



**HAL**  
open science

# Statistical modeling of the attitudes toward mathematics from a school survey data of Mayotte

Jean-Berky Nguala, Solym Manou-Abi

► **To cite this version:**

Jean-Berky Nguala, Solym Manou-Abi. Statistical modeling of the attitudes toward mathematics from a school survey data of Mayotte. 2023. hal-04208748

**HAL Id: hal-04208748**

**<https://hal.science/hal-04208748>**

Preprint submitted on 15 Sep 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Public Domain

# Statistical modeling of the attitudes toward mathematics from a school survey data of Mayotte

Jean Berky Nguala <sup>a,c</sup> and Solym M. Manou-Abi <sup>b,c</sup>

<sup>a</sup>Laboratoire d'Informatique et de Mathématiques, Université de La Réunion, EA 2525 Parc Technologique Universitaire, France.

<sup>b</sup>Institut Montpelliérain Alexander Grothendieck, UMR CNRS 5149, Place Eugène Bataillon, Montpellier, France

<sup>c</sup> Centre Universitaire de Formation et de Recherche, 8 Rue de l'université, Mayotte, France

## ARTICLE HISTORY

Compiled September 15, 2023

## ABSTRACT

Motivated by two survey data samples (student and teacher) obtained from institutional records in some school establishments of Mayotte, this paper has two main objectives. Firstly, a global application of statistical models for the analysis of students' attitude toward mathematics (difficulty, interest or perception, value of the level in mathematics and others attitudes such as social factors, school environment) together with the teachers' attitude toward mathematics (difficult courses, professional training experience, teaching preparation time, etc). Some didactic analysis have helped to reach conclusions to ensure the consistency of certain characterization. Secondly, we develop a theoretical methodology for parameter estimation of the density of the students' level in mathematics and teachers' attitude in the framework of stable and extreme values distributions which are suitable for studying non Gaussian and heavy-tailed distributions. More precisely, we described a new estimation method based on the tail regression estimation method for stable distributions. For the extreme family distributions, we consider a mixture estimation using the EM algorithm. Finally, we end with a brief discussion on a current experimentation of the statistical and didactic analysis results in the form of Lesson study involving several teachers from the island of Mayotte.

## KEYWORDS

Parameter estimation; Stable distribution; Attitude toward mathematics; Statistic analysis; Didactic analysis.

## 1. Introduction

Mathematical knowledge in general achieves greater usefulness when using modeling, Hernandez-Martinez and Vos (2018) and the reflective discussion is established through the modeling process. In the educational environment, sciences in general, mathematics and modeling process are necessary for students to learn other sciences, Brousseau (2006). Unfortunately, the lack of success of students in this area is regularly observed in various national and local assessments. Some results in 2021 and 2022 of the national assessments are summarized in Table 1, **Supplementary**

---

The authors gratefully acknowledge *all people involved in the collection and processing of the survey data especially students, teachers and, including some school heads of Mayotte.*

**Material**, for elementary students evaluations, Andreu et al. (2021) for further information and references therein. We also described in Table 2, **Supplementary Material**), teachers' choice of topics training in french and maths. Students' achievement may be influenced by various facts such as attitudes toward mathematics. This raises various questions for those involved in the education system. Among them, there is a lot of studies concerning the student's attitude toward mathematics (Aiken Jr (1970); Borasi (2014); Ma and Kishor (1997)) such as the difficulty, interest or perception, value of the level in mathematics, social factors, school environment, gender, etc. In particular, the difficulty in understanding and reading sentences highlights the link between difficulties in the language and performing school skills requiring comprehension of instructions. In addition, such difficulties in comprehension through official national language (french) are compounded by difficulties of understanding mathematical contents, see for instance Laborde (1991). This entails that mastering the language of instruction is often necessary to represent and solve problems especially in mathematics. For many students, such school skills are challenging; Julo (1995, 2002); Nguala (2005, 2023). For professional practice, the language and instruction plays a key role in the educational system, De Saussure (1989); Hache (2019); Houdement (2022). The results of various evaluations, both national and of continuous evaluations in the classroom, show that results are very low in french and mathematics (Houdement (2017)), particularly verified in the case of Mayotte island. It is within this framework that two survey were launched to collect data related to maths academic performance and relevant co-variables that can be effected by a number of factors, (Bernier (2020); Mazana, Suero Montero, and Olifage (2019); Michael (2015); Peteros, Gamboa, Etcuban, Dinauanao, Sitoy, and Arcadio (2019); Wakhata, Mutarutinya, and Balimuttajjo (2022)), such as attitudes toward mathematics, teachers instructional practices, etc.

