Révisions 1

Correction du sujet ESSEC

1. La fonction f est clairement continue et positive.

   Montrons la convergence de l'intégrale $\displaystyle\int_{-\infty}^{+\infty} f(t)dt$ et calculons-la.

   D'une part, si $A > \alpha$:

$$\begin{aligned}\int_{\alpha}^{A} \frac{1}{2\beta}\exp\left(-\frac{|t-\alpha|}{\beta}\right) &= \frac{1}{2\beta}\int_{\alpha}^{A}\exp\left(-\frac{t-\alpha}{\beta}\right)dt = \frac{1}{2\beta}\int_{\alpha}^{A}\exp\left(\frac{-t+\alpha}{\beta}\right)dt \\ &= \frac{1}{2\beta}\left[-\beta\exp\left(\frac{-t+\alpha}{\beta}\right)\right]_{\alpha}^{A} = \frac{1}{2}\left(-\exp\left(\frac{-A+\alpha}{\beta}\right)+1\right) \\ &\underset{A\to+\infty}{\longrightarrow} \frac{1}{2}.\end{aligned}$$

   D'autre part, si $B < \alpha$:

$$\begin{aligned}\int_{B}^{\alpha} \frac{1}{2\beta}\exp\left(-\frac{|t-\alpha|}{\beta}\right) &= \frac{1}{2\beta}\int_{B}^{\alpha}\exp\left(-\frac{-t+\alpha}{\beta}\right)dt = \frac{1}{2\beta}\int_{B}^{\alpha}\exp\left(\frac{t-\alpha}{\beta}\right)dt \\ &= \frac{1}{2\beta}\left[\beta\exp\left(\frac{t-\alpha}{\beta}\right)\right]_{B}^{\alpha} = \frac{1}{2}\left(1-\exp\left(\frac{B-\alpha}{\beta}\right)\right) \\ &\underset{B\to-\infty}{\longrightarrow} \frac{1}{2}.\end{aligned}$$

   Finalement, $\displaystyle\int_{-\infty}^{+\infty} f(t)dt$ converge et $\displaystyle\int_{-\infty}^{+\infty} f(t)dt = \int_{-\infty}^{\alpha} f(t)dt + \int_{\alpha}^{+\infty} f(t)dt = \frac{1}{2}+\frac{1}{2} = 1.$

   Par conséquent, f est une densité de probabilité.

2. - Pour $x \leq 0$:

$$\Psi(x) = \int_{-\infty}^{x} \frac{1}{2}\exp(t)dt = \lim_{A\to-\infty}\left[\frac{1}{2}\exp(t)\right]_{A}^{x} = \frac{1}{2}e^{x}.$$

   - Pour $x \geq 0$:

$$\Psi(x) = \Psi(0) + \int_{0}^{x}\frac{1}{2}\exp(-t)dt = \frac{1}{2} + \left[-\frac{1}{2}\exp(-t)\right]_{0}^{x} = 1 - \frac{1}{2}e^{-x}.$$

3. (a) Soit $Y = \beta X + \alpha$. Déterminons la fonction de répartition F_Y de Y ; pour tout $x \in \mathbb{R}$:

$$F_Y(x) = P(Y \leq x) = P(\beta X + \alpha \leq x) \underset{(\beta>0)}{=} P\left(X \leq \frac{x-\alpha}{\beta}\right) = \Psi\left(\frac{x-\alpha}{\beta}\right).$$

   Comme composée de $x \mapsto \dfrac{x-\alpha}{\beta}$ et de Ψ, toutes deux de classe $\mathcal{C}^1$ sur $\mathbb{R}$ (Ψ est la fonction de répartition d'une variable à densité continue), la fonction F_Y est de classe $\mathcal{C}^1$ sur $\mathbb{R}$; donc Y est une variable à densité et la dérivée de F_Y en est une densité :

$$\forall x \in \mathbb{R},\ {F_Y}'(x) = \frac{1}{\beta}\Psi'\left(\frac{x-\alpha}{\beta}\right) = \frac{1}{\beta}f\left(\frac{x-\alpha}{\beta}\right) = \frac{1}{\beta}\times\frac{1}{2}\exp\left(-\left|\frac{x-\alpha}{\beta}\right|\right) = f(x).$$

   Ainsi f est une densité de Y qui suit donc la loi $\mathcal{L}(\alpha,\beta)$.

