

ESTIMATION PARAMÉTRIQUE D'UN MODÈLE q -GAUSSIEN ET D'UN MODÈLE MÉLANGE DE q -GAUSSIENNES

Oumaima Ben Mrad¹ & Afif Masmoudi² & Yousri Slaoui³

¹ *Laboratoire de Probabilités et Statistique, Sfax, Tunisie et Laboratoire de Mathématiques et Applications, Poitiers, France. oumaima.benmrad.fss@gmail.com*

² *Laboratoire de Probabilités et Statistique, Sfax, Tunisie. afif.masmoudi@fss.usf.tn*

³ *Laboratoire de Mathématiques et Applications, Poitiers, France. Yousri.Slaoui@math.univ-poitiers.fr*

Résumé. Dans ce travail, nous abordons l'estimation des paramètres de la q -distribution q -Gaussienne $\mathcal{N}_q(\mu, \sigma^2)$ introduite par Ben Mrad et al. (2023), une généralisation de la q -distribution $\mathcal{N}_q(0, 1)$ introduite par Diaz et Pariguan (2009). Pour ce faire, nous commençons par introduire une nouvelle distribution discrète, appelée q -Gaussienne discrète $\mathcal{N}_q^d(\mu, \sigma^2)$, associée à chaque q -Gaussienne quantique introduite par Diaz et Pariguan (2009). Ensuite, nous procédons à l'estimation des paramètres de la distribution $\mathcal{N}_q^d(\mu, \sigma^2)$, ce qui implique une estimation des paramètres de la distribution q -Gaussienne $\mathcal{N}_q(\mu, \sigma^2)$. Nous appliquons pour cela les méthodes des moments et du maximum de vraisemblance. De plus, nous étudions un mélange fini de q -Gaussiennes discrètes et appliquons l'algorithme Espérance-Maximisation (E-M) pour estimer les paramètres du mélange. Enfin, nous menons une étude de simulation pour évaluer le modèle et les méthodes d'estimation.

Mots-clés. q -Calcul, q -Distribution, q -Gaussienne, Mélange fini, Estimation Paramétrique, Algorithme E-M.

Abstract. In this work, we address parameters estimation of the q -Gaussian q -distribution $\mathcal{N}_q(\mu, \sigma^2)$ introduced by Ben Mrad et al. (2023), a generalization of the q -distribution $\mathcal{N}_q(0, 1)$ introduced by Diaz and Pariguan (2009). To do this, we begin by introducing a new discrete distribution, called q -discrete Gaussian $\mathcal{N}_q^d(\mu, \sigma^2)$, associated for each quantum q -Gaussian introduced by Diaz and Pariguan (2009). Next, we proceed to estimate the parameters of $\mathcal{N}_q^d(\mu, \sigma^2)$ distribution, which involves estimating the parameters of the q -Gaussian q -distribution $\mathcal{N}_q(\mu, \sigma^2)$. We apply the methods of moments and maximum likelihood. Then, we study a finite mixture of q -discrete Gaussians and apply the Expectation-Maximization (E-M) algorithm to estimate the mixture parameters. Finally, we conduct a simulation study to evaluate the model and estimation methods.

Keywords. q -Calculus, q -Distribution, q -Gaussian, Finite mixture, Parametric estimation, E-M algorithm.

1 Introduction

La naissance du calcul quantique, également connu sous le nom de q -calcul, remonte à Léonard Euler, qui a introduit le paramètre q dans les séries infinies de Newton en 1740. Depuis sa création et jusqu'à nos jours, le q -calcul est un sujet de recherche et de développement majeur dans plusieurs domaines mathématiques et physiques. Ensuite, en 1910, Jackson (1910) a introduit les concepts de la q -intégrale et de la q -dérivée, ce qui a été une contribution marquante. Dans leur livre, Kac et Cheung (2002) ont repris et amélioré plus tard le q -calcul.

