Spé Maths
Terminale

Loi Grands Nombres

CORRIGÉ DE L'EXERCICE

Loi des grands nombres

16

Correction

1. Déterminons qui a raison :

Lou choisit un univers de probabilité à 21 éléments et émet l'hypothèse d'équiprobabilité. Or, même « indiscernables », il y a 2 dés distincts, disons D1 et D2. Pour avoir un même numéro, il faut que D2 montre exactement le même numéro que D1, alors que pour avoir deux numéros distincts donnés, on peut intervertir les rôles des deux dés. Une issue où *x* est différent de *y* a une probabilité double d'une issue ou *x* et *y* sont égaux.

Parmi les 21 éléments de l'univers de Lou, soit *p* la probabilité de chacun des 6 éléments de type $\{x, x\}$ et soit $2p$ la probabilité des 15 autres $\{x, y\}$ où *x* et *y* sont différents.

Calculons *p* :

La somme des probabilités de toutes les issues est égale à 1 : $6p + 15 \times (2p) = 36p = 1$. Nous obtenons : $p = \frac{1}{36}$. En conséquence, la probabilité d'avoir deux numéros identiques est bien le nombre $6 \times \frac{1}{36} = \frac{1}{6}$ et non le nombre $\frac{6}{21}$.

C'est la modélisation de Pablo qui est correcte et non celle de Lou.

2 et 3. Le protocole imaginé conduit à répéter indépendamment *n* fois une même épreuve de Bernoulli où l'on observe la réalisation de *A* ou celle de son contraire. La variable aléatoire S_n est somme de *n* lois de Bernoulli indépendantes et suivra de ce fait une loi binomiale $B(n, p)$ avec une valeur de *p* différente suivant les points de vue des deux personnages.

2.1 . Précisons, la nature et les paramètres de la loi de probabilité suivie par S_n sous l'hypothèse de Lou (en supposant donc que $p = \frac{2}{7}$) :

La variable aléatoire S_n suivrait dans ce cas la loi binomiale $B\left(n, \frac{2}{7}\right)$. Elle aurait donc pour espérance $n \times \frac{2}{7} = \frac{2n}{7}$ et pour variance $n \times \frac{2}{7} \times \frac{5}{7} = \frac{10n}{49}$.

$F_n = \frac{1}{n} \times S_n$ a pour espérance $\frac{1}{n} \times \frac{2n}{7} = \frac{2}{7}$ et pour variance : $\frac{1}{n^2} \times \frac{10n}{49} = \frac{10}{49n}$.

2.2. Déterminons des entiers *n* pour lesquels on a $P\left(\left|F_n - \frac{2}{7}\right| \geq 0{,}05\right) \leq 0{,}01$:

Dans le cas où la variable aléatoire de référence suit la loi de Bernoulli $\mathcal{B}(p)$, l'inégalité de concentration nous dit que pour tout réel $\delta > 0$:

$$P(|F_n - p| \geq \delta) \leq \frac{p(1-p)}{n\,\delta^2}$$

Appliquons cette inégalité avec $p = \frac{2}{7}$; $\delta = 0{,}05$. Nous obtenons :

$$P\left(\left|F_n - \frac{2}{7}\right| \geq 0{,}05\right) \leq \frac{10}{49n \times\ 0{,}05^2} = \frac{4000}{49n}$$

Si l'on veut que cette probabilité soit $\leq 0{,}01$, il s'agit de choisir *n* se sorte que $\frac{4000}{49n} \leq 0.01$, c'est-à-dire tel que $n \geq \frac{400000}{49}$ (soit $n \geq 8164$). Nous pouvons proposer $n_0 = 8164$.