(b) Avec les questions 3.(a) et 2. :

$$F_Y(x) = \Psi\left(\frac{x-\alpha}{\beta}\right) = \begin{cases} \frac{1}{2}\exp\left(\frac{x-\alpha}{\beta}\right) & \text{si } x \leq \alpha \\ 1 - \frac{1}{2}\exp\left(\frac{-x+\alpha}{\beta}\right) & \text{si } x \geq \alpha \end{cases}.$$

4. (a) La fonction $f : \begin{cases} \mathbb{R} & \to & \mathbb{R} \\ t & \mapsto & \frac{1}{2}\exp(-|t|) \end{cases}$ est une densité de $X \hookrightarrow \mathcal{L}(0,1)$. C'est une fonction paire.

- L'existence d'une espérance pour X équivaut à la convergence absolue de $\int_{-\infty}^{+\infty} tf(t)dt$. La fonction $t \mapsto tf(t)$ étant impaire, il s'agit de justifier la convergence de $\int_0^{+\infty} tf(t)dt$: il s'agit de l'intégrale d'une fonction continue et positive sur $[0,+\infty[$, négligeable devant $t \mapsto \frac{1}{t^2}$ au voisinage de $+\infty$.

  Comme f est paire : $E(X) = 0$.
- De même, l'intégrale $\int_0^{+\infty} t^2 f(t)dt$ converge par comparaison avec une intégrale de Riemann ; la fonction $t \mapsto t^2 f(t)$ étant impaire, ceci donne la convergence de $\int_{-\infty}^{+\infty} t^2 f(t)dt$ et :

$$E(X^2) = \int_{-\infty}^{+\infty} t^2 f(t)dt = 2\int_0^{+\infty} t^2 f(t)dt = 2\int_0^{+\infty} t^2 \frac{1}{2}e^{-t}dt = \int_0^{+\infty} t^2 e^{-t}dt.$$

  On reconnait le moment d'ordre 2 d'une variable aléatoire Z suivant la loi exponentielle de paramètre 1 :

$$m_2(Z) = V(Z) + E(Z)^2 = 1 + 1^2 = 2.$$

  Par conséquent :

$$V(X) = E(X^2) - E(X)^2 = 2 - 0^2 = 2.$$

(b) Soit $Y \hookrightarrow \mathcal{L}(\alpha,\beta)$. Soit $X \hookrightarrow \mathcal{L}(0,1)$. La variable aléatoire Y suit la même loi que $\beta X + \alpha$ d'après 3.(a) ; cette dernière admet une espérance et une variance comme X avec :

$$E(Y) = E(\beta X + \alpha) = \beta E(X) + \alpha = \alpha \quad ; \quad V(Y) = V(\beta X + \alpha) = \beta^2 V(X) = 2\beta^2.$$

5. (a) Déterminons la fonction de répartition F_X de X en utilisant la formule des probabilités totales avec le système complet d'événements $\{(V=1),(V=0)\}$; pour tout $x \in \mathbb{R}$:

$$\begin{aligned} F_X(x) &= P(X \leq x) \\ &= P\big((V=1)\cap(X\leq x)\big) + P\big((V=0)\cap(X\leq x)\big) \\ &= P\big((V=1)\cap(U\leq x)\big) + P\big((V=0)\cap(-U\leq x)\big) \\ &= P(V=1)\times P(U\leq x)\big) + P(V=0)\times P(-U\leq x) \\ &\qquad \text{(indépendance de } U \text{ et } V) \\ &= \frac{1}{2}F_U(x) + \frac{1}{2}\big(1 - F_U(-x)\big) \quad (\star) \end{aligned}$$

La fonction de répartition de $U \hookrightarrow \mathcal{E}(1)$ est :

$$F_U : \begin{cases} \mathbb{R} & \to & \mathbb{R} \\ y & \mapsto & \begin{cases} 0 & \text{si } y \leq 0 \\ 1 - e^{-y} & \text{si } y \geq 0 \end{cases} \end{cases}$$

En remplaçant dans $(\star)$:

- pour $x \leq 0$:

$$F_X(x) = 0 + \frac{1}{2}\left(1 - e^{-(-x)}\right) = \frac{1}{2}e^x;$$

- pour $x \geq 0$:
$$F_X(x) = \frac{1}{2}\left(1 - e^{-x}\right) + \frac{1}{2}(1-0) = 1 - \frac{1}{2}e^{-x}.$$

On retrouve la fonction de répartition Ψ déterminée en question 2. Par conséquent, X suit la loi $\mathcal{L}(0,1)$.