We aim in this paper to highlight the students' and teachers' profiles in order to produce may be an adequacy between them that can be necessary for a better mathematics content teaching and learning in various situations. We conduct statistical analysis and modeling by considering the relevant variables of the attitude toward mathematics. We point out related facts that may justify the need of training in the areas where students have the most difficulties namely in problem sloving (Ashkenazi and Eisner (2022); Houdement (2017)) and to identify some student's profiles related to math level. For the teacher's attitude level toward mathematics, we highlight the training desired for teachers and the different profile recorded. This may suggest the different actions that can allow to better improve the attitude to mathematics not only for students but also between students and teachers. We introduce some parameter estimation method to highlight non gaussian features, stable distributions and the generalized extreme value family distribution.

The originality of this work can be divided into two main objectives. On one hand, the statistical modeling and didactic cross-analysis both for students' and teachers' attitude toward mathematics. On the other hand, we develop a statistical methodology for parameter estimation of the density of the students' level in mathematics and teachers' attitude level in the theoretical framework of stable distributions which are suitable for studying non Gaussian and heavy-tailed distributions. We also deal with parameter estimation for extreme value distributions using stochastic algorithms. This paper is organized as follows: In section 2, the material and data collected and the main objective. We have divided the methodology in section 3 into three main

parts. In the first part, we introduced the didactic analysis through the approaches of semiotic representation. Secondly, we present the statistical analysis tools (variable importance and multidimensional exploration analysis) in section 2. The last part of section 3, deal with the main probability distributions of interest to fit the data distribution of students' maths level and teachers' attitude level towards mathematics as well as the proposed parameter estimation of the above two family distributions. More precisely, we have described a new estimation method based on the tail regression estimation method for stable distributions. For the extreme family distributions, we consider a mixture estimation using the EM algorithm. Section 4 consists of our results presentation and a detailed discussion that follows the main methodology. Finally, we end by a conclusion on how such results allow a current experimentation in the form of Lesson study involving several teachers from the island of Mayotte.

## 2. Material and Data

The present work concerns two survey sample databases related to students' and teachers' attitudes toward mathematics recorded from 167 students and 102 teachers from Mayotte's schools establishments. We refer to the **Supplementary Material** for a full list and description of variables recorded in Table 3 and 4. This preliminary processing of the data is shown partially in Figure 1 and 2 as well as the students' mean maths level and teachers' attitude level toward mathematics distributions. The main objective is the statistical modeling (statistical analysis and parameter estimation) of students' attitudes toward mathematics (perception, usefulness, interest, social factors, school environment) and teachers' attitude towards mathematics (experience, perception, preparation time for lesson design, ability) together with a didactic point of view. All the presented results are implemented using R software.

## 3. Methodology

This study was designed to better understand students' attitudes toward mathematics and teachers' attitude towards mathematics. To achieve this, we opted for a qualitative didactic and statistical research, that consists of several steps. Firstly, we describe the way we'll conduct the statistical analysis (variable selection and clustering) and didactic approaches (semiotic representation and contextualisation). The theoretical statistical modeling framework we have chosen concerns the parameter estimation of some probability densities family (stable distributions and Extremum value distributions) in order to model the students' math level and teachers' math level attitude distributions. Thus, we present three main estimation methods, namely mixture estimation using EM- algorithms for the family of generalized extremum distributions, a new tail regression estimation and log-moment estimation for stable distributions.

### 3.1. *Semiotic and didactic analysis*

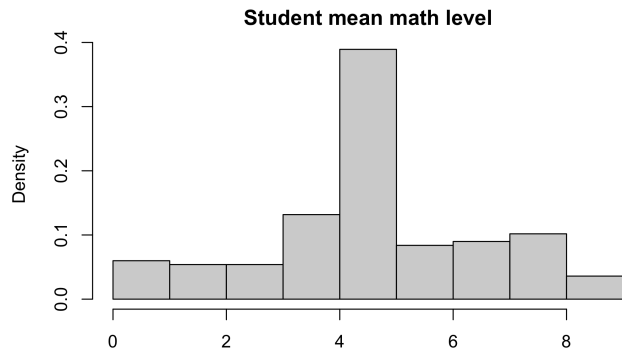
We conduct the didactic thought analysis upstream of the statistical formulations as follows. Language issues in the learning of mathematics is the first focus as also highlighted in Hache (2019); Laborde (1982). The second aspect takes account of the specific difficulties encountered by students in the language of instruction, which in

Description: df [167 x 20]

