
DS6 (vB) ESSEC I 2017 - Barème

Est-il possible que le marketing digital pose des problèmes de sécurité des données personnelles ? De récents travaux mettant en cause les outils de mesure de performance en temps réel des différentes campagnes de publicité sur internet, démontrent que certaines données très sensibles (préférences religieuses, sexuelles, etc.) peuvent être obtenues par des segmentations précises des audiences et sans aucune action de la part de l'utilisateur.

Dans ce problème, nous nous intéressons à une méthode proposée pour protéger ces données, méthode baptisée **confidentialité différentielle**.

Les parties I et II sont totalement indépendantes. Vous trouverez une aide **Python** en fin de sujet.

On considère un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ sur lequel sont définies les variables aléatoires qui apparaissent dans l'énoncé.

Partie I - Lois de Laplace - propriétés et simulation

Soit $\alpha \in \mathbb{R}$ et $\beta > 0$. On dit qu'une variable aléatoire réelle à densité suit une loi de Laplace de paramètre (α, β) , notée $\mathcal{L}(\alpha, \beta)$, si elle admet comme densité la fonction f donnée par :

$$\forall t \in \mathbb{R}, f(t) = \frac{1}{2\beta} \exp\left(-\frac{|t - \alpha|}{\beta}\right)$$

1. Vérifier que f est bien une densité de probabilité d'une variable aléatoire réelle.

- 1 pt : f continue

- 1 pt : f positive

- 2 pts : qq soit la méthode

Découpage possible : $\int_{\alpha}^{+\infty} f(t) dt = \frac{1}{2}$ (1 point) et $\int_{-\infty}^{\alpha} f(t) dt = \frac{1}{2}$ (1 point)

2. Déterminer la fonction de répartition, notée Ψ , de la loi $\mathcal{L}(0, 1)$.

- 1 pt : cas $x \leq 0$ ($\Psi(x) = \frac{1}{2}e^x$)

- 1 pt : cas $x > 0$ ($\Psi(x) = 1 - \frac{1}{2}e^{-x}$)

3. On suppose que X suit la loi $\mathcal{L}(0, 1)$.

a) Montrer que $Y = \beta X + \alpha$ suit la loi $\mathcal{L}(\alpha, \beta)$.

- 1 pt : $F_Y(x) = \Psi\left(\frac{x - \alpha}{\beta}\right)$

- 1 pt : expression de $F_Y(x) = \begin{cases} \frac{1}{2} \exp\left(\frac{x - \alpha}{\beta}\right) & \text{si } x \leq \alpha \\ 1 - \frac{1}{2} \exp\left(\frac{-x + \alpha}{\beta}\right) & \text{si } x > \alpha \end{cases}$

- 1 pt : Y est à densité

- 1 pt : $f_Y = F'_Y = f$ (pas de point si dérivation en α)

b) En déduire la fonction de répartition de la loi $\mathcal{L}(\alpha, \beta)$.

- 1 pt : expression de F trouvée dans la question précédente

4. *Espérance et variance.*

a) On suppose que X suit la loi $\mathcal{L}(0, 1)$.

Montrer que $\mathbb{E}(X)$ et $\mathbb{V}(X)$ existent et valent respectivement 0 et 2.

- 1 pt : X admet une esp $\Leftrightarrow \int_{-\infty}^{+\infty} te^{-|t|} dt$ (absolument) convergente

- 1 pt : $\int_0^{+\infty} te^{-|t|} dt = - \int_{-\infty}^0 te^{-|t|} dt$ car f impaire

- 1 pt : $\int_0^{+\infty} te^{-t} dt$ est convergente (on reconnaît l'espérance de $Z \hookrightarrow \mathcal{E}(1)$)

- 1 pt : $\mathbb{E}(X) = 0$

- 1 pt : $\int_{-\infty}^{+\infty} t^2 e^{-|t|} dt = 2 \int_0^{+\infty} t^2 e^{-t} dt$ (parité)

- 1 pt : on reconnaît le moment d'ordre 2 de $Z \hookrightarrow \mathcal{E}(1)$

- 1 pt : $E(Z^2) = 2$ par la formule de Kœnig-Huygens

- 1 pt : $\mathbb{E}(X^2) = 2$ et conclusion $\mathbb{V}(X^2) = 2$.

b) En déduire l'existence et les valeurs de l'espérance et de la variance d'une variable aléatoire réelle qui suit la loi $\mathcal{L}(\alpha, \beta)$.

