# Devoir Maison n° 7

**À rendre le 05/02/24**

## Exercice n°1

### Partie 1 : Étude d'une (suite de) fonction(s)

Pour tout $n \in \mathbb{N}$, on définit la fonction $g_n : [0, +\infty[ \to \mathbb{R}$ par :

$$g_n(x) = \frac{(\ln(1+x))^n}{(1+x)^2}$$

1. (a) Étudier les variations de la fonction $g_0$.
   On précisera la limite de $g_0$ en $+\infty$, ainsi que l'équation de la tangente en 0 .
   )

   (b) Donner l'allure de la courbe représentative de $g_0$.
2. Pour $n \geqslant 1$, justifier que $g_n$ est dérivable sur $[0, +\infty\ [$ et montrer que, pour tout $x \geq 0$,

$$g_n'(x) \geqslant 0 \quad \Longleftrightarrow \quad n \geqslant 2\ln(1+x)$$

   En déduire les variations de la fonction $g_n$ lorsque $n \geqslant 1$.
3. Calculer soigneusement $\lim_{x \to +\infty} g_n(x)$.
4. Montrer que, pour $n \geqslant 1, g_n$ admet un maximum sur $[0, +\infty[$ qui vaut

$$M_n = \left(\frac{n}{2e}\right)^n$$

   et déterminer la limite de $M_n$ lorsque $n$ tend vers $+\infty$.
5. Déterminer $\alpha > 1$ tel que, pour tout $n \geqslant 1$,

$$g_n(x) = o\left(\frac{1}{x^\alpha}\right), \quad x \to +\infty$$

### Partie 2 : Étude d'une suite d'intégrales

On pose pour tout $n \in \mathbb{N}$

$$I_n = \int_0^{+\infty} g_n(t)\mathrm{d}t$$

3. Montrer que l'intégrale $I_0$ est convergente et la calculer.
4. Montrer que pour tout entier $n \geqslant 1$, l'intégrale $I_n$ est convergente.
5. À l'aide d'une intégration par parties, montrer que, pour tout $n \in \mathbb{N}$,

$$I_{n+1} = (n+1)I_n$$

6. En déduire que, pour tout $n \in \mathbb{N}$,

$$I_n = n!$$

**Partie 3 : Une suite de variables aléatoires à densité**

Pour tout $n \in \mathbb{N}$, on définit la fonction $f_n$ sur $\mathbb{R}$ par :

$$f_n(x) = \begin{cases} 0, \text{ si } x < 0 \\ \dfrac{g_n(x)}{n!}, \text{ si } x \geq 0 \end{cases}$$

7. Montrer que $g_n$ peut être considérée comme une densité de probabilité.
   On note $X_n$ une v.a de densité $g_n$ et $F_n$ sa fonction de répartition.
8. Déterminer, pour tout $n \in \mathbb{N}$ et tout $x < 0, F_n(x)$.
9. Déterminer, pour $x \geq 0, F_0(x)$.
10. Soit $x \geqslant 0$ et $k \in \mathbb{N}^*$. Montrer que

$$F_k(x) - F_{k-1}(x) = -\frac{1}{k!}\frac{(\ln(1+x))^k}{1+x}$$

11. En déduire une expression de $F_n(x)$ pour $x \geqslant 0$ et $n \in \mathbb{N}^*$ faisant intervenir une somme qu'on ne cherchera pas à calculer.
12. Pour $x \in \mathbb{R}_+$fixé, déterminer la limite de $F_n(x)$ lorsque $n$ tend vers $+\infty$.
13. (KHÛBES) La suite de variables aléatoires $(X_n)$ converge-t-elle en loi ?
14. On introduit alors la variable $Y_n = \ln(1 + X_n)$.
    (a) Justifier que $Y_n$ est bien définie, puis que $Y_n(\Omega) \subset \mathbb{R}_+$.
    (b) Justifier que $Y_n$ admet une espérance et la calculer.
    (c) Justifier que $Y_n$ admet une variance et la calculer.
    (d) On note $H_n$ la fonction de répartition de $Y_n$. Montrer que, pour tout $x \in \mathbb{R}$, on a

$$H_n(x) = F_n(e^x - 1)$$

    (e) En déduire que $Y_n$ est une variable à densité et préciser une densité de $Y_n$.
    (f) Reconnaître la loi de $Y_0$. Déduire de la Partie 2. que $Y_0$ admet, pour tout $k \in \mathbb{N}^*$, un moment d'ordre $k$ et préciser sa valeur.

