Math ECG 2. 2022-2023

Mathématiques Appliquées - F. Gaunard
http://frederic.gaunard.com
ENC Bessières, Paris 17e.

Devoir surveillé n°3 - sujet B

Samedi 3 Décembre
Durée : 4 heures

Les questions précédées de $(*)$ sont réservées aux khubes.

Exercice 1

(1) Déterminer la nature de la série $\displaystyle\sum_{k\geq 2} \frac{\ln(k)}{k}$.

(2) On introduit la fonction f définie sur $\mathbb{R}_+^*$ par $f(x) = \dfrac{\ln(x)}{x}$.

(a) Peut-on prolonger f par continuité en 0?
(b) Justifier que f est dérivable sur $\mathbb{R}_+^*$ et calculer, pour tout $x > 0$, $f'(x)$.
(c) Dresser le tableau de variations complet de f, limites comprises.

(3) Dans toute la suite, on note, pour $n \geq 2$, $S_n = \displaystyle\sum_{k=2}^{n} \frac{\ln(k)}{k}$.

(a) Montrer que, pour tout entier $k \geq 3$, on a

$$f(k+1) \leq \int_k^{k+1} f(t)\mathrm{d}t \leq f(k).$$

(b) En déduire que, pour tout entier $n > 3$,

$$\int_3^n \frac{\ln(t)}{t}\mathrm{d}t + \frac{\ln(2)}{2} \leq S_n \leq \int_3^n \frac{\ln(t)}{t}\mathrm{d}t + \frac{\ln(2)}{2} + \frac{\ln(3)}{3}.$$

(c) En déduire alors un équivalent simple de S_n, lorsque $n \to +\infty$.

(4) Le but de cette question est de montrer qu'il existe un nombre réel κ tel que

$$S_n = \frac{(\ln(n))^2}{2} + \kappa + o(1), \qquad n \to +\infty.$$

(a) Montrer que

$$(\ln(n))^2 - (\ln(n-1))^2 = \frac{2\ln(n)}{n} + \frac{\ln(n)}{n^2} + o\left(\frac{\ln(n)}{n^2}\right), \qquad n \to +\infty.$$

(b) On pose alors $u_n = \dfrac{\ln(n)}{n} - \dfrac{1}{2}\left((\ln(n))^2 - (\ln(n-1))^2\right)$.
Montrer que

$$\sum_{k=2}^{n} u_k = S_n - \frac{(\ln(n))^2}{2}.$$

(c) Conclure.

Exercice 2

Supposons que vous soyez le chef de direction d'une franchise de camions ambulants (*Food Trucks*). Vous envisagez différentes villes pour ouvrir un nouveau point de vente. La chaîne a déjà des camions dans différentes villes et vous avez des données pour les bénéfices et les populations des villes. Vous souhaitez utiliser ces données pour vous aider à choisir la ville pour y ouvrir un nouveau point de vente.

On dispose d'un fichier `data.csv` et on utilise la bibliothèque `pandas`.

(1) On exécute les instructions suivantes qui donne l'affichage ci-après. Que contient le fichier `data.csv` importé ?

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

donnees=pd.csv_read('data.csv', sep=';')
donnees.head()
```

Affichage Python

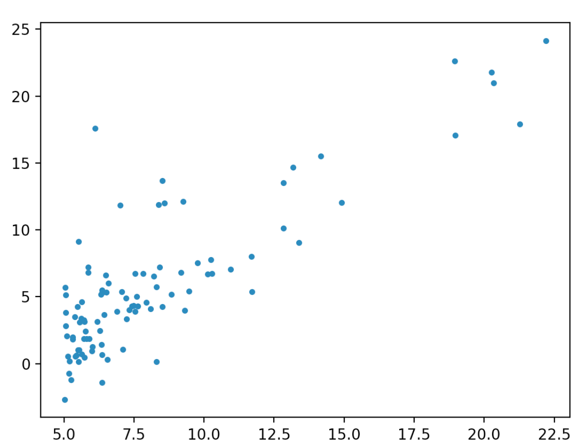
```
>>> donnees.head()
   Population (en 10k)  Profit (en 10k EUR)
0               6.1101              17.5920
1               5.5277               9.1302
2               8.5186              13.6620
3               7.0032              11.8540
4               5.8598               6.8233
```

(2) On ajoute les commandes suivantes