- 1 pt : existence par transformation affine

- 1 pt : $\mathbb{E}(Y) = \alpha$ (linéarité de l'espérance)

- 1 pt $\mathbb{V}(Y) = 2\beta^2$ (propriété de la variance)

5. *Simulation à partir d'une loi exponentielle.*

Soit U une variable aléatoire qui suit la loi exponentielle de paramètre 1 et V une variable aléatoire qui suit la loi de Bernoulli de paramètre $\frac{1}{2}$ et indépendante de U .

a) En utilisant le système complet naturellement associé à V , montrer que $X = (2V - 1)U$ suit la loi $\mathcal{L}(0, 1)$.

- 1 pt : formule des proba totales ($\mathbb{P}([X \leq x]) = \mathbb{P}([V = 1] \cap [X \leq x]) + \mathbb{P}([V = 0] \cap [X \leq x])$)

- 1 pt : $= \mathbb{P}([V = 1] \cap [U \leq x]) + \mathbb{P}([V = 0] \cap [-U \leq x])$

- 1 pt : indépendance de U et V

- 1 pt : $F_X(x) = \frac{1}{2} F_U(x) + \frac{1}{2}(1 - F_U(-x))$

- 1 pt : $F_X : x \mapsto \begin{cases} \frac{1}{2} e^x & \text{si } x \leq 0 \\ 1 - \frac{1}{2} e^{-x} & \text{si } x > 0 \end{cases}$

et on reconnaît $\mathcal{L}(0, 1)$

b) Compléter la définition **Python** ci-dessous pour que la fonction ainsi définie réalise la simulation d'une variable aléatoire qui suit la loi $\mathcal{L}(\alpha, \beta)$:

```
1 import numpy.random as rd
2 def Laplace(alpha, beta):
3     if rd.random() < 1/2:
4         V = 1
5     else:
6         V = 0
7     X = (2*V - 1) * rd.exponential(1)
8     return beta * X + alpha
```

```
- 1 pt : import  numpy.random as rd 
- 1 pt : if  rd.random()  < 1/2:
- 1 pt : return  beta * X + alpha 
```

Partie II - Lois ε -différentielles

Soit $\varepsilon > 0$. On dit que (X, Y) , un couple de variables aléatoires, est un couple ε -différentiel si, pour tout intervalle I de \mathbb{R} :

$$e^{-\varepsilon} \mathbb{P}([X \in I]) \leq \mathbb{P}([Y \in I]) \leq e^{\varepsilon} \mathbb{P}([X \in I])$$

Intuitivement, les lois de X et Y seront d'autant plus proches que le plus petit ε tel que (X, Y) soit un couple ε -différentiel est proche de 0.

6. Soit (X, Y, Z) un triplet de variables aléatoires réelles.

a) Montrer que si (X, Y) est ε -différentiel alors (Y, X) l'est aussi.

- 1 pt : multiplication par $e^{\varepsilon} > 0$ à gauche

- 1 pt : multiplication par $e^{-\varepsilon} > 0$ à droite

b) Montrer que si (X, Y) est ε -différentiel et (Y, Z) est ε' -différentiel alors (X, Z) est $(\varepsilon + \varepsilon')$ -différentiel.