En analyse quantique, on s'intéresse généralement aux q -analogues, où il n'existe pas une règle générale pour définir la notion du q -analogue. Une définition intuitive d'un q -analogue d'un tel objet mathématique A , est une famille d'objets $(A_q)_{q \in \mathbb{R} \setminus \{1\}}$, tels que A_q tend vers A lorsque q tend vers 1. En particulier, plusieurs mathématiciens ont été intéressés par la construction d'un q -analogue de la distribution Gaussienne classique qui revient à la construction d'un q -analogue de sa fonction densité, appelée fonction q -densité. Parmi eux on trouve Diaz et Pariguan qui ont introduit un q -analogue de la distribution Gaussienne centrée réduite notée $\mathcal{N}_q(0, 1)$. La fonction q -densité de cette q -distribution est donnée par

$$s_q(x) = \frac{1}{c(q)} E_{q^2}^{-\frac{q^2 x^2}{[2]_q}} \mathbb{1}_{[-v, v]}(x) \quad \forall 0 < q < 1. \quad (1.1)$$

Où $c(q) := \int_{-v}^v E_{q^2}^{-\frac{q^2 x^2}{[2]_q}} d_q x$ est la constante de normalisation de $\mathcal{N}_q(0, 1)$, $v := \frac{1}{\sqrt{1-q}}$, E_q^x est la fonction q -exponentielle introduite par Jackson (2010) définie par $E_q^x = \sum_{j=0}^{\infty} q^{j(j-1)/2} \frac{x^j}{[j]_q!}$. Avec,

$$[\alpha]_q = \frac{1-q^\alpha}{1-q}, \quad \forall \alpha \in \mathbb{R} \quad \text{et} \quad [n]_q! = \begin{cases} 1 & \text{si } n = 0 \\ [n]_q [n-1]_q \dots [1]_q & \text{si } n \geq 1. \end{cases}$$

Dans notre papier, Ben Mrad et al. (2023), nous avons généralisé cette q -Gaussienne en une q -Gaussienne de moyenne μ et de variance σ^2 de la façon suivante : En premier lieu, nous avons introduit la q -distribution $\mathcal{N}_q(0, \sigma^2)$, de fonction q -densité définie par

$$g_q(y) = \frac{1}{c(\sigma, q)} E_{q^2}^{-\frac{q^2 y^2}{\sigma^2 [2]_q}} \mathbb{1}_{[-\sigma v, \sigma v]}(y).$$

En deuxième lieu, nous avons caractérisé la q -Gaussienne $\mathcal{N}_q(\mu, \sigma^2)$ par :

$$Z \sim \mathcal{N}_q(\mu, \sigma^2), \quad \text{si } Z - \mu \sim \mathcal{N}_q(0, \sigma^2).$$

Notre objectif dans cette communication est d'estimer les paramètres q , μ et σ^2 de la q -Gaussienne $\mathcal{N}_q(\mu, \sigma^2)$ et d'étudier et d'estimer un modèle de mélange fini de q -Gaussiennes. Les résultats présentés dans cette communication fait l'objet d'un article intitulé "The discrete q -Gaussian distribution $\mathcal{N}_q(\mu, \sigma^2)$: Properties and parameters estimation" publié dans le journal "Pysics Letters A" (voir Ben Mrad et al. (2024)).

1.1 La distribution q -Gaussienne discrète $\mathcal{N}_q^d(\mu, \sigma^2)$

La définition suivante présente une nouvelle distribution discrète, appelée q -Gaussienne discrète centrée réduite $\mathcal{N}_q^d(0, 1)$, associée à chaque q -Gaussienne quantique introduite par Diaz et Pariguan (2009) donnée par l'Equation 1.1.

Définition 1.1. Soient $0 < q < 1$, $v = \frac{1}{\sqrt{1-q}}$ et $T := \{\pm q^j v; j \geq 0\} \cup \{0\}$. On dit qu'une variable aléatoire X de support T suit la distribution discrète q -Gaussienne, notée $\mathcal{N}_q^d(0, 1)$, si sa fonction de masse est donnée par

$$\mathbb{P}(X = x) = (1 - q)|x|s_q(x)\mathbb{1}_T(x).$$

Cette distribution peut être généraliser en une distribution discrète q -Gaussienne de moyenne μ et d'écart type σ en faisant un changement de variable affine à partir d'une variable aléatoire $X \sim \mathcal{N}_q^d(0, 1)$, comme le montre la proposition suivante.