	Age	School_Name	Classroom	Commune	Nbh_maths	Nbh_science	Nbh_French
1	16	Mamoudzou nord school	SECONDE	MAMOUDZOU	5.00	28.00	7.00
2	16	Mamoudzou nord school	SECONDE	KOUNGOU	2.00	4.00	2.00
3	15	Mamoudzou nord school	SECONDE	KOUNGOU	2.00	3.00	4.00
4	15	Mamoudzou nord school	SECONDE	KOUNGOU	2.00	1.50	2.00
5	15	Mamoudzou nord school	SECONDE	KOUNGOU	0.00	2.00	1.00
6	16	Mamoudzou nord school	SECONDE	KOUNGOU	0.20	1.00	5.00
7	17	Mamoudzou nord school	SECONDE	MAMOUDZOU	3.00	3.00	5.00
8	16	Mamoudzou nord school	SECONDE	MAMOUDZOU	9.00	20.00	4.00
9	16	Mamoudzou nord school	SECONDE	KOUNGOU	5.00	28.00	7.00
10	16	Dembeni high school	SECONDE	MAMOUDZOU	0.00	0.00	0.00

1-10 of 167 rows | 1-8 of 20 columns

(a) A short view of the students' database.



(b) An histogram of students' maths level distribution.

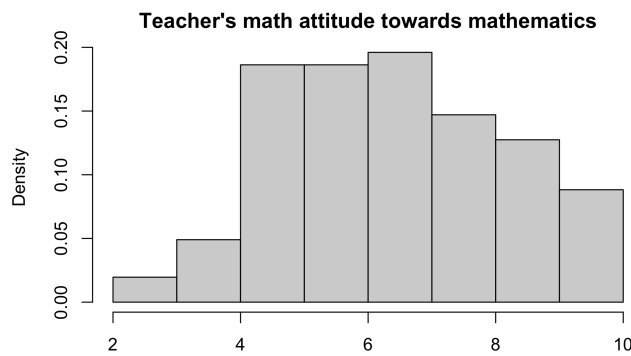
**Figure 1.** Partial view of the students' survey dataset and mean maths level distribution.

Description: df [102 x 10]

	Experience	Commune	Math_level_and_report_as_teacher
2022/04/05 1:49:41 PM GMT+2	22	TSOUNDZOU	7
2022/04/05 2:23:41 PM GMT+2	6	TSOUNDZOU	8
2022/04/05 8:09:45 PM GMT+2	4	CHIRONGUI	8
2022/04/06 7:11:32 PM GMT+2	3	KOUNGOU	7
2022/04/11 6:45:31 AM GMT+2	2	MAMOUDZOU	6
2022/04/15 4:05:28 AM GMT+2	12	MTSAMBORO	6
2022/04/29 3:29:12 PM GMT+2	3	MTSAMBORO	4.5
2022/05/04 7:36:36 PM GMT+2	2	BOUENI	4
2022/05/05 12:36:40 PM GMT+2	6	MAMOUDZOU	6
2022/05/05 12:49:10 PM GMT+2	1	MAMOUDZOU	10

1-10 of 102 rows | 1-4 of 10 columns

(a) A short view of the teachers' database.



(b) An histogram of teachers' attitude level distribution toward mathematics.

**Figure 2.** Partial view of the teachers' survey dataset and attitude level distribution toward mathematics

reality, they use only occasionally and only in class, Dias (2020); Faure (2020). Another aspect refers to the elements of contextualisation that may help the students to build an environment success of the tasks to be performed, Delcroix, Forissier, and Anciaux (2013); Julo (1995); Nguala (2005); Salone (2022). To take account of students' difficulties in the language of instruction, we refer to what is done with those working in didactics on common topic, Faure (2020). Moreover, the complexity of mathematical language and mathematical levels among students can be mainly due to the interweaving of the two codes of natural language and symbolic language as well as their attitudes toward mathematics. These two codes are registers of semiotic representation Duval (1993); that we are going to explore along the discussion.

### ***3.2. Statistical and Data Analysis***

#### *The regression analysis model for variable selection*

Here we use the Generalized Linear Model (GLM), Bingham and Fry (2010); Huet, Bouvier, Poursat, Jolivet, and Bouvier (2004) to explain students' level in mathematics as a function of the others co-variables of the attitudes toward mathematics. The GLM then specifies a relationship between the mean response  $Y(r)$  and the covariates  $X_i(r)$  through a specified monotonic link function  $g$  :

$$g(Y(r)) = \beta_0 + \beta_1 X_1(r) + \dots + \beta_p X_p(r)$$

where  $\beta_0$  is the intercept and  $\beta_l, l > 0$ , is a regression coefficient associated with the covariates  $X_l(r)$ , the  $l$ -th covariate of the respondent  $r$ . The above regression model can be fitted via ordinary least squares (OLS) in which case we find estimates of the parameters.

