DEVOIR NUMÉRO 5 VERSION B

ECG2 MATHS APPLIQUÉES

Exercice

On considère la fonction f qui à tout couple (x, y) de $\mathbb{R}^2$ associe le réel :

$$f(x, y) = x^4 + y^4 - 2\,(x-y)^2$$

1. Justifier que f est de classe $\mathcal{C}^1$ sur $\mathbb{R}^2$.

2. **a.** Calculer les dérivées partielles d'ordre 1 de f.

b. Montrer que le gradient de f est nul si, et seulement si, on a : $\begin{cases} x^3 - x + y & = & 0 \\ y^3 + x - y & = & 0 \end{cases}$.

c. En déduire que f possède trois points critiques : $(0,0)$, $(\sqrt{2}, -\sqrt{2})$, $(-\sqrt{2}, \sqrt{2})$.

3. **a.** Calculer les dérivées partielles d'ordre 2 de f.

b. Écrire la matrice hessienne de f en chaque point critique.

c. Déterminer les valeurs propres de chacune de ces trois matrices puis montrer que f admet un minimum local en deux de ses points critiques. Donner la valeur de ce minimum.

d. Déterminer les signes de $f(x,x)$ et $f(x,-x)$ au voisinage de $x = 0$.
Conclure quant à l'existence d'un extremum en le troisième point critique de f.

4. **a.** Pour tout (x, y) de $\mathbb{R}^2$, calculer $f(x, y) - (x^2-2)^2 - (y^2-2)^2 - 2(x+y)^2$.

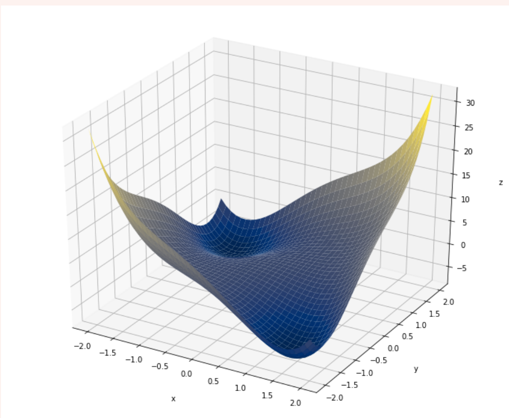
b. Que peut-on déduire de ce calcul quant au minimum de f ?

5. **a.** Compléter la deuxième ligne du programme suivant afin de définir la fonction f en Python.

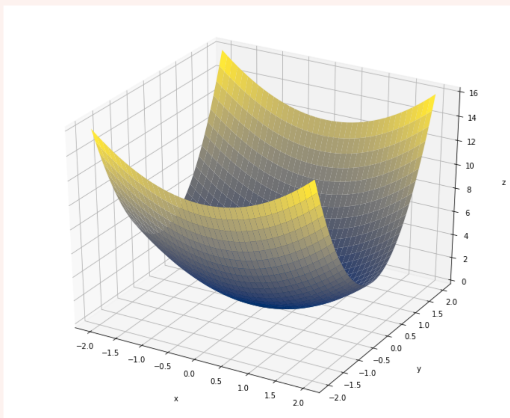
```python
def f(x,y):
    return ................
```

b. On utilise la fonction précédente pour tracer le graphe de f et Python renvoie l'une des trois nappes suivantes. Laquelle ? Justifier la réponse.

Graphe 1

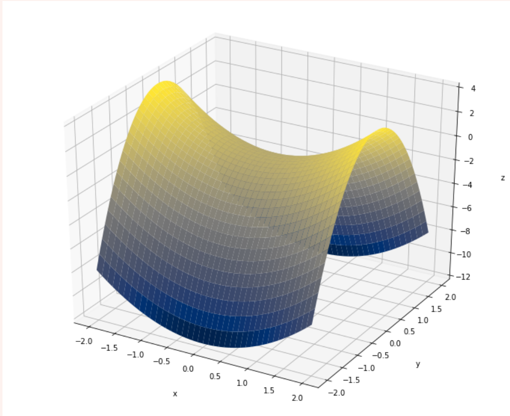


Graphe 2



Date: 30 Mars 2024 8h30-12h30.
http://louismerlin.fr.