3.1. Précisons, la nature et les paramètres de la loi de probabilité suivie par S_n sous l'hypothèse de Pablo (en supposant maintenant que $p = \frac{1}{6}$) :

La variable aléatoire S_n suivrait dans ce cas la loi binomiale $B\left(n,\ \frac{1}{6}\right)$. Elle aurait donc pour espérance $n \times \frac{1}{6} = \frac{n}{6}$ et pour variance $n \times \frac{1}{6} \times \frac{5}{6} = \frac{5n}{36}$. Il en résulte que $F_n = \frac{1}{n} \times S_n$ a pour espérance $\frac{1}{n} \times \frac{n}{6} = \frac{1}{6}$ et pour variance : $\frac{1}{n^2} \times \frac{5n}{36} = \frac{5}{36n}$.

3.2. Déterminons des entiers *n* pour lesquels on a $P\left(\left|F_n - \frac{1}{6}\right| \geq 0{,}05\right) \leq 0{,}01$:

Appliquons la même inégalité que dans **2.2** avec $p = \frac{1}{6}$; $\delta = 0{,}05$. Nous obtenons :

$$P\left(\left|F_n - \frac{1}{6}\right| \geq 0{,}05\right) \leq \frac{5}{36n \times\ 0{,}05^2} = \frac{2000}{36n} = \frac{500}{9n}$$

Si l'on veut que la probabilité soit plus petite que 0,01, il s'agit de choisir *n* se sorte que $\frac{500}{9n} \leq 0.01$, c'est-à-dire tel que $n \geq \frac{50000}{9}$ (soit $n \geq 5556$). Nous pouvons proposer $n_0 = 5556$.

4.1. Expliquons le choix de l'entier *n* dans la la simulation :

Les 2 personnages ont choisi $n = 8200$ (ce qui est visible dans « rand(1, 8201) ») soit une valeur un peu plus grande que 8164. Elle est *a fortiori* plus grande que 5556.

Nous pouvons appliquer les résultats des questions 2 et 3 et cela quelle que soit l'hypothèse retenue, aussi bien celle de Lou que celle de Pablo.

Si ces deux personnages veulent « se départager », il faut qu'ils puissent localiser la fréquence observée dans l'un ou l'autre de deux intervalles disjoints, l'un centré en $\frac{2}{7}$ (celui de Lou) ou l'autre centré en $\frac{1}{6}$ (celui de Pablo). Le choix $\delta = 0,05$ permet cette discrimination.

4.2. Les instructions x=rd.randint(1,6) et y=rd.randint(1,6) simulent le résultat des jets simultanés de deux dés censés être équilibrés. Elles renvoient deux entiers aléatoires entre 1 et 6, entiers dont l'algorithme compare ensuite les valeurs.

4.3. L'algorithme renvoie la fréquence des occurrences de *A*, calculée sur 8200 lancers simulés.

4.4. Sachant que $0,285 < \frac{2}{7} < 0,286$, l'intervalle $\left]\frac{2}{7} - 0,05\,;\,\frac{2}{7} + 0,05\right[$ est inclus dans l'intervalle $[0,235\,;0,336]$.

Les quatre fréquences observées sont toutes, nettement, à l'extérieur de cet intervalle.

Au seuil de confiance 99 %, nous pouvons affirmer qu'aucune des fréquences observées n'est conforme à l'étude théorique de Lou. Son hypothèse doit donc être rejetée.

Avec une probabilité au moins égale à 99 %, nous affirmons que l'analyse de Lou est *fausse*.

En revanche, on note que l'intervalle $\left]\frac{1}{6} - 0,05\,;\,\frac{1}{6} + 0,05\right[$ contient l'intervalle $[0,117\,;0,216]$ et que chacune des quatre fréquences observées est bel et bien dans cet intervalle.

Nous pouvons conclure que l'étude théorique de Pablo est *plausible*.

NB. Alors que nous pouvons affirmer avec un risque minime que l'hypothèse de Lou est manifestement *fausse*, nous ne pouvons pas *démontrer* pour autant que l'hypothèse de Pablo est « *exacte* ». Nous pouvons seulement dire qu'elle est *plausible* c'est-à-dire qu'il n'y a aucun argument qui la réfute. Les résultats expérimentaux de la simulation concordent bien avec son étude théorique.