Proposition 1.1. Soient $\mu \in \mathbb{R}$, $\sigma^2 > 0$ et $X \sim \mathcal{N}_q^d(0, 1)$. Alors, la variable aléatoire $Y := \sigma X + \mu$ suit la distribution discrète q -Gaussienne de moyenne μ et d'écart type σ , notée $Y \sim \mathcal{N}_q^d(\mu, \sigma^2)$, et sa fonction de masse est définie comme suit

$$\mathbb{P}(Y = y) = \frac{1 - q}{c(q)} \left| \frac{y - \mu}{\sigma} \right| E_{q^2}^{-\frac{q^2(y-\mu)^2}{[2]_q \sigma^2}} \mathbb{1}_T\left(\frac{y - \mu}{\sigma}\right). \quad (1.2)$$

Théorème 1.1. Soient $\mu \in \mathbb{R}$, $\sigma^2 > 0$ et $0 < q < 1$. Alors, on a

$$\mathcal{U}_{\{-\sigma + \mu; \sigma + \mu\}} \xrightarrow[q \rightarrow 0]{\mathcal{L}} \mathcal{N}_q^d(\mu, \sigma^2) \xrightarrow[q \rightarrow 1]{\mathcal{L}} \mathcal{N}(\mu, \sigma^2).$$

Où \mathcal{L} se réfère à la convergence en loi et $\mathcal{U}_{\{-\sigma + \mu; \sigma + \mu\}}$ se réfère à la distribution Uniforme discrète classique sur $\{-\sigma + \mu; \sigma + \mu\}$.

Le théorème suivant illustre l'expression des moments de la distribution proposée.

Théorème 1.2. Soit $Y \sim \mathcal{N}_q^d(\mu, \sigma^2)$. Alors, les moments de la variable aléatoire Y sont donnés par

$$\mathbb{E}(Y^n) = \mu^n + \sum_{\substack{k=2, \\ k \text{ pair}}}^n [n]_q \sigma^k \mu^{n-k} [k-1]_q !!.$$

Où,

$$[n]_q !! = \begin{cases} \prod_{k=1}^{\frac{n}{2}} [2k]_q = [n]_q [n-2]_q \dots [4]_q [2]_q & \text{si } n \text{ est pair} \\ \prod_{k=1}^{\frac{n+1}{2}} [2k-1]_q = [n]_q [n-2]_q \dots [3]_q [1]_q & \text{si } n \text{ est impair} \end{cases}$$

et $[n]_q = \frac{n!}{k!(n-k)!}$ se réfère au coefficient binomial classique et $n!$ se réfère à la factorielle de n .

En se basant sur le papier de Ben Mrad et al. (2023), nous remarquons que les moments de la distribution q -Gaussienne discrète $\mathcal{N}_q^d(\mu, \sigma^2)$ et les moments de la q -distribution q -Gaussienne $\mathcal{N}_q(\mu, \sigma^2)$ sont égaux. En conséquence, estimer les paramètres de $\mathcal{N}_q(\mu, \sigma^2)$ revient à estimer les paramètres de $\mathcal{N}_q^d(\mu, \sigma^2)$.

2 Estimation paramétrique de $\mathcal{N}_q^d(\mu, \sigma^2)$

Soient (x_1, x_2, \dots, x_n) n observations d'un échantillon de variables aléatoires indépendantes et identiquement distribuées suivant la distribution $\mathcal{N}_q^d(\mu, \sigma^2)$ de même fonction de masse donnée par

$$P(X = x) = \frac{1 - q}{c(q)} \left| \frac{x - \mu}{\sigma} \right| E_{q^2}^{\frac{-q^2(x-\mu)^2}{[2]_q \sigma^2}} \mathbb{1}_T \left(\frac{x - \mu}{\sigma} \right).$$

L'estimation par la méthode des moments est donnée dans la sous section suivante.

2.1 Estimation paramétrique de $\mathcal{N}_q^d(\mu, \sigma^2)$ par la méthode des moments

Dans ce cas, nous supposons que μ , σ^2 et q sont des paramètres inconnus et nous allons les estimer en utilisant la méthode des moments (MM) comme suit

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n x_i := \bar{x}.$$

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2.$$

En approximant le moment théorique centré réduit d'ordre 4 par le moment empirique centré réduit d'ordre 4, nous obtenons

$$\frac{1}{n} \sum_{i=1}^n \left(\frac{x_i - \bar{x}}{\hat{\sigma}} \right)^4 = \mathbb{E} \left(\left(\frac{X - \mu}{\sigma} \right)^4 \right).$$

Alors, nous trouvons

$$\frac{1}{n} \sum_{i=1}^n \left(\frac{x_i - \bar{x}}{\hat{\sigma}} \right)^4 = [3]_q!! = \frac{1 - q^3}{1 - q}.$$

En résolvant cette équation, nous obtenons un estimateur de $q \in]0, 1[$ par la méthode des moments qui est donné par

$$\hat{q} = \frac{-1 + \sqrt{1 - 4 \left(1 - \frac{1}{n} \sum_{i=1}^n \left(\frac{x_i - \bar{x}}{\hat{\sigma}} \right)^4 \right)}}{2}. \quad (2.1)$$

La sous section suivante décrit en détail l'estimation des paramètres par la méthode du maximum de vraisemblance.