(b) Voici le programme demandé :

```
def Laplace(alpha,beta):
  if rd.random() <= 1/2:
    V = 1
  else:
    V = 0
  X = (2*V-1)*rd.exponential(1)
  return(beta*X+alpha)
```

6. (a) Si pour tout intervalle I de $\mathbb{R}$ on a :
$$e^{-\varepsilon}P(X \in I) \underset{(1)}{\leq} P(Y \in I) \underset{(2)}{\leq} e^{\varepsilon}P(X \in I)$$
alors,
$$e^{-\varepsilon}P(Y \in I) \underset{(2)}{\leq} P(X \in I) \underset{(1)}{\leq} e^{\varepsilon}P(Y \in I).$$
En conclusion, si (X,Y) est ε-différentiel alors (Y,X) l'est aussi.

(b) Supposons que (X,Y) est ε-différentiel et que (Y,Z) est ε'-différentiel.
Pour tout intervalle I de $\mathbb{R}$:
$$\begin{array}{ccccc} e^{-\varepsilon}P(X \in I) & \underset{(1)}{\leq} & P(Y \in I) & \underset{(2)}{\leq} & e^{\varepsilon}P(X \in I) \\ e^{-\varepsilon'}P(Y \in I) & \underset{(3)}{\leq} & P(Z \in I) & \underset{(4)}{\leq} & e^{\varepsilon'}P(Y \in I) \end{array}$$
Les inégalités (1) et (3) donnent :
$$e^{-\varepsilon'}e^{-\varepsilon}P(X \in I) \leq e^{-\varepsilon'}P(Y \in I) \leq P(Z \in I)$$
donc
$$e^{-\varepsilon'-\varepsilon}P(X \in I) \leq P(Z \in I) \quad (\star).$$
Les inégalités (4) et (2) donnent :
$$P(Z \in I) \leq e^{\varepsilon'}P(Y \in I) \leq e^{\varepsilon'}e^{\varepsilon}P(X \in I)$$
donc
$$P(Z \in I) \leq e^{\varepsilon'+\varepsilon}P(X \in I) \quad (\star\star).$$
En résumé ($(\star)$ et $(\star\star)$) :
$$e^{-\varepsilon'-\varepsilon}P(X \in I) \leq P(Z \in I) \leq e^{\varepsilon'+\varepsilon}P(X \in I).$$
Le couple (X,Z) est donc $(\varepsilon + \varepsilon')$-différentiel.

7. • Si le couple (X,Y) est ε-différentiel, il suffit d'appliquer, pour chaque $n \in J$, la définition à l'intervalle $I = [z_n, z_n] = \{z_n\}$ pour obtenir :
$$\forall n \in J,\ e^{-\varepsilon}P(X = z_n) \leq P(Y = z_n) \leq e^{\varepsilon}P(X = z_n).$$

- Réciproquement, supposons que :

$$\forall n \in J,\ e^{-\varepsilon}P(X = z_n) \leq P(Y = z_n) \leq e^{\varepsilon}P(X = z_n).$$

Soit I un intervalle de $\mathbb{R}$. Si I ne contient aucun z_n $(n \in J)$, alors les probabilités $P(X \in I)$ et $P(Y \in I)$ sont nulles et on a donc bien évidemment les inégalités souhaitées.
Sinon, notons K le sous-ensemble de J tel que $\big(X(\Omega) \cup Y(\Omega)\big) \cap I = \{z_n | n \in K\}$. Il suffit alors de sommer les inégalités :

$$e^{-\varepsilon}P(X = z_n) \leq P(Y = z_n) \leq e^{\varepsilon}P(X = z_n)$$

pour tout $n \in K$; on obtient :

$$e^{-\varepsilon}\underbrace{\sum_{n\in K} P(X = z_n)}_{=P(X\in I)} \leq \underbrace{\sum_{n\in K} P(Y = z_n)}_{=P(Y\in I)} \leq e^{\varepsilon}\underbrace{\sum_{n\in K} P(X = z_n)}_{=P(X\in I)}$$

et le couple (X, Y) est donc ε-différentiel.