#### *Multidimensional exploration Analysis*

We conduct principal component analysis (PCA) (see Jolliffe (2005); Patlolla (2020)) on significant variables after the above variable selection method. This analysis is a pre-processing step before moving on to clustering. The unsupervised classification designates a corpus of methods whose objective is to draw up or find an existing typology characterizing a set of  $n$  observations, from  $p$  characteristics measured on each of the observations. We use the complementarity between clustering and principal component methods to better highlight the main feature of our data set. The objective is to obtain a hierarchical tree, in other words a sequence of nested partitions going from a partition in which each individual is considered as a group to one in which all the individuals belong to the same group. The first dimensions of PCA will extract most of the information while the last ones are limited to noise and the number of dimensions retained for the grouping can be chosen with several methods. The `hpc` function of R software allows us to visualize the tree, the partition and the main components in a convenient way. An unsupervised clustering (Praene, Malet-Damour, Radanielina, Fontaine, and Riviere (2019)) is done afterwards based on the principal components in this paper. The results are plotted to correlate the obtained clustered students' and teachers' profiles.

### 3.3. Parameter estimation methods

We start this section, with the main probability distributions of interest for our study: stable and the family of extreme value distributions. It will be the target distributions chosen for some quantitative variables such as students' level in mathematics and teachers' attitudes level towards mathematics.

#### 3.3.1. Stable distributions

Despite the lack of non-analytical expression for their probability density function, the class of  $\alpha$ -stable distributions are popular in many fields as economics, finance, and insurance. For convenience, we provide the definitions of  $\alpha$ -stable distributions.

**Definition 3.1.** A random variable  $X$  is said to follow a stable distribution, denoted by  $X \sim S(\alpha, \beta, \sigma, \mu)$ , if its characteristic function  $\varphi_X(t) = \mathbb{E}(e^{itX})$ ,  $t \geq 0$ , is expressed as follows.

$$\begin{aligned} \phi(t) = & \exp\left(-\sigma^\alpha |t|^\alpha \left(1 - i\beta \left(\tan\left(\frac{\pi\alpha}{2}\right)\right) \text{sign}(t)\right) + i\mu t\right) \mathbb{I}_{\{\alpha \neq 1\}} \\ & + \exp\left(-\sigma |t| \left(1 + i\beta \frac{2}{\pi} \text{sign}(t) \log(|t|)\right) + i\mu t\right) \mathbb{I}_{\{\alpha = 1\}} \end{aligned}$$

where  $i^2 = -1$  and  $\text{sign}(\cdot)$  is the well-known sign function. This kind of parametrization is called the 1-parametrization in Nolan (2012)); see also Samoradnitsky (2017) for other parameterizations (each having advantages and drawbacks). The parameter  $\alpha \in (0, 2]$  is the index of stability also called the tail index,  $\beta \in [-1, 1]$  the skewness parameter,  $\sigma > 0$  the scale and  $\mu \in \mathbb{R}$  the location parameter.

The case  $\mu = 0$  corresponds to a strictly  $\alpha$ -stable distribution. Moreover, it is said to be symmetric if  $\beta = 0$ . In the case where  $\beta = 1$  and  $\alpha < 1$ , it is the class of positive stable distributions. A practical difficulty when working with  $\alpha$ -stable distributions for most cases is the lack of closed formulas for the density and distribution functions.

#### 3.3.2. Generalized Extreme value distributions

Now, let us turn to the family of Extreme Value distributions (EVD) arising as limit distributions for maximums or minimums (extreme values) of a sample of independent, identically distributed random variables, as the sample size increases, see Kotz and Nadarajah (2000). They are important when data distribution reveals events which occur with very low probability. Thus, these distributions are important in statistics and are contained in the following generalized extreme value (GEV) distribution.

**Definition 3.2.** The probability density function for the GEV distribution with location parameter  $\mu \in \mathbb{R}$ , scale parameter  $\sigma > 0$  and shape parameter  $\zeta \in \mathbb{R}$  is

$$f(z/\mu, \sigma, \varepsilon) = \begin{cases} \exp\left(-\left(1 + \varepsilon\left(\frac{z-\mu}{\sigma}\right)\right)^{-1/\zeta}\right) & \text{if } \zeta \neq 0, \\ \exp\left(-\left(\frac{z-\mu}{\sigma}\right)^{-1/\zeta}\right) & \text{if } \zeta = 0 \end{cases}$$

for all  $z$  such that  $1 + \zeta\left(\frac{z-\mu}{\sigma}\right) > 0$ . The case  $\zeta > 0$  corresponds to the Frechet type distribution while  $\zeta < 0$  corresponds to the Weibull type distribution. The case  $\zeta \rightarrow 0$  is the Gumbel distribution.