- 1 pt : $e^{-\varepsilon} \mathbb{P}([X \in I]) \leq \mathbb{P}([Y \in I])$ et $e^{-\varepsilon'} \mathbb{P}([Y \in I]) \leq \mathbb{P}([Z \in I])$ implique $e^{-(\varepsilon+\varepsilon')} \mathbb{P}([X \in I]) \leq \mathbb{P}([Z \in I])$

- 1 pt : $\mathbb{P}([Y \in I]) \leq e^{\varepsilon} \mathbb{P}([X \in I])$ et $\mathbb{P}([Z \in I]) \leq e^{\varepsilon'} \mathbb{P}([Y \in I])$ implique $\mathbb{P}([Z \in I]) \leq e^{\varepsilon+\varepsilon'} \mathbb{P}([X \in I])$

7. Soit (X, Y) un couple de variables aléatoires réelles discrètes.

On suppose que $X(\Omega) \cup Y(\Omega) = \{z_n \mid n \in J\}$ où J est un sous ensemble non vide de \mathbb{N} .

Montrer que (X, Y) est ε -différentiel si et seulement si

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

- 1 pt : (\Rightarrow) (appliquer ε -différentiabilité à $I = \{z_n\}$)

- 3 pts : (\Leftarrow)

8. *Premier exemple.*

Dans cette question, on suppose que X suit la loi géométrique de paramètre $\frac{1}{2}$, Z suit la loi de Bernoulli de paramètre $p \in]0, 1[$ et elles sont indépendantes. On pose $Y = X + Z$.

a) Déterminer la loi de Y .

- 1 pt : $Y(\Omega) = \mathbb{N}^*$

- 1 pt : **FPT** $\mathbb{P}([Y = k]) = (1 - p) \mathbb{P}([X = k]) + p \mathbb{P}([X = k - 1])$

- 1 pt : **cas** $k = 1$ ($\mathbb{P}([Y = 1]) = \frac{1-p}{2}$)

- 1 pt : **cas** $k \geq 2$: $\mathbb{P}([Y = k]) = \frac{1+p}{2^k}$

b) Établir que pour tout $k \in \mathbb{N}^*$, $1 - p \leq \frac{\mathbb{P}([Y = k])}{\mathbb{P}([X = k])} \leq \frac{1}{1 - p}$.

- 2 pts : **cas** $k = 1$ ($1 - p \leq \frac{\mathbb{P}([Y = 1])}{\mathbb{P}([X = 1])} = 1 - p \leq \frac{1}{1 - p}$)

- **2 pts** : cas $k \geq 2$ ($\frac{\mathbb{P}([Y = k])}{\mathbb{P}([X = k])} = 1 + p$)

c) En déduire que (X, Y) est $-\ln(1 - p)$ -différentiel.

- **1 pt** : d'après qst précédente $e^{\ln(1-p)} \leq \frac{\mathbb{P}([Y = k])}{\mathbb{P}([X = k])} \leq e^{-\ln(1-p)}$

d) Que se passe-t-il lorsque p s'approche de 0 ou lorsqu'il s'approche de 1 ? Était-ce prévisible ?

- **1 pt** : si p s'approche de 0, alors la probabilité $\mathbb{P}([X = Y]) = \mathbb{P}([Z = 0]) = 1 - p$ se rapproche de 1.

Les lois de X et Y sont donc d'autant plus proche que p s'approche de 0.

- **1 pt** : Lorsque p s'approche de 1, $-\ln(1 - p)$ s'approche de $+\infty$.

9. On suppose que X et Y sont deux variables à densité de densités respectives f et g et de fonction de répartition F et G .

a) On suppose que pour tout $t \in \mathbb{R}$: $e^{-\varepsilon} f(t) \leq g(t) \leq e^{\varepsilon} f(t)$.

Montrer que (X, Y) est ε -différentiel.

- **1 pt** : croissance de l'intégrale (bornes bien ordonnées)

b) On suppose dans la suite de cette question que (X, Y) est ε -différentiel.