8. (a) Clairement, $Y(\Omega) = \mathbb{N}^*$. Déterminons $P(Y = k)$:
   - pour $k = 1$, on a $(Y = 1) = (X = 1) \cap (Z = 0)$ et par indépendance de X et Z :

   $$P(Y = 1) = P\big((X = 1) \cap (Z = 0)\big) = P(X = 1) \times P(Z = 0) = \frac{1-p}{2}.$$

   - pour $k \geq 2$, on a $(Y = k) = \big((X = k) \cap (Z = 0)\big) \cup \big((X = k-1) \cap (Z = 1)\big)$, l'union étant disjointe puis par indépendance de X et Z :

   $$P(Y = k) = P(X = k)P(Z = 0) + P(X = k-1)P(Z = 1) = \frac{1-p}{2^k} + \frac{p}{2^{k-1}} = \frac{1+p}{2^k}.$$

   On peut bien entendu vérifier que :

   $$\sum_{k=1}^{+\infty} P(Y = k) = \frac{1-p}{2} + (1+p)\sum_{k=2}^{+\infty}\frac{1}{2^k} = \frac{1-p}{2} + (1+p) \times \frac{1}{4} \times 2 = \frac{1-p+1+p}{2} = 1.$$

   (b) Pour $k = 1$:

   $$\frac{P(Y = 1)}{P(X = 1)} = \frac{\frac{1-p}{2}}{\frac{1}{2}} = 1 - p \in \left[1-p, \frac{1}{1-p}\right].$$

   Pour $k \geq 2$:

   $$\frac{P(Y = k)}{P(X = k)} = \frac{\frac{1+p}{2^k}}{\frac{1}{2^k}} = 1 + p$$

   Or :

   $$1 - p \leq 1 + p \text{ (c'est clair)} \quad \text{et} \quad 1 + p \leq \frac{1}{1-p} \text{ car } (1+p)(1-p) = 1 - p^2 \leq 1.$$

   On a donc bien :

   $$\forall k \in \mathbb{N}^*,\ 1 - p \leq \frac{P(Y = k)}{P(X = k)} \leq \frac{1}{1-p}.$$

   (c) On peut réécrire le résultat ci-dessus ainsi :

   $$\forall k \in \mathbb{N}^*,\ e^{-\left(-\ln(1-p)\right)} \leq \frac{P(Y = k)}{P(X = k)} \leq e^{-\ln(1-p)}$$

   et on conclut avec la caractérisation de la propriété ε-différentiel pour les couples de variables aléatoires discrètes (question 7.) que (X, Y) est $-\ln(1-p)$-différentiel.

(d) Lorsque p tend vers 0 : par encadrement dans la relation obtenue en 8.b), $\dfrac{P([Y=k])}{P([X=k])}$ se rapproche de 1, c'est-à-dire que les lois de X et Y sont très proches l'une de l'autre. C'est cohérent avec le fait que la variable de Bernoulli Z a alors une probabilité très forte de prendre la valeur 0, donc Y de prendre la même valeur que X.

Si p se rapproche de 1 : alors Z prend très probablement la valeur 1, et Y est plus proche de $X+1$ que de X, ce qui est cohérent avec le fait que $-\ln(1-p)$ devient d'autant plus grand que p est proche de 1.

9. (a) Soit I un intervalle de $\mathbb{R}$ de bornes a et b, avec a réel ou $-\infty$, b réel ou $+\infty$, et $a<b$.
La croissance de l'intégrale appliquée sur I aux inégalités :

$$\forall t \in I,\ e^{-\varepsilon} f(t) \leq g(t) \leq e^{\varepsilon} f(t)$$

s'écrit :

$$e^{-\varepsilon} \underbrace{\int_a^b f(t)dt}_{=P(X\in I)} \leq \underbrace{\int_a^b g(t)dt}_{=P(Y\in I)} \leq e^{\varepsilon} \underbrace{\int_a^b f(t)dt}_{=P(X\in I)}.$$

Donc (X,Y) est ε-différentiel.