2.2 Estimation de μ et σ^2 par la méthode du maximum de vraisemblance

Dans ce cas, nous supposons que q est un paramètre connu et on note par $\Theta = (\mu, \sigma^2)$ l'ensemble des paramètres inconnus. Notre but est d'estimer Θ par la méthode du maximum de vraisemblance.

Les fonctions strictement concave vraisemblance et log-vraisemblance associées au vecteur des paramètres $\Theta = (\mu, \sigma^2)$ sont données, respectivement, par

$$l(x_1, x_2, \dots, x_n; \Theta) = \prod_{i=1}^n \frac{(1-q)}{c(q)} \left| \frac{x_i - \mu}{\sigma} \right| E_{q^2}^{-\frac{q^2(x_i - \mu)^2}{\sigma^2 [2]_q}}.$$

$$L(x_1, x_2, \dots, x_n; \Theta) = \sum_{i=1}^n \log(1-q) - \log(c(q)) + \log|x_i - \mu| - \log(\sigma) + \log \left(E_{q^2}^{-\frac{q^2(x_i - \mu)^2}{\sigma^2 [2]_q}} \right).$$

les dérivées partielle de $L(x_1, x_2, \dots, x_n; \Theta)$ sont données par

$$\begin{cases} \frac{\partial L}{\partial \mu} = \sum_{i=1}^n \frac{-1}{x_i - \mu} + \sum_{l \geq 0} \frac{2q^{2l+2}(x_i - \mu)(1-q)t}{1 - (1-q)q^{2l+2}(x_i - \mu)^2 t} \\ \frac{\partial L}{\partial t} = \frac{n}{2t} + \sum_{i=1}^n \sum_{l \geq 0} \frac{-q^{2l+2}(x_i - \mu)^2(1-q)}{1 - (1-q)q^{2l+2}(x_i - \mu)^2 t}. \end{cases} \quad (2.2)$$

Par conséquence, l'estimateur du maximum de vraisemblance de Θ est obtenu en résolvant le système 2.2 comme étant égal à 0.

Dans la sous section suivante, nous examinons un mélange fini de distributions discrètes q -Gaussiennes et nous estimons les paramètres du mélange en appliquant l'algorithme (E-M).

2.3 Estimation paramétrique d'un mélange fini de q -Gaussiennes discrète

Soient $\mu \in \mathbb{R}$, $\sigma^2 > 0$ et X une variable aléatoire suivant la distribution $\mathcal{N}_q^d(\mu, \sigma^2)$. La fonction de masse de X est donnée par

$$\mathbb{P}(x; \mu, \sigma^2, q) = \frac{1-q}{c(q)} \left| \frac{x - \mu}{\sigma} \right| E_{q^2}^{-\frac{q^2 \left(\frac{x - \mu}{\sigma} \right)^2}{[2]_q}} \mathbb{1}_T \left(\frac{x - \mu}{\sigma} \right).$$

La fonction de masse du mélange de q -Gaussiennes discrètes est donnée par

$$\mathbb{P}(x | \Theta) = \sum_{k=1}^K \pi_k \frac{1-q}{c(q)} \left| \frac{x - \mu_k}{\sigma_k} \right| E_{q^2}^{-\frac{q^2 \left(\frac{x - \mu_k}{\sigma_k} \right)^2}{[2]_q}} \mathbb{1}_T \left(\frac{x - \mu_k}{\sigma_k} \right). \quad (2.3)$$

