Exercice 2 (exercice important)

$X_1, \dots, X_n$ v.a. i.i.d. à valeurs dans A.

$D \subset A$ tq $p = \mathbb{P}(X_1 \in D) \neq 0$.

$n \geq 1$ $S_n = \sum_{j=1}^{n} \mathbb{1}_{\{X_j \in D\}}$.

1) • Remarquons $Y_j = \mathbb{1}_{\{X_j \in D\}} = \begin{cases} 1 & \text{si } X_j \in D \\ 0 & \text{sinon} \end{cases}$

est une variable de Bernoulli de paramètre $\mathbb{P}(X_j \in D) = p$.

• On remarque cela car les 2 principaux théorèmes (loi des grands nombres et TCL) font intervenir $\sum_{j=1}^{n} Y_j$.

$$\mathbb{E}[S_n] = \mathbb{E}\left[\sum_{j=1}^{n} \mathbb{1}_{[X_j \in D]}\right] = \mathbb{E}\left[\sum_{j=1}^{n} Y_j\right]$$

$$= \sum_{j=1}^{n} \mathbb{E}[Y_j]$$

Par linéarité de l'espérance

$$\mathrm{Var}(S_n) = \mathrm{Var}\left(\sum_{j=1}^{n} Y_j\right)$$

Y_j indépendantes

$$= \sum_{j=1}^{n} \mathrm{Var}(Y_j)$$

$\mathrm{Var}(Y_j) = \mathrm{Var}(Y_1)$ $\forall j$ car les variables Y_j sont i.d. (identiquement distribuées)

$$= \sum_{j=1}^{n} \mathrm{Var}(Y_1)$$

$$= \sum_{j=1}^{n} p(1-p)$$

$$= n p(1-p)$$

⚠ Pas besoin de mettre i.i.d., c'est que l'indépendance que l'on utilise !

⚠ La variance n'est pas linéaire !

2) On applique la loi des grands nombres (LFGN)

On a : • Y_i i.i.d.
• $\mathbb{E}[Y_1] = p$

A METTRE SUR VOTRE COPIE !

(sinon pas de points ☹)

Par la LFGN, on a donc

$$\frac{S_n}{n} = \frac{\sum_{j=1}^{n} Y_j}{n} \xrightarrow[n \to +\infty]{p.s.} \mathbb{E}[Y_1] = p$$

Fait en cours :

3) EQM (erreur quad. moy.)

majoration "uniforme" → on veut majorer par une quantité dépendant de n (mais pas de p)

$$\mathbb{E}\left[\left(\frac{S_n}{n}-p\right)^2\right] = \mathbb{E}\left[\left(\frac{S_n}{n} - \underbrace{\mathbb{E}\left(\frac{S_n}{n}\right)}\right)^2\right]$$

en effet on a : $\mathbb{E}[S_n] = np$

donc par linéarité de l'épérance

$$\mathbb{E}\left[\frac{S_n}{n}\right] = \frac{1}{n}\mathbb{E}[S_n] = p$$

$$= \mathrm{Var}\left(\frac{S_n}{n}\right)$$

$$= \frac{1}{n^2}\mathrm{Var}(S_n)$$

$$= \frac{1}{n^2}\, np(1-p)$$

$$= \frac{1}{n}\, p(1-p)$$

CLASSIQUE !

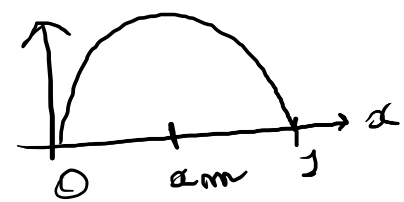
$p \in (0,1)$ donc $p(1-p) \leq \frac{1}{4}$

Il suffit d'étudier la fonction $f : x \mapsto x(1-x)$

On cherche un max : dérivée $f'(x) = 1-2x$

$f'(x_m) = 0 \Rightarrow x_m = \frac{1}{2}$.

$\cdot\ f(x_m) = \frac{1}{4}$



On a donc

$$\mathbb{E}\left[\left(\frac{S_n}{n}-p\right)^2\right] \leqslant \frac{1}{4n}.$$

on utilise 3)

4) inégalité de Tchebychev. Soit $\varepsilon > 0$

$$\mathbb{P}\left(\left|\frac{S_n}{n}-p\right| \geqslant \varepsilon\right) \leqslant \frac{\mathbb{E}\left[\left(\frac{S_n}{n}-p\right)^2\right]}{\varepsilon^2} \leqslant \frac{1}{4n\varepsilon^2}$$

Quand on a $\frac{1}{\varepsilon^a}$, $a \geqslant 1$, à droite et une $\mathbb{P}$ type $\mathbb{P}(- \geqslant \varepsilon)$ à gauche

ON PENSE TCHEBYCHEV !

5) Théorème central limite (TCL)

On a : • Y_i i.i.d
• $Var(Y_1) = p(1-p)$

À ECRIRE SUR VOTRE COPIE !

Par le TCL, on a :

$$\sqrt{n}\left(\frac{S_n}{n} - \underbrace{p}_{= \mathbb{E}[Y_1]}\right) \xrightarrow[n \to +\infty]{\mathcal{L}} \mathcal{N}(0, \underbrace{p(1-p)}_{= Var(Y_1)})$$