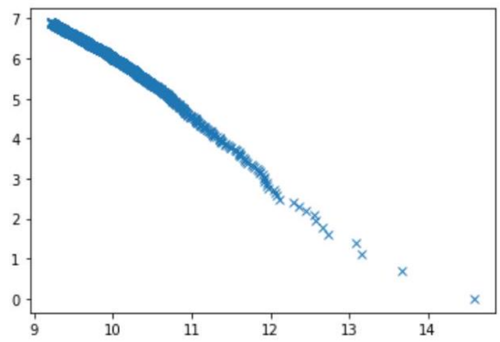
## Exercice n°3 : Loi de Pareto-Zipf

On souhaite modéliser la loi de probabilité de la variable aléatoire $T$ qui à une ville, choisie au hasard parmi les villes françaises, associe l'effectif de sa population. On note $n$ le nombre de villes.

1. (a) Pour 2018, on dispose d'un fichier Data1.csv de données provenant de l'INSEE et on utilise la bibliothèque Pandas. On exécute les instructions suivantes :

```
import pandas as pd; import numpy.random as rd;import numpy as np;import matplotlib.pyplot as p

dataset=pd.read_csv('Data1.csv',sep=";")

données=dataset[["Libellé","Population" ]]

données=données.sort_values(by =["Population","Libellé"], ascending=False, ignore_index=True)

logPopu=[np.log(t) for t in données["Population"]]

logRang=[np.log(k) for k in range (1,len(logPopu)+1)]

plt.plot(logPopu,logRang,'*')
```

ce qui produit le graphique suivant :




Expliquer pourquoi le programme et le graphique précédent justifient qu'il existe deux réels $a$ et $b$ tels que, pour toute ville, le réel $\frac{a}{t^b}$, où $t$ est l'effectif de celle-ci, est une approximation raisonnable du rang de celle-ci dans la liste des villes classées par ordre décroissant de population ?

Quelle grandeur peut-on calculer pour confirmer ce que l'on constate graphiquement ? Quelle méthode peut-on utiliser pour obtenir le meilleur couple $(a, b)$ en un certain sens ?

(b) On suppose que l'on a pour toutes les villes, $r = \frac{a}{t^b}, r$ étant le rang de cette ville, $t$ son effectif, $a$ et $b$ deux réels identiques pour toutes les villes. Si $x$ est l'effectif d'une des villes françaises, quelle est, en fonction de $a, b, x$ et $n$, la proportion de villes dont la population urbaine est supérieure ou égale à $x$ ?

- On suppose désormais que $T$ suit la loi de Pareto de paramètre $\theta > 1$ et $x_0 > 0$, c'est à dire qu'elle admet pour densité la fonction $f$ définie par :

$$f(x) = \begin{cases} \theta \frac{x_0^\theta}{x^{\theta+1}} \text{ si } x \geq x_0 \\ 0 \text{ sinon} \end{cases}$$

On note $\text{Par}(\theta, x_0)$ cette loi.

2. (a) Déterminer la fonction de répartition $F$ de $T$.

   (b) En calculant $\mathbb{P}([T > x])$, montrer que ce résultat est cohérent avec le résultat de la question 1 et exprimer $x_0$ et $\theta$ en fonction de $a, b$ et $n$.

   (c) Montrer que $\mathbb{E}(T)$ existe et vaut $\frac{\theta}{\theta-1} x_0$.

3. (a) Montrer que pour tout $t \geq x \geq x_0, \mathbb{P}_x(T > t) = \left(\frac{x}{t}\right)^\theta$.

   (b) On pose pour tout $t \in \mathbb{R}, F_x(t) = \mathbb{P}_x([T \leq t])$. De quelle loi $F_x$ est-elle la fonction de répartition ? Quelle est alors l'espérance de cette loi ?

4. Soit $\delta \in ]1, +\infty[$.

   - On suppose que $Y$ est une variable aléatoire à valeurs dans $[x_0, +\infty[$ dont $f_Y$ est une densité, continue sur $[x_0, +\infty[$, $F_Y$ la fonction de répartition. On suppose que $\forall x \geq x_0, \mathbb{P}([Y \geq x]) > 0$.

   (a) Soit $x \geq x_0$. Montrer que la fonction $G_x$ définie sur $\mathbb{R}$ par $G_x : t \mapsto \mathbb{P}_{[Y \geq x]}(Y \leq t)$ est la fonction de répartition d'une variable aléatoire à densité dont on déterminera une densité.

   - On suppose que pour tout $x \geq x_0$, l'espérance d'une variable aléatoire de fonction de répartition $G_x$ existe et vaut $\delta x$.

   (b) En déduire que pour tout $x \geq x_0, (\delta - 1)x f_Y(x) = \delta (1 - F_Y(x))$.

   (c) Résoudre l'équation différentielle $(1 - \delta)xy' - \delta y = 0$ sur $[x_0, +\infty[$.

   (d) En conclure que $Y$ suit une loi de Pareto dont on précisera les paramètres.