Où $\Theta = (\pi_1, \dots, \pi_K, \mu_1, \dots, \mu_K, \sigma_1^2, \dots, \sigma_K^2, q)$ représente le vecteur des paramètres du modèle de mélange, $\theta_k := \{\mu_k, \sigma_k^2, q\}$ est l'ensemble des paramètres pour le $k^{\text{ème}}$ composant du mélange et $\pi_1, \pi_2, \dots, \pi_K$ représentent les poids du mélange tels que ($0 < \pi_k < 1$ et $\sum_{k=1}^K \pi_k = 1$). Nous nous intéressons actuellement à l'estimation du vecteur des paramètres du modèle mélange. Nous avons proposé d'appliquer l'algorithme Espérance-Maximisation (E-M) introduit par Dempster et al. (1977) pour estimer le vecteur des paramètres $\{\mu_k, \sigma_k^2; 1 \leq k \leq K\}$ et la méthode des moments pour estimer le paramètre commun inconnu q . Pour cela, nous considérons (X_1, X_2, \dots, X_N) comme étant N variables aléatoires indépendantes tirées du mélange donné par l'Équation (2.3), et (x_1, x_2, \dots, x_N) comme N observations associées à (X_1, X_2, \dots, X_N) . La fonction de vraisemblance incomplète est donnée par

$$l(x_1, x_2, \dots, x_N; \Theta) = \prod_{i=1}^N \mathbb{P}(x_i | \Theta) = \prod_{i=1}^N \sum_{k=1}^K \pi_k \frac{1-q}{c(q)} \left| \frac{x_i - \mu_k}{\sigma_k} \right| E_{q^2}^{-\frac{q^2 \left(\frac{x_i - \mu_k}{\sigma_k} \right)^2}{[2]_q}} \mathbb{1}_T \left(\frac{x_i - \mu_k}{\sigma_k} \right).$$

Afin d'estimer le vecteur des paramètres Θ , nous associons à chaque point de données observé X_i un vecteur aléatoire discret latent $Z_i = (Z_{i1}, \dots, Z_{iK})$. Ce vecteur aléatoire latent est modélisé par une distribution de Bernoulli multivariée avec un vecteur de probabilités (π_1, \dots, π_K) , c'est-à-dire $\mathbb{P}(Z_i = z_i) = \prod_{k=1}^K \pi_k^{z_{ik}}$ où $z_i = (z_{i1}, \dots, z_{iK}) \in \{0, 1\}^K$ et $\sum_{k=1}^K z_{ik} = 1$. La fonction vraisemblance complètes est donc donnée par

$$l_c(x_1, \dots, x_n, z_1, \dots, z_n | \Theta) = \prod_{i=1}^n \prod_{k=1}^K (\pi_k \mathbb{P}_k(x_i | \theta_k))^{z_{ik}}.$$

Ainsi, la log-vraisemblance des données complètes s'écrit comme suit :

$$L_c(x_1, \dots, x_n, z_1, \dots, z_n | \Theta) = \sum_{i=1}^n \sum_{k=1}^K z_{ik} [\log(\pi_k) + \log(\mathbb{P}_k(x_i | \theta_k))].$$

En raison de la présence des données manquantes, nous suggérons, dans notre travail, d'estimer les paramètres par l'algorithme (E-M), qui alterne deux étapes lorsqu'on donne les paramètres initiaux $\Theta^{(0)}$:

- **Étape Espérance** : L'espérance conditionnelle de la fonction log-vraisemblance des données complètes sachant les observations et le vecteur $\Theta^{(l)}$, qui représente le vecteur des paramètres à la l^{me} itération, est donnée par

$$\begin{aligned} Q(\Theta || \Theta^{(l)}) &= \mathbb{E}_{\Theta^{(l)}}(L_c(x_1, \dots, x_n, z_1, \dots, z_n | \Theta) | x_1, \dots, x_n) \\ &= \sum_{i=1}^n \sum_{k=1}^K \mathbb{E}_{\Theta^{(l)}}(z_{ik} | x_1, \dots, x_n) [\log(\pi_k) + \log(\mathbb{P}_k(x_i | \theta_k))]. \end{aligned}$$