Graphe 3



Problème 1

Partie 1 : Loi de Gumbel. Soient m un réel et a un réel strictement positif. On introduit la fonction $f_{m,a}$ définie sur $\mathbb{R}$ par

$$\forall x \in \mathbb{R}, \qquad f_{m,a}(x) = \frac{1}{a}\exp\left(\frac{m-x}{a} - e^{\frac{m-x}{a}}\right).$$

1. Montrer que $f_{m,a}$ peut être considérée comme une densité de probabilité.

On dira qu'une variable aléatoire Z ayant pour densité $f_{m,a}$ suit la loi de Gumbel de paramètre m et a et on écrira $Z \hookrightarrow \mathcal{G}(m,a)$.

2. Déterminer l'expression de la fonction de répartition $F_{m,a}$ d'une v.a $Z \hookrightarrow \mathcal{G}(m,a)$.

3. Montrer que

$$Z \hookrightarrow \mathcal{G}(0,1) \iff aZ + mZ \hookrightarrow \mathcal{G}(m,a).$$

4. Soit $X \hookrightarrow \mathcal{E}(1)$. Montrer que $-\ln(X) \hookrightarrow \mathcal{G}(0,1)$.

5. Python. Déduire des deux questions précédentes l'écriture d'une fonction d'en-tête `def Gumbel(m, a)` qui renvoie une simulation d'une loi de Gumbel de paramètres m et a.

6. On note Ψ la fonction de répartition de la loi $\mathcal{G}(0,1)$. Soit $\alpha \in]0;1[$. Montrer qu'il existe deux réels c_α et d_α (que l'on déterminera) tels que

$$\Psi(c_\alpha) = 1-\alpha, \quad \text{et} \quad \Psi(d_\alpha) = \alpha.$$

Partie 2 : $n-$échantillon de lois exponentielles et loi de Gumbel. Soit $n \in \mathbb{N}^*$. On considère un $n-$échantillon $(X_1, X_2, ..., X_n)$ de la loi exponentielle $\mathcal{E}(\lambda)$ de paramètre $\lambda > 0$. On pose $L_n = \min(X_1, ..., X_n)$ et $M_n = \max(X_1, ..., X_n)$.

7. Déterminer les fonctions de répartition des variables L_n puis M_n.

8. Quelle est la loi de la variable $Y_n = n\lambda L_n$?

9. On considère maintenant $Z_n = \lambda M_n - \ln(n)$.

- **a.** Déterminer la fonction de répartition F_n de Z_n.
- **b.** Pour $t \in \mathbb{R}$, déterminer la limite $F(t)$ de $F_n(t)$ lorsque $n \to +\infty$.
- **c.** En déduire soigneusement que Z_n converge en loi vers une variable aléatoire Z suivant une loi de Gumbel dont on précisera les paramètres.

Partie 3 : Construction de deux intervalles de confiance. La durée de vie d'une ampoule est modélisée par une variable aléatoire $X \hookrightarrow \mathcal{E}(\lambda)$ où $\lambda > 0$ est inconnu. On cherche à estimer la durée de vie moyenne $\mu = E(X) = 1/\lambda$ et on dispose d'un échantillon de n ampoules (dont les durées de vie sont supposées indépendantes).