```
table=donnees.rename(columns={'Population (en 10k)' : 'pop',
                              'Profit (en 10k EUR)':'profit'})

X=table['pop']
Y=table['profit']
plt.grid()
plt.plot(X,Y, '.')
plt.show()
```

Affichage Python



(a) Que représente cette figure ?
(b) Expliquer pourquoi la figure ci-dessus permet de conjecturer qu'il existe deux réels a, b tels que $ax+b$, où x est le nombre d'habitants de la ville (en dizaines de milliers d'habitants), est une approximation raisonnable du profit (en dizaines de milliers d'euros) d'un *Food Truck* installé dans cette même ville.
(c) Quelle quantité pourrait-on calculer pour conforter cette approximation? Donner une suite d'instruction en Python permettant de la calculer.
(d) On suppose qu'on a été en mesure de répondre à la question précédente correctement. L'exécution des commandes affiche alors une valeur de `0.8378733891854535`. Est-ce cohérent?
(e) Il y a 182354 habitants à *Legumeville* et pas encore de *Food Truck*. Quelle(s) commande(s) Python permettraient d'estimer raisonnablement le profit suivant l'installation d'un camion dans cette localité ?

(3) Votre société a beau être établie en zone euro, son siège social est dans le Delaware aux Etats-Unis, et on décide d'exprimer le profit en dollars. Sachant qu'un euro vaut au moment de faire le calcul 1.03 dollar, que devient la covariance des séries statistiques habitants/profit ? Même question avec le coefficient de corrélation linéaire.

Exercice 3

On considère les fonctions e_1, e_2, e_3 et e_4 définies sur $\mathbb{R}_+^*$ par

$$e_1(x) = x, \quad e_2(x) = x^2, \quad e_3(x) = x\ln(x) \quad \text{et} \quad e_4(x) = x^2\ln(x)$$

et on note E l'espace vectoriel engendré par e_1, e_2, e_3 et e_4

$$E = \mathrm{Vect}(e_1, e_2, e_3, e_4).$$

Ainsi, E est un espace vectoriel dont les vecteurs sont des fonctions définies sur $\mathbb{R}_+^*$, son vecteur nul 0_E est la fonction identiquement nulle sur $\mathbb{R}_+^*$.

(1) Le but de cette question est de montrer que la famille (e_1, e_2, e_3, e_4) est libre. Pour cela, on considère quatre réels a, b, c et d tels que

$$ae_1 + be_2 + ce_3 + de_4 = 0_E,$$

c'est à dire que, **pour tout** $x \in \mathbb{R}$,

$$ax + bx^2 + cx\ln(x) + dx^2\ln(x) = 0 \qquad (\star)$$

(a) Avec un choix judicieux de x dans la relation $(\star)$, montrer que $a + b = 0$.
(b) Établir que, pour tout $x > 1$,

$$\frac{a}{x\ln(x)} + \frac{b}{\ln(x)} + \frac{c}{x} + d = 0.$$

En déduire que $d = 0$.
(c) Utiliser le même type de raisonnement pour montrer que $b = 0$.
(d) Montrer finalement que $a = b = c = d = 0$.

(2) Expliciter une base et préciser la dimension de E.

(3) On note u l'application qui à toute fonction f de E associe la fonction $g = u(f)$ définie par

$$\forall x \in \mathbb{R}_+^*, \qquad g(x) = u(f)(x) = xf'(x).$$

(a) Montrer que u est une application linéaire.
(b) Déterminer $u(e_1), u(e_2), u(e_3)$ et $u(e_4)$.
(c) En déduire que u est un endomorphisme de E.

(4) (a) Donner la matrice A de u dans la base (e_1, e_2, e_3, e_4).
(b) Montrer que u est un automorphisme de E.
(c) Montrer que
$$\dim(\mathrm{Ker}(A-I)) = \dim(\mathrm{Ker}(A-2I)) = 1.$$

(5) En écrivant $A = D + N$ où D est une matrice diagonale et N une matrice telle que $N^2 = 0$, montrer que, pour tout $n \in \mathbb{N}^*$,
$$A^n = \begin{pmatrix} 1 & 0 & n & 0 \\ 0 & 2^n & 0 & n2^{n-1} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 2^n \end{pmatrix}.$$

Problème

Une urne contient des boules blanches (en proportion p), des boules noires (en proportion q) et des boules rouges (en proportion r). On a donc $p + q + r = 1$ (on suppose que $p, q, r \in]0;1[$). On effectue des tirages successifs **avec remise** dans cette urne. Pour $i \in \mathbb{N}^*$, on note T_i la variable qui vaut 1 si la $i-$ème boule tirée est blanche, -1 si elle est noire et 0 si elle est rouge.
Les tirages étant effectués avec remise, les variables (T_i) sont donc mutuellement indépendantes.

Partie A

On note ensuite X_1 la variable aléatoire égale au numéro du tirage qui amène pour la première fois une boule blanche et X_2 celui correspondant au tirage où sort pour la première fois la **deuxième** boule blanche.

Par exemple, si les premiers tirages donnent *Noire, Rouge, Noire, Blanche, Rouge, Blanche, ...*, on a $X_1 = 4$ et $X_2 = 6$.

(1) Expliciter, pour tout $i \in \mathbb{N}^*$, la loi de T_i. Calculer son espérance et sa variance.

(2) Reconnaître la loi de X_1. Rappeler son espérance et sa variance.

(3) Compléter la fonction Python suivante, afin qu'elle simule les variables X_1 et X_2. Pourquoi la fonction ne prend-elle en argument que a proportion p de boules blanches?