Nous notons $\tau_{ik}^{(l)} := \mathbb{E}_{\Theta^{(l)}}(z_{ik} \mid x_1, \dots, x_n) = \frac{\pi_k^{(l)} \mathbb{P}_K(x_i; \theta_k^{(l)})}{\sum_{k=1}^K \pi_k^{(l)} \mathbb{P}_K(x_i; \theta_k^{(l)})}$, la probabilité à postériori

que X_i appartienne à la $k^{\text{ème}}$ composante du mélange à la l^{th} itération. L'expression de la quantité intermédiaire $Q(\Theta \parallel \Theta^{(l)})$ devienne alors

$$Q(\Theta \parallel \Theta^{(l)}) = \sum_{i=1}^n \sum_{k=1}^K \tau_{ik}^{(l)} [\log(\pi_k \mathbb{P}_k(x_i \mid \theta_k))].$$

- **Étape Maximisation** : Il s'agit de maximiser $Q(\Theta \parallel \Theta^{(l)})$ globalement par rapport à Θ .

$$\Theta^{(l+1)} = \underset{\Theta}{\text{Argmax}} Q(\Theta \parallel \Theta^{(l)}).$$

Pour $\varepsilon > 0$ choisie par l'utilisateur assez proche de zéro, et une norme sélectionnée sur l'espace des paramètres, l'algorithme (E-M) s'arrête lorsque la norme de la différence entre deux estimations successives des paramètres est suffisamment petite, c'est-à-dire $\|\Theta^{(l+1)} - \Theta^{(l)}\| \leq \varepsilon$.

Théorème 2.1. Pour $t_k = \frac{1}{\sigma_k^2}$, l'algorithme (E-M) à l'itération $(l+1)$ donne les résultats suivants.

1. L'estimateur de la proportion π_k du mélange est donné par

$$\pi_k^{(l+1)} = \frac{1}{n} \sum_{i=1}^n \tau_{ik}^{(l)}.$$

2. Les paramètres μ_k et σ_k^2 de la $k^{\text{ème}}$ composante du mélange peuvent être trouvés en résolvant les équations suivantes :

$$\sum_{i=0}^n \tau_{ik}^{(l)} \frac{\partial \log(\mathbb{P}_k(x_i \mid \mu_k, t_k))}{\partial \mu_k} = 0 \text{ et } \sum_{i=0}^n \tau_{ik}^{(l)} \frac{\partial \log(\mathbb{P}_k(x_i \mid \mu_k, t_k))}{\partial t_k} = 0.$$

La section suivante présente des études de simulation, qui ont pour but d'évaluer la performance des estimateurs obtenus.

3 Étude numérique

En utilisant la technique classique de simulation par inversion de la fonction de répartition, un échantillon de taille $N = 5000$ est simulé à partir de la distribution discrète q -Gaussienne $\mathcal{N}_q^d(\mu, \sigma^2)$, avec les paramètres réels $\mu = 4$ et $\sigma = 0,8$ et pour différentes valeurs de q . La Figure 1 présente un résumé des résultats obtenus à partir de notre simulation. D'après la Figure 1, lorsque q s'approche de 1 et de 0, la distribution discrète q -Gaussienne présente le même histogramme que les distributions Gaussienne et Uniforme ordinaires, respectivement, comme l'indique le Théorème 1.1.

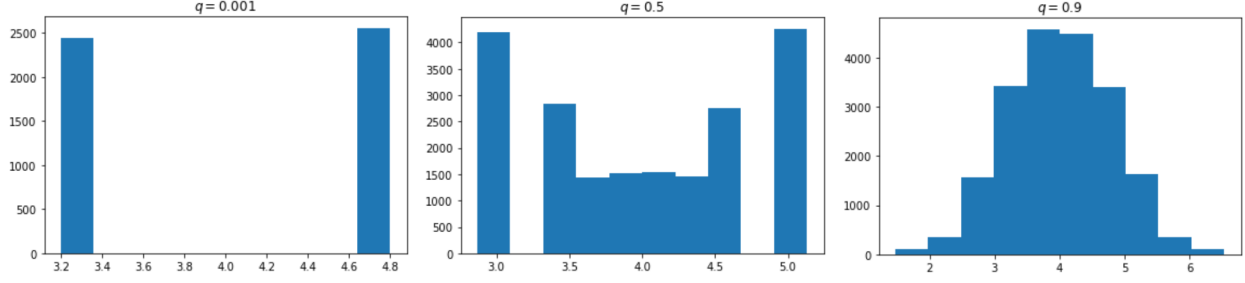


FIGURE 1 – Histogramme d'un échantillon suivant $\mathcal{N}_q^d(4, 0.8^2)$ pour différentes valeurs de q .

