Exercice 4 3 points

Dans cet exercice, on considère le premier chiffre des entiers naturels non nuls, en écriture décimale.
Par exemple, le premier chiffre de 2017 est 2 et le premier chiffre de 95 est 9.
Dans certaines circonstances, le premier chiffre d'un nombre aléatoire non nul peut modélisé par une variable aléatoire X telle que pour tout entier c compris 1 et 9, $P(X=c)=\frac{\ln(c+1)-\ln(c)}{\ln(10)}$.
Cette loi est appelée loi de Benford.

1. Que vaut $P(X=1)$?

2. On souhaite examiner si la loi de Benford est un modèle valide dans deux cas particuliers.

2.a. Premier cas

Un fichier statistique de l'INSEE indique la population des communes en France au 1^{er} janvier 2016 (champ : France métropolitaine et départements d'outre-mer de la Guadeloupe, de la Guyane, de la Martinique et de la Réunion).
À partir de ce fichier, on constate qu'il y a 36 677 communes habitées.
Parmi elles, il y a 11 094 communes dont la population est un nombre qui commence par le chiffre 1.
Cette observation vous semble-t-elle compatible avec l'affirmation : « le premier chiffre de la population des communes en France au 1^{er} janvier 2016 suit la loi de Benford » ?

2.b. Deuxième cas

Pour chaque candidat au baccalauréat de la session 2017, on considère sa taille en centimètres.
On désigne par X la variable aléatoire égale au premier chiffre de la taille en centimètres d'un candidat pris au hasard.
La loi de Benford vous semble-t-elle une loi adaptée pour X ?

CORRECTION

1. $P(X=1)=\dfrac{\ln(1+1)-\ln(1)}{\ln(10)}$ = 0,301 à 10^{-3} près.

Remarque (*résultats non demandés*)
On peut déterminer la loi de Benford en donnant les résultats sous la forme d'un tableau.
Pour tout nombre entier c compris entre 1 et 9 :
$p_c=P(X=c)=\dfrac{\ln(c+1)-\ln(c)}{\ln(10)}$

x_c	1	2	3	4	5	6	7	8	9
p_c	0.301	0.176	0.125	0.097	0.079	0.067	0.058	0.051	0.046

On peut vérifier que :
$p_1+p_2+\ldots+p_9=1$.

2.a. $f=\dfrac{11094}{36677}=0,302$ est la fréquence des communes dont le premier chiffre de la population est 1.

Pour déterminer si cette observation est compatible avec l'affirmation, on calcule un intervalle de fluctuation asymptotique au seuil de 95 %.
$n=36677\geqslant 30 \quad p=0,301 \quad np\geqslant 5 \quad n(1-p)\geqslant 5$.

$$I=\left[0,301-1,96\times\sqrt{\frac{0,301\times 0,639}{36677}}\,;0,301+\sqrt{\frac{0,301\times 0,699}{36677}}\right].$$

En utilisant la calculatrice, on obtient : $1,96\times\sqrt{\dfrac{0,301\times 0,699}{36677}}=0,0047$ à 10^{-4} près.

$I=[0,2963\,;0,3057]$
f=0,302 appartient à cet intervalle
Conclusion
Cette observation est compatible avec l'affirmation : » le premier chiffre de la population des communes en France au 1er janvier 2016 suit la loi de Benford », avec un risque d'erreur de 5 %.

2.b. L'immense majorité des candidats ont une taille comprise entre 100 cm et 199 cm, quelques candidats seulement ont une taille supérieure ou égale à 200 cm ou inférieure ou égale à 99 cm.
Aucun candidat n'a une taille de 30 cm, 40 cm, . . .
Donc $P(X=1)$ est une valeur voisine de 1.
On peut donc affirmer que la loi de Benford n'est pas adaptée pour X.