Soit $h > 0$ et $t \in \mathbb{R}$ où f et g sont continues.

Montrer que :

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

En conclure que : $e^{-\varepsilon} f(t) \leq g(t) \leq e^{\varepsilon} f(t)$.

- **1 pt** : appliquer définition ε -différentiabilité à $I = [t, t+h]$

- **1 pt** : F et G de classe \mathcal{C}^1 car f et g continues

- **1 pt** : passage à la limite quand h tend vers 0

10. Deuxième exemple : lois de Cauchy.

a) Montrer que $\int_{-\infty}^{+\infty} \frac{1}{t^2 + 1} dt$ converge. On admet que cette intégrale est égale à π .

- **1 pt** : $t \mapsto \frac{1}{1+t^2}$ continue sur \mathbb{R} , donc $\int_{-\infty}^{+\infty} \frac{1}{1+t^2} dt$ est impropre en $+\infty$ et $-\infty$

- **2 pts** : $\int_0^{+\infty} \frac{1}{1+t^2} dt$ converge (1pt pour équivalent + 1 pt pour le critère avec intégrale de Riemann impropre en $+\infty$ d'exposant 2, donc convergente)

- **1 pt** : $\int_{-\infty}^0 \frac{1}{1+t^2} dt = \int_0^{+\infty} \frac{1}{1+t^2} dt$ par parité

b) On définit, pour $a > 0$, la fonction f_a sur \mathbb{R} par, pour tout $t \in \mathbb{R}$, $f_a(t) = \frac{a}{\pi(t^2 + a^2)}$.

Montrer que f_a est une densité de probabilité d'une variable aléatoire à densité.

- **1 pt** : f_a continue sur \mathbb{R}

- **1 pt** : f_a positive

- **2 pts** : changement de variable $u = \frac{t}{a}$ (1 pt pour classe \mathcal{C}^1 , 1 pt pour reste)

c) On suppose que X et Y sont deux variables aléatoires admettant comme densités respectives f_1 et f_a avec $a > 1$.

Montrer que (X, Y) est $\ln(a)$ -différentiel.

- 1 pt : question 10 (on veut montrer : $e^{-\ln(a)} f(t) \leq g(t) \leq e^{\ln(a)} f(t)$)

- 1 pt : $\frac{1}{a} \frac{1}{t^2 + 1} \leq \frac{a}{t^2 + a^2} \Leftrightarrow 0 \leq (a^2 - 1)t^2$

- 1 pt : $\frac{a}{t^2 + a^2} \leq \frac{a}{t^2 + 1}$

11. Une première interprétation.

On suppose que (X, Y) est un couple ε -différentiel et que U est une variable de Bernoulli de paramètre $p \in]0, 1[$ indépendante de X et Y .

On définit la variable aléatoire Z par :

$$\forall \omega \in \Omega, Z(\omega) = \begin{cases} X(\omega) & \text{si } U(\omega) = 1 \\ Y(\omega) & \text{sinon.} \end{cases}$$

a) Soit I un intervalle de \mathbb{R} telle que $\mathbb{P}([Z \in I]) \neq 0$.

Montrer que :

$$\mathbb{P}_{[Z \in I]}([U = 1]) = p \frac{\mathbb{P}([X \in I])}{p \mathbb{P}([X \in I]) + (1 - p) \mathbb{P}([Y \in I])}$$

En déduire que :

$$\frac{p}{p + (1 - p) e^\varepsilon} \leq \mathbb{P}_{[Z \in I]}([U = 1]) \leq \frac{p}{p + (1 - p) e^{-\varepsilon}}$$

- 1 pt : proba conditionnelle $\mathbb{P}_{[Z \in I]}([U = 1]) = \frac{\mathbb{P}([Z \in I] \cap [U = 1])}{\mathbb{P}([Z \in I])}$

- 1 pt : $\mathbb{P}([Z \in I] \cap [U = 1]) = p \mathbb{P}([X \in I])$

- 2 pts : FPT $\mathbb{P}([Z \in I]) = p \mathbb{P}([X \in I]) + (1 - p) \mathbb{P}([Y \in I])$ dont indépendance

- 1 pt : conséquence

b) Si ε est proche de zéro, le fait de disposer d'une information sur la valeur de Z change-t-il notablement le paramètre de la loi de U et par conséquent la probabilité d'en déduire la valeur prise par U ?