```
import numpy as np
import numpy.random as rd

def simul_X1_X2(p):
    x1=.......
    while ................ :
        x1=.............
    x2=..........
    while ................ :
        x2=................
    return [x1, x2]
```

(4) (a) Expliciter la loi conjointe du couple (X_1, X_2).
(b) En déduire la loi marginale de X_2.
(c) Montrer que X_2 admet une espérance et que $E(X_2) = \dfrac{2}{p}$.

(5) On note $U_2 = X_2 - X_1$.

(a) Pour $j \in \mathbb{N}^*$, déterminer $P(U_2 = j)$.
En déduire que X_1 et U_2 suivent la même loi, puis que U_2 admet une variance et préciser sa valeur.

(b) Montrer que U_2 est indépendante de X_1.

(c) Exprimer X_2 en fonction de U_2 et X_1 et en déduire que X_2 admet une variance qu'on explicitera.

(d) Que vaut $\mathrm{cov}(X_1, X_2)$? Les variables X_1 et X_2 sont-elles indépendantes?

(e) On rajoute les instructions Python suivantes à la suite de la fonction précédente. Que peut-on prévoir quant à l'affichage après exécution de celles-ci ?

```
p=1/4;
L=[ ]; M=[ ];
for k in range(10000):
    [X1,X2]=simul_X1_X2(p)
    L=L.append(X1)
    M=M.append(X2)
U=[M[k]-L[k] for k in range(10000)]
print(np.mean([L[k]*U[k] for k in range(10000)])-np.mean(L)*np.mean(U))
```

Partie B

On note W la variable correspondant au nombre de boules rouges obtenues avant l'obtention de la première boule blanche.

(6) Pour $i \in \mathbb{N}^*$, déterminer la loi conditionnelle de W sachant $[X_1 = i]$.

(7) En déduire que la loi de W est donnée par la somme, pour $k \in \mathbb{N}$,

$$P(W = k) = p\left(\frac{r}{q+r}\right)^k \sum_{i=k+1}^{+\infty} \binom{i-1}{k} \left(\frac{q}{q+r}\right)^{i-1-k} (1-p)^{i-1}$$

(8) Vérifier que, pour tout $i \in \mathbb{N}^*$ et pour tout $k \in [\![0, i]\!]$,

$$k\binom{i}{k} = i\binom{i-1}{k-1}$$

(9) En admettant qu'il est licite de permuter les sommes

$$\sum_{k=0}^{+\infty} \sum_{i=k+1}^{+\infty} \cdots = \sum_{i=1}^{+\infty} \sum_{k=0}^{i-1} \cdots ,$$

et que W admet une espérance, montrer que

$$E(W) = \frac{r}{p}$$

Partie C

On note, pour $n \in \mathbb{N}^*$, $S_n = T_1 + ... + T_n$.

(10) Quelle est la loi de S_1? Préciser son espérance et sa variance.

(11) Expliciter l'espérance et la variance de S_n.

(12) Soit $t > 0$. On pose $V_n = t^{S_n}$.

(a) Déterminer la loi, l'espérance et la variance de V_1.

(b) En déduire l'expression de $E(V_n)$.