10. Dans cette question, on suppose que la seule information dont on dispose est la durée de vie de l'ampoule qui a *grillé* le plus tôt.

a. À l'aide de la Question (8), proposer un estimateur $\tilde{L}_n$ de μ, construit à partir de L_n, qui soit sans biais.

b. Quel est son risque quadratique ?

c. Montrer que, pour tout $\varepsilon > 0$

$$P\left(\left|\tilde{L}_n - \mu\right| > \varepsilon\right) = 1 - e^{-1+\lambda\varepsilon} + e^{-1-\lambda\varepsilon}.$$

d. L'estimateur $\tilde{L}_n$ est-il convergent ?

e. Soit $\alpha \in]0;1[$. Montrer que si $Y \hookrightarrow \mathcal{E}(1)$, alors

$$P\left(Y < -\ln\left(1 - \frac{\alpha}{2}\right)\right) = P\left(Y > -\ln\left(\frac{\alpha}{2}\right)\right) = \frac{\alpha}{2}.$$

f. Montrer alors que l'intervalle $I_{\alpha,n}$ ci-dessous est un intervalle de confiance au seuil $1-\alpha$ pour μ :

$$I_{\alpha,n} = \left[\frac{\tilde{L}_n}{-\ln(\alpha/2)}, \frac{\tilde{L}_n}{-\ln(1-\alpha/2)}\right].$$

11. Dans cette question, on suppose que l'on connaît la durée de vie de la dernière ampoule. Soit $\alpha \in]0;1[$.

a. Expliciter les valeurs de $c_{\alpha/2}$ et $d_{\alpha/2}$.

b. À l'aide de la Question (9c), montrer que $J_{\alpha,n}$ est un intervalle de confiance asymptotique au seuil $1-\alpha$ pour μ :

$$J_{\alpha,n} = \left[\frac{M_n}{\ln(n) - \ln\left(-\ln(1-\alpha/2)\right)}; \frac{M_n}{\ln(n) - \ln\left(-\ln(\alpha/2)\right)}\right]$$

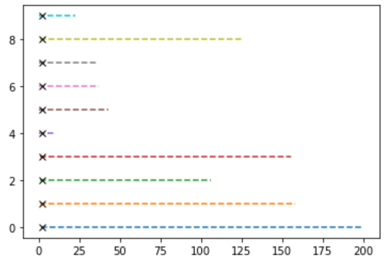
12. Python. On ajoute les commandes suivantes dont l'exécution produit les affichages ci-contre. Commenter et interpréter. Vaut-il mieux connaître L_n ou M_n ?

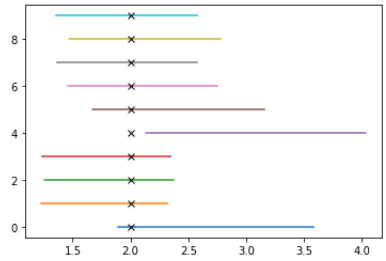
```python
def IDC1(alpha, X):
    L, n = min(X), len(X)
    return [n*L/(-np.log(alpha/2)), n*L/(-np.log(1-alpha/2))]

def h(t) :
    return -np.log(-np.log(t))
def IDC2(alpha, X):
    M, n = max(X), len(X)
    return [M/(np.log(n)+h(1-alpha/2)), M/(np.log(n)+h(alpha/2))]

alpha=0.05
for i in np.arange(10):
    sample=[rd.exponential(2) for k in range(1000)]
    [A, B]=IDC1(alpha, sample)
    plt.plot([A,B], [i, i], '--')
    plt.plot([2], [i], 'kx')
plt.show() # Figure 1
for i in np.arange(10):
    sample=[rd.exponential(2) for k in range(1000)]
    [A, B]=IDC2(alpha, sample)
    plt.plot([A,B], [i, i])
```

```
    plt.plot([2], [i], 'kx')
plt.show() # Figure 2
```





Problème 2

Dans tout l'exercice, les variables aléatoires sont définies sur le même espace probabilisé $(\Omega, \mathcal{T}, P)$.

Partie I : Préliminaires.

Dans cette partie, λ désigne un réel strictement positif.