TABLE 1 – Estimation de q obtenue à l'aide de la méthode des moments (MM) pour un échantillon de taille $N = 1000$.

Modèle	$\mathcal{N}_q^d(0, 0.1^2)$			$\mathcal{N}_q^d(-4, 1.2^2)$			$\mathcal{N}_q^d(4, 0.8^2)$		
q	0.2	0.5	0.9	0.2	0.5	0.9	0.2	0.5	0.9
\hat{q}	0.1995	0.5002	0.9036	0.2008	0.50006	0.9022	0.1995	0.4997	0.902

Le tableau ci-dessous présente les estimations de q pour trois modèles q -Gaussiens discrets, pour différentes valeurs de q et une taille d'échantillon égale à $N = 1000$, en utilisant la méthode des moments considérée dans l'Équation (2.1). D'après le Tableau 1, le paramètre estimé \hat{q} est très proche de la valeur réelle de q .

Afin d'évaluer la performance des deux méthodes d'estimation, à savoir la méthode des moments (MM) et la méthode du maximum de vraisemblance (MLE) présentées dans la Section 2, nous avons mené une étude de simulation à l'aide de 500 échantillons de tailles différentes, dont $N = 200, 500$ et 1000 observations, qui ont été générés à partir des différentes distributions discrètes q -Gaussiennes $\mathcal{N}_q^d(\mu, \sigma^2)$. Sur la base de ces échantillons, nous avons calculé l'erreur quadratique moyenne (Root Mean Squared Error RMSE) définie comme suit

$$RMSE(\hat{\mu}, \mu) = \sqrt{MSE(\hat{\mu}, \mu)},$$

où, l'erreur quadratique moyenne (Mean Squared Error MSE) est définie par

$$MSE(\hat{\mu}, \mu) = \frac{1}{n} \sum_{i=1}^n (\hat{\mu}^i - \mu)^2.$$

De même, nous avons calculé (RMSE) pour l'écart-type σ , noté, $RMSE(\hat{\sigma}, \sigma)$.

Le Tableau 2 présente l'erreur quadratique moyenne relative (RMSE) de μ et de σ calculée à l'aide des Équations (3) et (3). Comme le montre le Tableau 2, pour tous les modèles considérés, la RMSE diminue avec l'augmentation de la taille de l'échantillon, ce qui conduit à des estimations plus proches aux vrais paramètres. En outre, les deux méthodes d'estimation donnent des valeurs similaires pour chaque modèle, avec une performance légèrement meilleure pour la méthode (MM) que pour la méthode (MLE). Cependant, sauf dans les cas où $q = 0, 2$, la méthode des moments fournit des résultats significativement plus précis que la méthode du maximum de vraisemblance pour l'estimation de σ .

TABLE 2 – Erreur quadratique moyenne relative (RMSE) des paramètres μ et σ .

Modèle	q	N	MM		MLE	
			RMSE($\mu, \hat{\mu}$)	RMSE($\sigma, \hat{\sigma}$)	RMSE($\mu, \hat{\mu}$)	RMSE($\sigma, \hat{\sigma}$)
$\mathcal{N}_q^d(0, 0.1^2)$	0.2	200	0.00683	0.00176	0.0073	0.06613
		500	0.00436	0.00112	0.00555	0.06572
		1000	0.00321	0.00078	0.00396	0.06502
	0.5	200	0.00705	0.00307	0.00715	0.00216
		500	0.0043	0.00195	0.00427	0.00212
		1000	0.00308	0.000001	0.00329	0.00211
	0.9	200	0.00717	0.00471	0.00722	0.00469
		500	0.00458	0.00308	0.00451	0.00293
		1000	0.00302	0.00208	0.00321	0.00213
$\mathcal{N}_q^d(-4, 1.2^2)$	0.2	200	0.08305	0.02281	0.08765	0.08116
		500	0.05604	0.01424	0.06786	0.08108
		1000	0.03824	0.00916	0.05319	0.08021
	0.5	200	0.08643	0.03825	0.08861	0.04362
		500	0.0547	0.02405	0.05526	0.03251
		1000	0.039	0.01611	0.03868	0.02753
	0.9	200	0.08472	0.05621	0.08387	0.05402
		500	0.05416	0.03489	0.05306	0.03393
		1000	0.03752	0.02509	0.03713	0.02446
$\mathcal{N}_q^d(4, 0.8^2)$	0.2	200	0.05365	0.01485	0.05739	0.04732
		500	0.03509	0.00858	0.04328	0.04581
		1000	0.02497	0.00631	0.03441	0.04212
	0.5	200	0.05699	0.02529	0.05864	0.04515
		500	0.03734	0.01666	0.03737	0.04166
		1000	0.02436	0.01082	0.02683	0.03791
	0.9	200	0.05438	0.03682	0.06281	0.03756
		500	0.03666	0.02362	0.03641	0.02482
		1000	0.0232	0.016	0.02429	0.01613