In this section we consider statistical methods in order to estimate the densities of occurrence of some main quantitative variables such as students' level in mathematics and teachers' attitudes towards mathematics. The probability densities chosen for the estimation are the above family following of Extreme Value and Stable distributions.

### 3.3.3. The proposed estimation method for stable distributions

Recall that, in the case of stable distributions, the tail exponent  $\alpha$  determines the rate at which the tails of the distribution decay. When  $\alpha < 2$ , the variance is infinite and the tails are asymptotically equivalent to a Pareto law. In this sub section and, for practical issues, we shall be concerned with a given sample of positive data. We establish  $\alpha$ -stables distributions parameters estimators.

**Theorem 3.3.** : Assume that, the random variable  $X \sim S(\alpha, \beta, \sigma, 0)$  follows a strictly stable distribution. We propose the following consistent estimators  $\hat{\alpha}$   $\hat{\sigma}$  of  $\alpha, \sigma$  respectively for large  $k$ .

$$\hat{\alpha}_k = \frac{\sum_{l=1}^k \log(x_l) \log(y_l) - \frac{1}{k} \left( \sum_{l=1}^k \log(y_l) \right) \left( \sum_{l=1}^k \log(x_l) \right)}{\frac{1}{k} \left( \sum_{l=1}^k \log(x_l) \right)^2 - \sum_{l=1}^k \log(x_l)^2},$$

$$\hat{b}_k = \frac{1}{k} \sum_{l=1}^k \log(y_l) + \frac{\hat{\alpha}_k}{k} \sum_{l=1}^k \log(x_l).$$

In the symmetric case ( $\beta = 0$ ), given an estimation  $\hat{\alpha}$  of  $\alpha$ , we have :

$$\hat{\sigma} = \left( \frac{e^{\hat{b}_k} 1/\hat{\alpha}}{C_{\hat{\alpha}}} \right).$$

**Proof.** We first remark that, since the tails of any  $\alpha$ -stable distributions are asymptotically equivalent to a Pareto law, we have from Samoradnitsky (2017) that

$$\lim_{x \rightarrow +\infty} x^\alpha \mathbb{P}(X > x) = C_\alpha (1 + \beta) \sigma^\alpha$$



where for  $\alpha \neq 1$  we have

$$C_\alpha = \left( 2 \int_0^{+\infty} x^{-\alpha} \sin(x) dx \right)^{-1} = \frac{1}{\pi} \Gamma(\alpha) \sin\left(\frac{\pi\alpha}{2}\right).$$

Let  $(X_i)_{i=1\dots n}$  be a sequence of sample observations of  $X$  with size  $n$ . In this case,  $\frac{\sum_{i=0}^{n-1} \mathbb{I}_{\{Y_i > x\}}}{n}$  asymptotically approximate  $\mathbb{P}(X > x)$  for large  $n$ . Hence for a given  $k$  largest positive values  $x_l \gg 1$  of  $x$ , we introduce the following statistical model, under classical assumptions as in a linear regression model. More precisely,

$$y_l := \log\left(\frac{\sum_{i=1}^n \mathbb{I}_{\{X_i > x_l\}}}{n}\right) \approx -\alpha \log(x_l) + \log(C_\alpha(1 + \beta)\sigma^\alpha) + \varepsilon_i$$

where the noises  $\varepsilon_i$  are uncorrelated and centered. Set  $b = \log(C_\alpha(1 + \beta)\sigma^\alpha)$ . We use the Least Square Estimation (LSE) which consists in minimizing the following objective function:

$$\mathbb{G} = \sum_{i=1}^k [\log(y_i) - b + \alpha \log(x_i)]^2$$

Thus it is straightforward to obtain the desired estimators by derivation of  $\mathbb{G}$  with respect to  $\alpha$  and  $b$ . The consistency of these estimators can easily be obtained under the above assumptions on the noises  $\varepsilon_i$ .  $\square$

As concerned the other parameters of the strictly stable distributions, namely  $\sigma$  and  $\beta \neq 0$ , we will use the following method called the logarithmic moment, see Kuruoglu (2001). This method was proposed to establish a moment type method for parameter estimation even though the first moment (for  $\alpha < 1$ ), second and higher moments ( $\alpha > 1$ ) of stable distributions are not finite. The following result based on the paper of Kuruoglu (2001) give a practical method through the moments of the logarithm of the  $\alpha$ -stable random variable; which are finite for certain parameter values.