(b) Choisissons l'intervalle $I=[t,t+h]$ dans la définition de (X,Y) est ε-différentiel :

$$e^{-\varepsilon} P(X\in[t,t+h]) \leq P(Y\in[t,t+h]) \leq e^{\varepsilon} P(X\in[t,t+h])$$

c'est-à-dire (en divisant par h (>0)):

$$e^{-\varepsilon} \frac{F(t+h)-F(t)}{h} \leq \frac{G(t+h)-G(t)}{h} \leq e^{\varepsilon} \frac{F(t+h)-F(t)}{h}$$

Par passage à la limite quand h tend vers 0, les fonctions F et G étant dérivables (et même de classe $\mathcal{C}^1$) en t car f et g sont continues en t :

$$e^{-\varepsilon} F'(t) \leq G'(t) \leq e^{\varepsilon} F'(t)$$

c'est-à-dire :

$$e^{-\varepsilon} f(t) \leq g(t) \leq e^{\varepsilon} f(t).$$

10. (a) La fonction $t \mapsto \dfrac{1}{1+t^2}$ étant paire, il suffit d'en montrer l'intégrabilité sur $[0,+\infty[$:

- elle est continue sur $[0,+\infty[$ donc l'intégrale est impropre uniquement en $+\infty$;
- au voisinage de $+\infty$, elle est équivalente à $\dfrac{1}{t^2}$ dont l'intégrale (de Riemann) converge.

Par le théorème de comparaison sur des intégrales de fonctions positives, on en déduit que $\displaystyle\int_0^{+\infty} \frac{1}{1+t^2}dt$ et donc $\displaystyle\int_{-\infty}^{+\infty} \frac{1}{1+t^2}dt$ convergent.

(b) La fonction f_a est clairement continue et positive. Elle est intégrable avec la même explication qu'en (a) et avec le changement de variable affine $u=\dfrac{t}{a}$, on a :

$$\int_{-\infty}^{+\infty} f_a(t)dt = \frac{1}{a\pi}\int_{-\infty}^{+\infty} \frac{1}{\left(\frac{t}{a}\right)^2+1}dt = \frac{1}{\cancel{a}\pi}\int_{-\infty}^{+\infty} \frac{1}{u^2+1}\cancel{a}du = \frac{1}{\pi}\times\pi = 1.$$

Par conséquent, f_a est une densité de probabilité.

(c) D'après la question 9.(a), il suffit de prouver que :

$$\forall t \in \mathbb{R},\ e^{-\ln(a)} f(t) \leq g(t) \leq e^{\ln(a)} f(t)$$

c'est-à-dire (en simplifiant par π) :

$$\forall t \in \mathbb{R},\ \frac{1}{a} \times \frac{1}{t^2+1} \leq \frac{a}{t^2+a^2} \leq a \times \frac{1}{t^2+1}.$$

La seconde inégalité est claire ($t^2+a^2 \geq t^2+1$) ; détaillons la première :

$$\begin{aligned} \frac{1}{a} \times \frac{1}{t^2+1} \leq \frac{a}{t^2+a^2} \quad & \Leftrightarrow \quad \frac{t^2+a^2}{t^2+1} \leq a^2 \\ & \Leftrightarrow \quad \frac{t^2+a^2}{t^2+1} - a^2 \leq 0 \\ & \Leftrightarrow \quad \frac{t^2+a^2-a^2t^2-a^2}{t^2+1} \leq 0 \\ & \Leftrightarrow \quad \underbrace{\frac{(1-a^2)t^2}{t^2+1} \leq 0}_{\text{ce qui est vrai } (1-a^2<0)} \quad . \end{aligned}$$

En conclusion, (X,Y) est $\ln(a)$-différentiel.