Les paramètres estimés à l'aide de l'algorithme (E-M) d'un mélange de trois distributions q -Gaussienne discrètes pour une valeur fixée de q sont présentés dans le tableau suivant. Le Tableau

TABLE 3 – Paramètres estimés d'un mélange de trois q -Gaussiennes discrètes.

Paramètres	Vrais paramètres	Paramètres estimés
π	[0.37, 0.34, 0.29]	[0.373, 0.336, 0.291]
μ	[0, -4, 4]	[-0.051, -3.974, 3.981]
σ	[0.1, 1.2, 0.8]	[0.946, 1.198, 0.787]

3, montre que les valeurs estimées des paramètres sont proches des valeurs réelles, ce qui indique que l'algorithme (E-M) est robuste et il fournit des estimations précises pour les paramètres d'un mélange fini de q -Gaussiennes discrètes.

4 conclusion

Dans ce travail, nous avons introduit une nouvelle distribution discrète, appelée distribution discrète q -Gaussienne, notée par $\mathcal{N}_q^d(0, 1)$. Nous avons généralisé cette distribution à $\mathcal{N}_q^d(\mu, \sigma^2)$. Par la suite, nous avons estimé le paramètre q à l'aide de la méthode des moments, et les nouveaux paramètres μ et σ^2 à la fois à l'aide de la méthode des moments et la méthode du maximum de vraisemblance. Ce qui implique une estimation des paramètre du modèle q -Gaussien $\mathcal{N}_q(\mu, \sigma^2)$. Ensuite, nous avons examiné un mélange fini de q -Gaussiennes discrètes tout en estimant ses paramètres en appliquant l'algorithme (E-M). Enfin, nous avons réalisé une étude de simulation afin d'évaluer la performances du modèle et des estimateurs proposés.

Bibliographie

- Ben Mrad, O., Masmoudi, A., et Slaoui, Y. (2023), *Some properties of q -Gaussian distributions*, Communications in Statistics - Theory and Methods.
- Ben Mrad, O., Masmoudi, A., et Slaoui, Y. (2024), *The discrete q -Gaussian distribution $\mathcal{N}_q(\mu, \sigma^2)$: Properties and parameters estimation*, Physics Letters A.
- Díaz, R. et Pariguan, E. (2009), *On the Gaussian q -distribution*, Journal of Mathematical Analysis and Applications, 357, pp. 1-9.
- Jackson, F. H. (1910), *q -Difference Equations*, American Journal of Mathematics, 32, pp. 305-314.
- Kac, V. et Cheung, P. (2002), *Quantum calculus*, Universitext. New York : Springer-Verlag.
- Dempster, A. P., Laird, N. M. et Rubin, D. B. (1977), *Maximum likelihood from incomplete data via the EM algorithm*, Journal of the royal statistical society : series B (methodological), 39, pp. 1-22.
- Díaz, R., Ortiz, C. et Pariguan, E. (2010), *On the k -gamma q -distribution*, Central European Journal of Mathematics, 8, pp. 448-458.
- Díaz, R. et Pariguan, E. (2007), *On hypergeometric functions and Pochhammer k -symbol*, Revista Matemática de la Universidad del Zulia. Divulgaciones Matemáticas, 15, pp. 179-192.
- Díaz, R. et Teruel, C. (2005), *q, k -generalized gamma and beta functions*, Journal of Nonlinear Mathematical Physics, 12, pp. 118-134.
- Fitouhi, A. and Bettaibi, N. (2007), *Applications of the Mellin transform in quantum calculus*, Journal of Mathematical Analysis and Applications, 328, pp. 518-534.
- Yamano, T. (2002), *Some properties of q -logarithm and q -exponential functions in Tsallis statistics*, Physica A : Statistical Mechanics and its Applications, 305, pp. 486-496.