**Proposition 3.4.** : *Consider a strictly  $\alpha$ -stable random variable  $X \sim S(\alpha, \beta, \sigma, 0; 1)$  and set  $\theta = \arctan(\beta \tan(\frac{\alpha\pi}{2}))$ . Then, we have*

$$\begin{aligned} L_1 &= E[\log(|X|)] = \gamma \left( \frac{1}{\alpha} - 1 \right) - \frac{1}{\alpha} \log(|\cos(\theta)|) + \log(\sigma) \\ L_2 &= E[(\log(|X|) - E[\log(|X|)])^2] = \left( \frac{\pi^2}{6\alpha^2} + \frac{\pi^2}{12} - \frac{\theta^2}{\alpha^2} \right) \\ L_3 &= E[(\log(|X|) - E[\log(|X|)])^3] = c_1 \left( 1 - \frac{1}{\alpha^3} \right) \end{aligned}$$

where  $\gamma = 0.57721\dots$  is Euler constant and constant  $c_1 = 1.2020569$  relies to the Riemann zeta function.

From the above proposition, we derive the following result.

**Theorem 3.5.** : *Let  $(Y_i)_{i=1}^n = (\log |X_i|)_{i=1}^n$ , where we assume that  $X = (X_1, \dots, X_n)$  is a sequence of independent strictly  $\alpha$ -stable random variables  $X \sim S(\alpha, \beta, \sigma, 0; 1)$  for*

$\alpha \neq 1$ . Given an estimation  $\hat{\alpha}$  of  $\alpha$ , set

$$\bar{Y} = Y - \mathbb{E}(Y) \quad \text{and} \quad \bar{Y}_n = \frac{\bar{Y}_1 + \dots + \bar{Y}_n}{n}.$$

Define

$$|\hat{\theta}_n| = \frac{\pi^2}{6} + \left(\frac{\pi^2}{12} - \bar{Y}_n^2\right)\hat{\alpha}^2; \quad \hat{\beta}_n = \frac{\tan(\hat{\theta}_n)}{\tan(\hat{\alpha}\frac{\pi}{2})}$$

$$\hat{\sigma}_n = |\cos \theta_n|^{1/\hat{\alpha}} e^{\bar{Y}_n + \gamma(1 - \frac{1}{\hat{\alpha}})}.$$

Then, we have

$$\begin{cases} \hat{\beta}_n \xrightarrow{p.s.} \beta \\ \hat{\sigma}_n \xrightarrow{p.s.} \sigma \end{cases}$$

when  $n \rightarrow +\infty$ .

**Proof.** We recall firstly Proposition 4 in Kuruoglu (2001) and by equating the sample moments (first, second first and third moments) and the real moments, we easily derive the estimators formula. Finally, the strong law of large numbers for the random variables

$$\left(\log Y_i - \mathbb{E}(\log Y_i)\right)^3, \left(\log Y_i - \mathbb{E}(\log Y_i)\right)^2, \log Y_i$$

together with the continuous mapping theorem give the consistency of the proposed estimators.  $\square$

### 3.3.4. Mixture estimation method using EM Algorithm for EVD

The EM algorithm, known as the most popular inferential tool, was developed by Dempster et al Dempster, Laird, and Rubin (1977) to find the maximum likelihood of the parameters of probabilistic models. Let  $X = (x_1, x_2, \dots, x_n)$  be a random sample of size  $n$  containing observations, where  $x_i$  is a realisation of a random variable with probability density function  $f_j$ . A mixture estimation which is a flexible tool (see (Li and Barron, 1999; Redner and Walker, 1984) ) to model a known smooth probability density function as a weighted sum of parametric density functions. The density of a mixture model with  $k$  components for one observation  $x_i$  is given by the mixture density  $p(x_i) = \sum_{j=1}^k p_j f_j(x_i, \theta_j)$  where  $p = (p_1, \dots, p_k)$  contains the corresponding mixture proportions with  $\sum_{i=1}^k p_i = 1$  and  $f_j(x_i, \theta_j)$  is the density component of mixture  $j$  and  $\theta_j, j = 1, 2, \dots, k$  are vectors of component specific parameters for each density and  $\theta = (\theta_1, \dots, \theta_n)$  denotes the vector of all parameters of the model. The log likelihood of the model for a sample of size  $n$  is then given by

$$L(\theta) = \sum_{i=1}^n \log\left(\sum_{j=1}^k p_j f_j(x_i, \theta_j)\right).$$

The sum of term appearing inside the logarithm makes the optimization quite difficult. There are, of course, many general iterative procedures which are suitable for finding an approximate solution of the likelihood equations. The parameters can be estimated using the EM algorithm. For this purpose, let us introduce latent variables  $Z = (Z_1, \dots, Z_k)$ , which are categorical variables taking on values  $1, \dots, k$  with probabilities  $p_1, \dots, p_k$  such that  $\mathbb{P}(X_i/Z_i = j) = f_j(x_i, \theta_j)$  for  $j = 1, \dots, k$ . The EM algorithm works in two (E and M) steps. In the E-step, it estimates the expected value of the latent variables  $Z_i$  given the current parameter estimates. In the M-step, it maximizes the likelihood of the observed data based on the expected values obtained in the E-step. This alternating process continues until convergence is achieved, resulting in the optimal parameter estimates for the model. The EM algorithm is summarised as follows.

