MATHEMATIQUES
Loi des grands nombres : entraînement savoir-faire (corrigé)

Exercice 1

1. **a.** La variable aléatoire S_n est la somme de n variables aléatoires de Bernoulli indépendantes et de paramètre p.
Donc la variable aléatoire S_n suit la loi binomiale de paramètres n et p.

b. Ainsi on a
- $E(S_n) = np$
- $V(S_n) = np(1-p)$
- $\sigma(S_n) = \sqrt{np(1-p)}$

2. **a.** La fonction f est définie, continue et dérivable sur l'intervalle $[0;1]$.
Soit $x \in [0;1]$.
On a $f(x) = x(1-x) = x - x^2$.
Donc $f'(x) = 1 - 2x$.
$f'(x) > 0$
$1 - 2x > 0$
$1 > 2x$
$\frac{1}{2} > x$
Ainsi la fonction f est croissante sur l'intervalle $\left[0;\frac{1}{2}\right]$ et décroissante sur l'intervalle $\left[\frac{1}{2};0\right]$.

b. Le maximum de la fonction f sur l'intervalle $[0;1]$ est donc atteint en $x = \frac{1}{2}$.
Or $f\left(\frac{1}{2}\right) = \frac{1}{2} - \left(\frac{1}{2}\right)^2 = \frac{1}{2} - \frac{1}{4} = \frac{1}{4}$.
Donc, pour tout $x \in [0;1]$, on a effectivement $f(x) \leqslant \frac{1}{4}$.

3. On applique l'inégalité de Bienaymé-Tchebichev à la variable aléatoire $\frac{S_n}{n}$.

Ainsi on obtient $p\left(\left|\frac{S_n}{n} - E\left(\frac{S_n}{n}\right)\right| \geqslant a\right) \leqslant \frac{V\left(\frac{S_n}{n}\right)}{a^2}$.

Or $E\left(\frac{S_n}{n}\right) = \frac{1}{n}E(S_n) = \frac{1}{n} \times np = p$.

De plus $V\left(\frac{S_n}{n}\right) = \frac{1}{n^2}V(S_n) = \frac{1}{n^2} \times np(1-p) = \frac{p(1-p)}{n}$.

Ainsi, l'inégalité devient $p\left(\left|\frac{S_n}{n} - p\right| \geqslant a\right) \leqslant \frac{p(1-p)}{na^2}$.

Avec la question précédente, on a démontré que pour tout $p \in [0;1]$, on a $p(1-p) \leqslant \frac{1}{4}$.

Cela permet de conclure que $p\left(\left|\frac{S_n}{n} - p\right| \geqslant a\right) \leqslant \frac{p(1-p)}{na^2} \leqslant \frac{1}{4na^2}$.

4. On cherche les entiers naturels n tels que $p\left(\left|\frac{S_n}{n} - p\right| < 10^{-2}\right) \geqslant 0,95$.

On passe à l'évènement contraire pour obtenir $1 - p\left(\left|\frac{S_n}{n} - p\right| \geqslant 10^{-2}\right) \geqslant 0,95$.

Donc on obtient $-p\left(\left|\frac{S_n}{n} - p\right| \geqslant 10^{-2}\right) \geqslant -0,05$.

Ainsi on a $p\left(\left|\dfrac{S_n}{n}-p\right|\geqslant 10^{-2}\right)\leqslant 0,05$.

Or, avec la question précédente, on a démontré $p\left(\left|\dfrac{S_n}{n}-p\right|\geqslant a\right)\leqslant \dfrac{1}{4na^2}$ pour $a>0$.

On utilise cette inégalité avec $a=10^{-2}$.

D'où $p\left(\left|\dfrac{S_n}{n}-p\right|\geqslant 10^{-2}\right)\leqslant \dfrac{1}{4n\times 10^{-4}}$.

Il suffit de choisir n tel que $\dfrac{1}{4n\times 10^{-4}}\leqslant 0,05$.

D'où $\dfrac{1}{n}\leqslant 0,05\times 4\times 10^{-4}$.

Ainsi $\dfrac{1}{n}\leqslant 0,2\times 10^{-4}$.

Donc, par passage à l'inverse $n\geqslant 50\ 000$.
Il faut donc utiliser un échantillon de taille supérieure à 50 000 pour obtenir la probabilité p à 10^{-2} près avec une probabilité de 0,95.
Cela commence à faire beaucoup...