1. Soit X une variable aléatoire suivant la loi exponentielle de paramètre λ.

a. Déterminer la fonction $x \mapsto P\left([X > x]\right)$ (appelée *fonction de survie de* X).

b. Pour tous nombres réels strictement positifs x et y, calculer la probabilité conditionnelle

$$P_{[X>x]}\left([X > x + y]\right).$$

Justifier alors que si X modélise la durée de vie d'un phénomène, on dise que ce dernier est "sans vieillissement" (on dit aussi que la loi exponentielle possède la propriété *d'absence de mémoire*).

2. Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires mutuellement indépendantes, de même loi exponentielle de paramètre λ. Pour tout entier naturel n non nul, on pose $S_n = \sum\limits_{k=1}^{n} X_k$.

a. Déterminer l'espérance de S_n.

b. Démontrer par récurrence que, pour tout $n \in \mathbb{N}^*$, la variable S_n admet pour densité la fonction f_n définie par

$$f_n(t) = \begin{cases} 0 & \text{si } t < 0 \\ \frac{\lambda^n}{(n-1)!} e^{-\lambda t} t^{n-1} & \text{si } t \geqslant 0 \end{cases}.$$

Pour cela, on admettra que si U et V sont des variables aléatoires indépendantes admettant respectivement pour densité les fonctions f_U et f_V, alors la variable aléatoire $U + V$ est aussi à densité et admet pour densité la fonction f_{U+V} définie sur $\mathbb{R}$ par

$$f_{U+V}(t) = \int_{-\infty}^{+\infty} f_U(x) f_V(t - x) dx.$$

Partie II : Loi de Pareto.

Soit a et b des réels strictement positifs. Par définition, on dit qu'une variable aléatoire suit la loi de Pareto de paramètres a et b si elle admet pour densité la fonction f définie sur $\mathbb{R}$ par

$$f(t) = \begin{cases} 0 & \text{si } t < b \\ a\frac{b^a}{t^{a+1}} & \text{si } t \geqslant b \end{cases}.$$

Soit X une variable aléatoire de loi de Pareto de paramètres a et b.

3. Vérifier que $\displaystyle\int_{-\infty}^{+\infty} f(t)dt = 1$, calculer l'espérance et la variance de X, en précisant à quelles conditions chacune de ces quantités existe.

4. Déterminer la fonction de répartition de X. Préciser la fonction de survie $x \mapsto P([X > x])$.

5. Démontrer que, pour tout réel y positif ou nul, la probabilité conditionnelle $P_{[X>x]}([X > x + y])$ tend vers 1 quand x tend vers $+\infty$. Comment s'interprète ce résultat sur un phénomène modélisé par X ?

6. On pose dans cette question $Y = \ln\left(\frac{X}{b}\right)$.

a. Démontrer que Y suit une loi exponentielle dont on précisera le paramètre.

b. Déduire de la question précédente une fonction Python nommée simul_X(a,b) qui permette de simuler la variable aléatoire X.

Partie III : Estimation des paramètres d'une loi de Pareto.

Les instants aléatoires des arrivées de paquets (symboles binaires représentant de l'information de type audio, vidéo, données,...) dans un canal de communication sont modélisés par une variable aléatoire suivant une loi de Pareto de paramètres α et β ($\alpha > 0$ et $\beta > 0$).

Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires indépendantes, toutes de même loi que X. On dit qu'un estimateur T_n de θ est *sans biais* si $E(T_n) = \theta$.

7. On suppose tout d'abord que le paramètre β fait partie des caractéristiques connues du canal de communication ; on se propose de déterminer un estimateur de α par une méthode dite "du maximum de vraisemblance". Pour cela, n désignant un entier naturel et $x_1, \cdots, x_n$ des réels supérieurs ou égaux à β, on introduit la fonction L, à valeurs dans $\mathbb{R}$ et définie sur $]0, +\infty[$ par

$$L(a) = f_a(x_1) \times \cdots \times f_a(x_n) = \prod_{k=1}^{n} f_a(x_k),$$

où f_a est la fonction définie sur $\mathbb{R}$ par

$$f_a(t) = \begin{cases} 0 & \text{si } t < \beta \\ a\frac{\beta^a}{t^{a+1}} & \text{si } t \geqslant \beta \end{cases}.$$

a. Exprimer $L(a)$ puis $\ln(L(a))$.