- 1 pt : si ε proche de 0, alors $\frac{p}{p + (1 - p)e^\varepsilon}$ et $\frac{p}{p + (1 - p)e^{-\varepsilon}}$ proches de p

- 1 pt : alors $p = \mathbb{P}([U = 1])$ proche de $\mathbb{P}_{[Z \in I]}([U = 1])$

- 1 pt : connaissance sur Z change peu la proba de déduire la valeur prise par U

Partie III - Confidentialité différentielle

- Soit $d \in \mathbb{N}^*$. On considère $D = \llbracket 0, d \rrbracket$ et n un entier naturel plus grand que 2.
- On dira que deux éléments de D^n , a et b , sont voisins si ils ne diffèrent que d'une composante au plus. On note \mathcal{V} l'ensemble des couples de voisins.
- On considère q une application de D^n dans \mathbb{R} .

Concrètement, un élément de D^n représente une table d'une base de donnée et q une requête sur cette base. Étant donné $a = (a_1, \dots, a_n)$, on s'intéresse au problème de la confidentialité de certains des a_i lorsque les autres a_i sont connus, ainsi que D , q et $q(a)$.

12. Dans cette question on suppose que a_2, \dots, a_n sont connus et on cherche à protéger a_1 .

a) Quelle est probabilité d'obtenir la bonne valeur de a_1 si l'on choisit une valeur au hasard dans $\llbracket 0, d \rrbracket$?

- 1 pt : $\frac{1}{d+1}$

b) Dans cette question $q(a_1, \dots, a_n) = \sum_{i=1}^n a_i$.

Montrer que si $q(a)$ est publique alors on sait déterminer la valeur de a_1 .

- 1 pt : si $q(a)$ publique, alors $a_1 = q(a) - a_2 - \dots - a_n$

On dit que l'on dispose d'un procédé de ε -confidentialité de D^n pour q si :

(c1) pour tout $a \in D^n$, on dispose d'une variable aléatoire réelle X_a ;

(c2) pour tout $(a, b) \in \mathcal{V}$, (X_a, X_b) est ε -différentiel.

(c3) pour tout $a \in D^n$, $\mathbb{E}(X_a) = q(a)$.

13. Majoration de la probabilité de trouver a_1 .

Dans cette question, nous allons justifier en partie la terminologie.

On suppose à nouveau que a_2, \dots, a_n sont connus, que l'on cherche à protéger a_1 et que :

× le public connaît des d'intervalles I_0, \dots, I_d disjoints de réunion \mathbb{R} tels qu'avec les valeurs fixées de a_2, \dots, a_n , si $q(a) \in I_j$ alors $a_1 = j$. Cela signifie que si $q(a)$ est publique alors a_1 aussi.

× on dispose d'un procédé de ε -confidentialité de D^n pour q et que l'on rend X_a publique à la place de $q(a)$.

On considère alors que l'expérience aléatoire modélisée par $(\Omega, \mathcal{A}, \mathbb{P})$ comporte comme première étape le choix au hasard de a_1 dans $\llbracket 0, d \rrbracket$ et on définit :

× A_1 la variable aléatoire associée à ce choix ;

× pour tout $j \in \llbracket 0, d \rrbracket$, $Y_j = X_{(j, a_2, \dots, a_n)}$.

On suppose que A_1 et Y_j sont indépendantes pour tout $j \in D$.