11. (a) On a, en utilisant au dénominateur la formule des probabilités totales avec le système complet d'événements $\{(U=1),(U=0)\}$:

$$\begin{aligned} P_{(Z\in I)}(U=1) &= \frac{P\big((Z\in I)\cap(U=1)\big)}{P(Z\in I)} \\ &= \frac{P\big((X\in I)\cap(U=1)\big)}{P\big((Z\in I)\cap(U=1)\big)+P\big((Z\in I)\cap(U=0)\big)} \\ &= \frac{P\big((X\in I)\cap(U=1)\big)}{P\big((X\in I)\cap(U=1)\big)+P\big((Y\in I)\cap(U=0)\big)}. \end{aligned}$$

Par indépendance de U et X d'une part, U et Y d'autre part :

$$\begin{aligned} P_{(Z\in I)}(U=1) &= \frac{P(X\in I)P(U=1)}{P(X\in I)P(U=1)+P(Y\in I)P(U=0)} \\ &= p\frac{P(X\in I)}{pP(X\in I)+(1-p)P(Y\in I)} \end{aligned}$$

Puisque (X,Y) est ε-différentiel :

$$\begin{aligned} & e^{-\varepsilon}P(X\in I) \leq P(Y\in I) \leq e^{\varepsilon}P(X\in I) \\ \Leftrightarrow\ & (1-p)e^{-\varepsilon}P(X\in I) \leq (1-p)P(Y\in I) \leq (1-p)e^{\varepsilon}P(X\in I) \quad (\text{car } 1-p>0) \\ \Leftrightarrow\ & (p+(1-p)e^{-\varepsilon})P(X\in I) \leq pP(X\in I)+(1-p)P(Y\in I) \leq (p+(1-p)e^{\varepsilon})P(X\in I) \\ \Leftrightarrow\ & \frac{1}{(p+(1-p)e^{-\varepsilon})P(X\in I)} \geq \frac{1}{pP(X\in I)+(1-p)P(Y\in I)} \geq \frac{1}{(p+(1-p)e^{\varepsilon})P(X\in I)} \\ \Leftrightarrow\ & \frac{p}{p+(1-p)e^{-\varepsilon}} \geq p\frac{P(X\in I)}{pP(X\in I)+(1-p)P(Y\in I)} \geq \frac{p}{p+(1-p)e^{\varepsilon}} \quad (\text{car } pP(X\in I)\geq 0) \\ \Leftrightarrow\ & \frac{p}{p+(1-p)e^{-\varepsilon}} \geq P_{(Z\in I)}(U=1) \geq \frac{p}{p+(1-p)e^{\varepsilon}} \end{aligned}$$

(b) Si ε est proche de 0, alors e^{ε} et $e^{-\varepsilon}$ sont proches de 1 et les deux nombres

$$\frac{p}{p+(1-p)e^{\varepsilon}} \quad \text{et} \quad \frac{p}{p+(1-p)e^{-\varepsilon}}$$

sont donc proches de p ; avec l'encadrement établi en (a), la probabilité conditionnelle $P_{(Z\in I)}(U=1)$ est donc proche de la probabilité $p = P(U=1)$.
Par conséquent, la connaissance d'une information sur Z lorsque ε est petit change peu la probabilité d'en déduire la valeur prise par U.

12. (a) En choisissant une valeur au hasard dans $[\![0,d]\!]$, la probabilité d'obtenir a_1 vaut $\dfrac{1}{d+1}$ (l'ensemble D contient $d+1$ éléments).

(b) Si $q(a)$ est publique, alors on obtient a_1 ainsi :

$$a_1 = q(a) - a_2 - \ldots - a_n.$$

13. (a) En écrivant :

$$(R = A_1) = \bigcup_{j=0}^{d} \left((R=j)\cap(A_1=j)\right)$$

l'union étant disjointe, on a :

$$P(R=A_1) = \sum_{j=0}^{d} P\big((R=j)\cap(A_1=j)\big).$$

Étant donnée la définition de R :

$$\big((R=j)\cap(A_1=j)\big) = \big((Y_j\in I_j)\cap(A_1=j)\big)$$

d'où l'égalité demandée.