b. On considère la fonction φ de $]0, +\infty[$ dans $\mathbb{R}$ définie par

$$\varphi(a) = n\ln(a) + na\ln(\beta) - (a + 1)\sum_{k=1}^{n} \ln(x_k).$$

(i) Démontrer que la fonction φ admet un maximum, atteint en un unique réel w que l'on exprimera en fonction de $x_1, \cdots, x_n$ et de β.

(ii) Que peut-on dire de w pour la fonction L ?

c. On pose dorénavant, pour tout n de $\mathbb{N}^*$,

$$W_n = \frac{n}{\sum\limits_{k=1}^{n} \ln\left(\frac{X_k}{\beta}\right)}$$

(la suite $(W_n)_{n\in\mathbb{N}^*}$ est appelée *estimateur du maximum de vraisemblance*).

(i) Justifier que la variable aléatoire $\sum\limits_{k=1}^{n} \ln\left(\frac{X_k}{\beta}\right)$ admet pour densité la fonction f_n définie dans **I.2.b** en prenant $\lambda = \alpha$.

(ii) À l'aide du théorème de transfert, en déduire que W_n admet pour espérance $\frac{n\alpha}{n-1}$ lorsque $n \geqslant 2$, puis proposer un estimateur sans biais de α construit sur W_n.

d. On pose, pour tout n de $\mathbb{N}^*$, $W'_n = \frac{n-1}{n} W_n$.

(i) Soit n un entier naturel supérieur ou égal à 3.

En admettant que le moment d'ordre 2 de W'_n est égal à $\frac{(n-1)\alpha^2}{n-2}$, calculer la variance de W'_n puis établir, à l'aide de l'inégalité de Bienaymé-Tchebychev, que pour tout réel ε strictement positif, on a l'égalité

$$P\left(\left[W'_n - \varepsilon < \alpha < W'_n + \varepsilon\right]\right) \geqslant 1 - \frac{\alpha^2}{\varepsilon^2(n-2)}.$$

(ii) On suppose dans cette question (et elle seule) que α est strictement compris entre 1 et 2. Déterminer un entier naturel N tel que, pour tout entier supérieur ou égal à N,

$$\left[W'_n - \frac{1}{10}, W'_n + \frac{1}{10}\right]$$

soit un intervalle de confiance du paramètre α au niveau de confiance 0.95.

8. On suppose maintenant que la paramètre α est déjà identifié et qu'il vérifie $\alpha > 2$.

a. Pour tout entier strictement positif n, on pose $Y_n = c_n \sum_{k=1}^{n} X_k$ où c_n est un réel choisi de manière à ce que Y_n soit un estimateur sans biais de β.

(i) Calculer c_n.

(ii) Quelle est la limite de la variance de Y_n quand n tend vers $+\infty$? (on dit que l'estimateur est convergent).

b. Pour tout entier strictement positif, on pose $Z_n = \min(X_1, \cdots, X_n)$.

(i) Déterminer la fonction de répartition de Z_n, puis reconnaître sa loi et préciser son espérance. Quelle est la limite de l'espérance lorsque n tend vers $+\infty$?

(ii) Pour tout entier strictement positif n, on pose $Z'_n = d_n Z_n$ où d_n est choisi de telle sorte que $(Z_n)_{n\in\mathbb{N}^*}$ soit un estimateur sans biais de β.

Quelle est la limite de la variance de Z'_n quand n tend vers $+\infty$.

(iii) Entre $(Y_n)_{n\in\mathbb{N}^*}$ et $(Z_n)_{n\in\mathbb{N}^*}$, quel estimateur de β préférez-vous ?