× la variable aléatoire réelle R par :

$$\forall \omega \in \Omega, \text{ si } A_1(\omega) = j \text{ alors on détermine l'unique } k \text{ tel que } Y_j(\omega) \in I_k \text{ et on pose } R(\omega) = k$$

× $\theta = \mathbb{P}([R = A_1])$.

a) Montrer que $\theta = \sum_{j=0}^d \mathbb{P}([Y_j \in I_j] \cap [A_1 = j])$.

- 1 pt : formule des proba totales sur le SCE associé à A_1 , i.e. $([A_1 = i])_{i \in \llbracket 1, d \rrbracket}$

- 1 pt : $[R = i] \cap [A_1 = i] = [Y_i = i] \cap [A_1 = i]$

b) En déduire que $\theta = \frac{1}{d+1} \sum_{j=0}^d \mathbb{P}([Y_j \in I_j])$.

- 1 pt : indépendance Y_j et A_1

- 1 pt : $\mathbb{P}([A_1 = j]) = \frac{1}{d+1}$

c) En conclure que :

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

- 1 pt : (j, a_2, \dots, a_n) et $(0, a_2, \dots, a_n)$ voisins donc (Y_j, Y_0) est ε -différentiel

- 1 pt : pour $I = I_j$, on a $e^{-\varepsilon} \mathbb{P}([Y_j \in I_j]) \leq \mathbb{P}([Y_0 \in I_j])$

- **1 pt** : on somme sur $\llbracket 1, d \rrbracket$
- **1 pt** : $\frac{1}{d+1} \sum_{j=0}^d \mathbb{P}([Y_j \in I_j]) \leq \frac{1}{d+1} (\mathbf{e}^\varepsilon - (\mathbf{e}^\varepsilon - 1) \mathbb{P}([Y_0 \in I_0]))$
- **1 pt** : $-(\mathbf{e}^\varepsilon - 1) \mathbb{P}([Y_0 \in I_0]) < 0$

d) On pose $\rho = \frac{1}{d+1}$ et $\tau = \frac{\theta - \rho}{\rho}$.

Donner une majoration de τ . Que représente cette quantité ?

Qu'en déduire concernant la méthode de confidentialité présentée dans cette question lorsque ε est proche de 0 ?

- **1 pt** : $\tau \leq \mathbf{e}^\varepsilon - 1$
- **1 pt** : quand ε tend vers 0, la méthode de confidentialité n'augmente que peu la proba d'obtenir la bonne valeur de a par rapport à un choix aléatoire

On pose $\delta = \max_{(a,b) \in \mathcal{V}} |q(a) - q(b)|$ et on suppose que $\delta > 0$.

14. Dans cette question, pour tout $a \in D^n$, on pose $X_a = q(a) + Y$ où Y suit la loi de Laplace de paramètre $(0, \beta)$.

a) Pour tout $a \in D^n$, déterminer $\mathbb{E}(X_a)$ et une densité de probabilité f_a de la loi de X_a en fonction de $q(a)$ et de β .

- **1 pt** : d'après 3.a), $X_a \hookrightarrow \mathcal{L}(q(a), \beta)$

- **1 pt** : $\mathbb{E}(X_a) = q(a)$

- **1 pt** : $f_a(t) = \frac{1}{2\beta} \exp\left(-\frac{|t - q(a)|}{\beta}\right)$

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

En déduire que pour tout $(a, b) \in \mathcal{V}$, (X_a, X_b) est $\frac{\delta}{\beta}$ -différentiel.

- **1 pt** : inégalité triangulaire $(|t - q(b)| - |t - q(a)|) \leq \delta$

- **1 pt** : exp strictement croissante sur \mathbb{R} donc $\exp\left(-\frac{|t - q(a)|}{\beta}\right) \leq \exp\left(\frac{\delta}{\beta}\right) \exp\left(-\frac{|t - q(b)|}{\beta}\right)$

- **1 pt** : $f_a(t) \leq \exp\left(\frac{\delta}{\beta}\right) f_b(t)$ et $f_b(t) \leq \exp\left(\frac{\delta}{\beta}\right) f_a(t)$

c) Comment choisir β pour disposer alors d'un procédé de ε -confidentialité de D^n pour q ?