(b) Par indépendance de Y_j et A_1 pour tout $j\in D$:

$$P(R=A_1) = \sum_{j=0}^{d} P(Y_j\in I_j)\times \underbrace{P(A_1=j)}_{=\frac{1}{d+1}\text{ d'après 12.(a)}} = \frac{1}{d+1}\sum_{j=0}^{d} P(Y_j\in I_j).$$

(c) Pour tout $j\in[\![1,d]\!]$, les n-uplets $(j,a_2,\ldots,a_n)$ et $(0,a_2,\ldots,a_n)$ étant voisins (seule la première composante diffère), d'après l'hypothèse (c2) de l'énoncé le couple (Y_j,Y_0) est ε-différentiel :

$$\text{pour tout } I \text{ intervalle de } \mathbb{R},\quad \underbrace{e^{-\varepsilon}P(Y_j\in I) \le P(Y_0\in I)}_{(\star)} \le e^{\varepsilon}P(Y_j\in I).$$

Écrivons pour tout $j\in[\![1,d]\!]$, l'inégalité $(\star)$ avec l'intervalle $I=I_j$:

$$\forall j\in[\![1,d]\!],\ e^{-\varepsilon}P(Y_j\in I_j)\le P(Y_0\in I_j)$$

et sommons ces inégalités membre à membre pour $j\in\{1,\ldots,d\}$:

$$\sum_{j=1}^{d} e^{-\varepsilon}P(Y_j\in I_j) \le \sum_{j=1}^{d} P(Y_0\in I_j)$$

soit encore, en multipliant par e^{ε} et en ajoutant $P(Y_0\in I_0)$:

$$\sum_{j=0}^{d} P(Y_j\in I_j) \le \left(e^{\varepsilon}\underbrace{\sum_{j=1}^{d} P(Y_0\in I_j)}_{=1-P(Y_0\in I_0)}\right) + P(Y_0\in I_0).$$

On en déduit :

$$\underbrace{\frac{1}{d+1}\sum_{j=0}^{d} P(Y_j \in I_j)}_{=\theta} \leq \frac{1}{d+1}\big(e^{\varepsilon} - (e^{\varepsilon}-1)P(Y_0 \in I_0)\big).$$

La deuxième inégalité demandé est claire car $-(e^{\varepsilon}-1)P(Y_0 \in I_0) < 0$ ($\varepsilon > 0$). En conclusion :

$$\theta \leq \frac{1}{d+1}\big(e^{\varepsilon} - (e^{\varepsilon}-1)P(Y_0 \in I_0)\big) \leq \frac{e^{\varepsilon}}{d+1}.$$

(d) On déduit de l'inégalité $\theta \leq \dfrac{e^{\varepsilon}}{d+1}$ que :

$$\tau = \frac{\theta-\rho}{\rho} \leq \frac{\dfrac{e^{\varepsilon}}{d+1} - \dfrac{1}{d+1}}{\dfrac{1}{d+1}} = e^{\varepsilon} - 1.$$

Lorsque ε est proche de 0, ce majorant est proche de 0, c'est-à-dire que la probabilité $P(R = A_1)$ est majorée par un nombre proche de $\dfrac{1}{d+1}$, qui est la probabilité d'obtenir la bonne valeur de a_1 en choisissant un entier au hasard de D. La méthode de confidentialité n'augmentera donc que peu (d'autant peu que ε est proche de 0) la probabilité d'obtenir la bonne valeur de a par rapport à un choix aléatoire.

14. (a) D'après 3.(a), la variable X_a suit la loi de Laplace $\mathcal{L}(q(a), \beta)$.
Donc $E(X_a) = q(a)$ et une densité de X_a est donnée par :

$$f_a : \begin{cases} \mathbb{R} & \to & \mathbb{R} \\ t & \mapsto & \dfrac{1}{2\beta}\exp\left(-\dfrac{|t-q(a)|}{\beta}\right) \end{cases}$$

(b) D'après l'inégalité triangulaire renversée, on a pour tout $(a,b) \in \mathcal{V}$ et $t \in \mathbb{R}$:

$$|t-q(b)| - |t-q(a)| \leq \big|\big(t-q(a)\big) - \big(t-q(b)\big)\big| = |q(b)-q(a)| \leq \delta$$

donc :

$$-|t-q(a)| \leq \delta - |t-q(b)| \quad \text{d'où} \quad -\frac{|t-q(a)|}{\beta} \leq \frac{\delta}{\beta} - \frac{|t-q(b)|}{\beta}.$$