---

**Algorithm 1** EM ALGORITHM

---

**Input:**  $\theta^{(0)}$  initial value of the model parameters;

- 1:  $X$  a set of the observed data distribution ;
- 2:  $\epsilon$  a threshold for the convergence of the algorithm.

**Output:**  $\theta$  optimal parameter estimates

- 3: **E**-step. Compute the conditional expectation of the complete log-likelihood function at current  $t$ -th iteration:

$$Q(\theta, \theta^{(t)}) = \mathbb{E}\left(L_c(\theta, p, Z) | X, \theta^{(t)}\right)$$

- 4: **M**-step. Maximize the above conditional expectation:

$$\theta^{(t+1)} = \operatorname{argmax}_{\theta} (Q(\theta | \theta^{(t)}))$$

- 5: **while**

$$|Q(\theta | \theta^{(t+1)}) - Q(\theta | \theta^{(t)})| > \epsilon$$

**do**

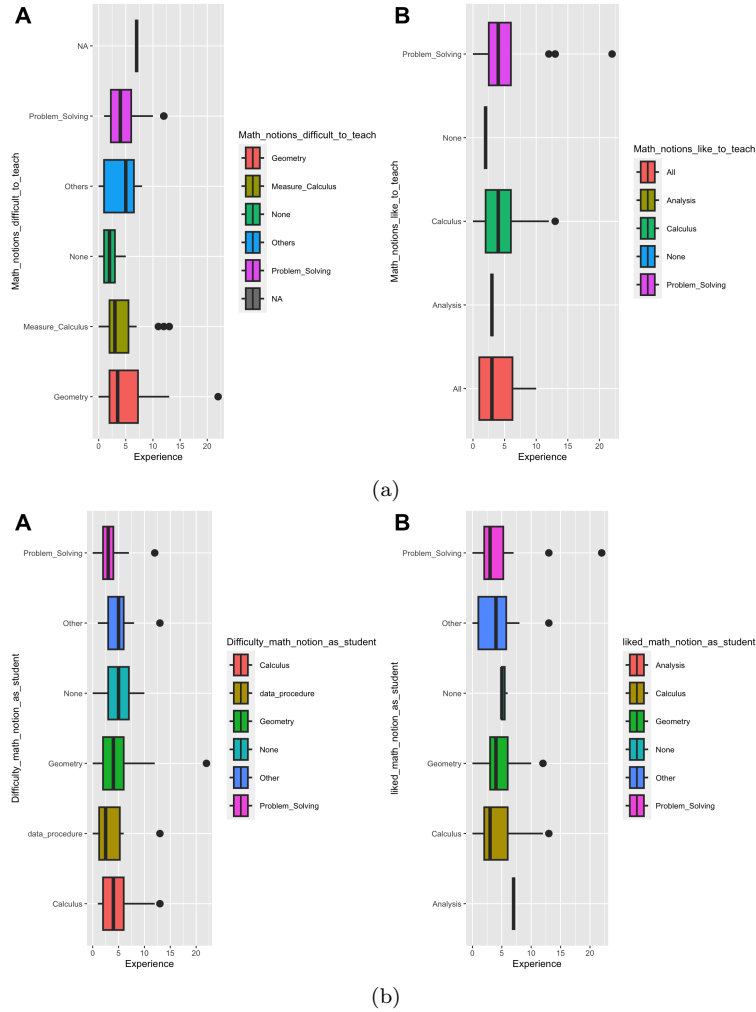
- 6: Repeat the **E** and **M**-steps.

7:

8: **end while**

9: **return**  $\theta$

---



**Figure 3.** Experience and related teachers' conception of mathematical concepts and professional training desired.

#### 4. Results and Discussion

The results presented in this section is concerned with the above mentioned survey sample databases. Let us start with the statistical and didactic cross analysis. The analysis of the mathematics' elementary concepts that teachers find difficult to teach reveals a high proportion of number and calculation, geometry and problem solving, whereas the concepts they would like to teach are essentially number and calculation, and problem solving for generally young teachers (less than 5 years of experience), Figure 3. Moreover we remark that many teachers (for instance young teachers) who had difficulties in problem solving, geometry, numbers and calculations previously as students want to overcome such difficulties. In summary, some particular teachers who have difficulties in a given area originally as a student (geometry, problem solving, numbers and calculations) and who seek professional training in these areas to fill their gaps, like to teach them. Professional training in mathematics depends on the experience of the teachers as shown in Figure 4.