- **2 pts** : $\beta = \frac{\delta}{\varepsilon}$

15. Dans cette question, pour tout $a = (a_1, \dots, a_n)$ appartenant à D^n , $q(a) = \sum_{k=1}^n a_k$.

a) Quelle est la valeur de δ ?

- **1 pt** : pour a et b voisins de D , $q(a) - q(b) = a_{i_0} - b_{i_0} \in \llbracket -d, d \rrbracket$

- **1 pt** : $|q(a) - q(b)| \leq d$, donc $\delta \leq d$

On utilise dans la suite le procédé de ε -confidentialité tel qu'il a été défini dans la question 14 mais au lieu de publier la valeur X_a , on procède ainsi :

× si $X_a < \frac{1}{2}$ on publie 0 ;

× si $X_a \in [k - \frac{1}{2}, k + \frac{1}{2}[$ où $k \in \llbracket 1, nd - 1 \rrbracket$, on publie k ;

× sinon on publie nd .

b) Montrer que la valeur aléatoire Z_a publiée vérifie : $Z_a = \begin{cases} 0 & \text{si } X_a < \frac{1}{2} \\ \lfloor X_a + \frac{1}{2} \rfloor & \text{si } X_a \in [\frac{1}{2}, nd - \frac{1}{2}[\\ nd & \text{si } X_a \geq nd - \frac{1}{2} \end{cases}$

- 1 pt : cas $X_a < \frac{1}{2}$

- 1 pt : cas $X_a \geq nd - \frac{1}{2}$

- 2 pts : cas $X_a \in [\frac{1}{2}, nd - \frac{1}{2}[$ (1 pt pour $X_a + \frac{1}{2} \in [k, k + 1[$, 1 pt pour $\lfloor X_a + \frac{1}{2} \rfloor = k = Z_a$)

c) Écrire un script **Python** qui pour d , n et ε saisis par l'utilisateur, génère une valeur aléatoire de $a \in D^n$ puis affiche $q(a)$ et Z_a .

- 1 pt : question 14 et 15.a) donnent $X_a = q(a) + Y$ avec $Y \hookrightarrow \mathcal{L}(0, \frac{d}{\varepsilon})$

- 6 pts

```
1 import numpy as np
2
3 d = int(input('Entrez un entier d plus grand que 1 :'))
4 n = int(input('Entrez un entier n plus grand que 2 :'))
5 eps = float(input('Entrez un réel epsilon strictement positif :'))
6
7 a = rd.randint(0,d+1,n)
8 q = np.sum(a)
9 X = q + Laplace(0, d/eps)
10 if X < 1/2:
11     Z = 0
12 elif X < n*d - 1/2:
13     Z = np.floor(X + 1/2)
14 else:
15     Z = n*d
16
17 print(q)
18 print(Z)
```

d) Pour $n = 1000$, $d = 4$ et ε choisi par l'utilisateur, écrire un script **Python** qui estime la valeur moyenne de $\frac{|Z_a - q(a)|}{q(a)}$ (on considèrera que $q(a)$ est toujours non nul).

- 6 pts


```
1  d = 4
2  n = 1000
3  N = 10000
4  eps = float(input('Entrez un réel epsilon strictement positif :'))
5
6  tabT = np.zeros(N)
7  a = rd.randint(0,d+1,n)
8  q = np.sum(a)
9  for i in range(N):
10     X = q + Laplace(0, d/eps)
11     if X < 1/2:
12         Z = 0
13     elif X < n*d - 1/2:
14         Z = np.floor(X + 1/2)
15     else:
16         Z = n*d
17     tabT[i] = abs(Z-q) / q
18  print(np.sum(tabT) / N)
```