Par croissance de l'exponentielle :

$$\exp\left(-\frac{|t-q(a)|}{\beta}\right) \leq \exp\left(\frac{\delta}{\beta}\right)\exp\left(-\frac{|t-q(b)|}{\beta}\right)$$

et en multipliant par $\dfrac{1}{2\beta}$ (> 0) :

$$\forall t \in \mathbb{R},\ \forall (a,b) \in \mathcal{V},\ f_a(t) \leq \exp\left(\frac{\delta}{\beta}\right) f_b(t).$$

Pour tous $t \in \mathbb{R}$ et $(a,b) \in \mathcal{V}$, on a donc :

$$f_a(t) \leq \exp\left(\frac{\delta}{\beta}\right) f_b(t) \quad \text{et} \quad f_b(t) \leq \exp\left(\frac{\delta}{\beta}\right) f_a(t)$$

soit :

$$\exp\left(-\frac{\delta}{\beta}\right) f_b(t) \leq f_a(t) \leq \exp\left(\frac{\delta}{\beta}\right) f_b(t)$$

ce qui prouve, avec la question 9., que le couple (X_a, X_b) est $\dfrac{\delta}{\beta}$-différentiel.

(c) Pour tout $a \in D$, on dispose d'une variable aléatoire X_a d'espérance $q(a)$ (propriétés (c1) et (c3)).
Si $\varepsilon = \dfrac{\delta}{\beta}$, alors pour tout $(a, b) \in \mathcal{V}$, le couple (X_a, X_b) est ε-différentiel (propriété (c2)).

On obtient donc un procédé de ε-confidentialité de D^n pour q en prenant par exemple $\beta = \dfrac{\delta}{\varepsilon}$.

15. (a) Soient a et b éléments voisins de D, c'est-à-dire que a et b différent d'au plus une composante :

$$\exists i_0 \in [\![1, n]\!],\ \forall i \in [\![1, n]\!] \setminus \{i_0\},\ a_i = b_i.$$

Alors :

$$q(a) - q(b) = \sum_{i=1}^{n} a_i - \sum_{i=1}^{n} b_i = a_{i_0} - b_{i_0} \in [\![-d, d]\!]$$

(différence de deux entiers compris entre 0 et d).

On a donc l'inégalité :

$$|q(a) - q(b)| \leq d$$

qui est atteinte (pour $a_{i_0} = 0$ et $b_{i_0} = d$ par exemple).
On en déduit :

$$\delta = \max_{(a,b)\in\mathcal{V}} \big(|q(a) - q(b)|\big) = d.$$

(b) Si $X_a \in \left[k - \dfrac{1}{2}, k + \dfrac{1}{2}\right[$ avec $k \in [\![1, nd - 1]\!]$, alors $X_a + \dfrac{1}{2} \in [k, k+1[$ donc :

$$\left\lfloor X_a + \frac{1}{2} \right\rfloor = k = Z_a.$$

Les deux autres cas sont évidents.

(c) Les question 14. et 15.(a) nous donnent :

$$\forall a \in D,\ X_a = q(a) + Y \quad \text{avec} \quad Y \hookrightarrow \mathcal{L}\left(0, \frac{d}{\varepsilon}\right).$$

On utilise bien entendu la question 5. pour générer Y :

```
d = input('d = ')
n = input('n = ')
eps = input('epsilon = ')

a = rd.randint(0, d+1, n)
print('q(a) = ', np.sum(a))
X = np.sum(a)+Laplace(0,d/eps)

if X < 1/2:
    Z = 0
elif X < n*d-1/2:
    Z = np.floor(X+1/2)
else:
    Z = n*d

print('Z_a = ',Z)
```

(d) On reprend le script précédent, sans les affichages de $q(a)$ et Z_a, en calculant la moyenne des nombres $\dfrac{|Z_a - q(a)|}{q(a)}$ pour 10000 (par exemple) valeurs de a générées aléatoirement.

```python
d = 4
n = 1000
eps = input('epsilon = ')

M = 0
for i in range(10000):
    a = rd.randint(0, d+1, n)
    X = np.sum(a)+Laplace(0,d/eps)

    if X < 1/2:
        Z = 0
    elif X < n*d-1/2:
        Z = np.floor(X+1/2)
    else:
        Z = n*d
    M = M + np.abs(Z-np.sum(a))/np.sum(a)
print(M/10000)
```