniforme)** *Calculer la droite des moindres carrés du nombre de décès en fonction des semaines lorsque les poids sont uniformes.*
3. **(Modèle 2 : pondéré)** *Calculer la droite des moindres carrés du nombre de décès en fonction des semaines lorsque les poids sont respectivement* $0,1$ *(semaine 1),* $0,15$ *(semaine 2),* $0,25$ *(semaine 3) et* $0,5$ *(semaine 4).*
4. *Pour les deux modèles et pour chacune des semaines suivantes, prévoir le nombre total de décès :*
   - *Semaine 5 : du lundi 7 février au dimanche 13 février.*
   - *Semaine 6 : du lundi 14 février au dimanche 20 février.*
   - *Semaine 7 : du lundi 21 février au dimanche 27 février.*

Solution (corrigé très synthétique) :
La droite des moindres carrés a pour équation $y = ax + b$ avec

$$a = \frac{cov(X,Y)}{var(X)} \text{ et } b = \overline{Y} - a\overline{X}.$$

2) **Modèle 1 :** On a $\overline{X} = 2,5$, $var(X) = 1,25$, $\overline{Y} = 1709,75$ et $cov(X,Y) = 170,625$ d'où $a = 136,5$ et $b = 1368,5$.

3) **Modèle 2 :** On a $\overline{X} = 3,15$, $var(X) = 1,0275$, $\overline{Y} = 1787,45$ et $cov(X,Y) = 116,5325$ d'où $a = 113,4136$ et $b = 1430,197$.

4) Le nombre de décès prévu dus au Covid-19 la semaine 5 vaut : $1368,5 + 5 \times 136,5 = 2051$ (modèle 1) ou $1430,197 + 5 \times 113,4136 \approx 1997$ (modèle 2).
Le nombre de décès prévu dus au Covid-19 la semaine 6 vaut : $1368,5 + 6 \times 136,5 \approx 2187$ (modèle 1) ou $1430,197 + 6 \times 113,4136 \approx 2111$ (modèle 2).
Le nombre de décès prévu dus au Covid-19 la semaine 7 vaut : $1368,5 + 7 \times 136,5 = 2324$ (modèle 1) ou $1430,197 + 7 \times 113,4136 \approx 2224$ (modèle 2).