Thus, young teachers want general training in fundamental aspects such as geometry, problem solving, numbers and calculations, pedagogy and didactics, except for

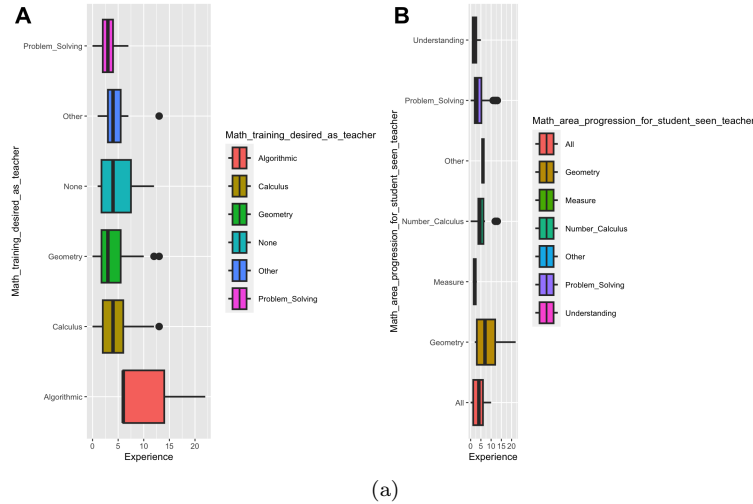
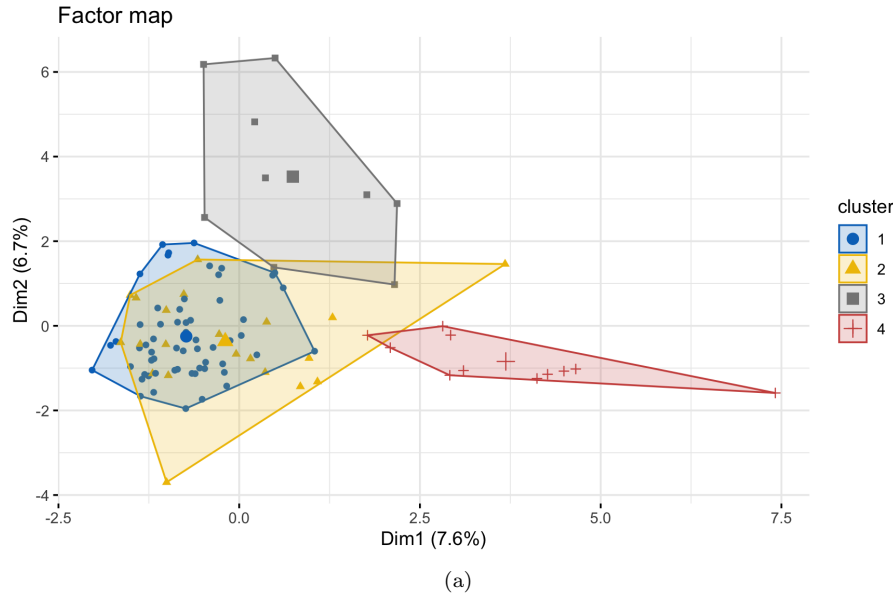


Figure 4. Experience and professional training desired

a few experienced teachers (more than 5 years of experience). There is a desire for algorithmic training among the latter teachers. Such observations, if confirmed in a global study, should make it possible to envisage the teaching-learning process for a constructive engagement and the future development of the teaching profession; see Kansanen and Meri (1999); Pino-Fan, Assis, and Castro (2015). A generalized Linear Model apply to the data was used to select variables (according to the p-value in Fig. 1 of the **Supplementary Material**) allow to conduct the above mentioned unsupervised clustering method. We found that the selected variables are involved in each cluster group (4 in total), see Figure 5.

Clusters 1 and 2 are composed of young teachers as opposed to Clusters 3 and 4. What these groups have in common is that they all enjoy teaching calculus and they mostly want training (65 per cent of them) in calculus and problem-solving skills to help their pupils make progress in these areas. In Cluster 1, as students, the difficulties of teachers were mainly calculus and problem solving. Geometry is a difficult subject to teach and is in demand as the desired professional training for clusters 2 to 4. Moreover, Algorithm is a special subject training requested by teachers in Cluster 4. In Clusters 2 and 4, they find it difficult to teach geometry and calculus or numbers whereas in Clusters 3 and 4 they find difficult to teach problem solving. In Cluster 3, they find it difficult to teach geometry, numbers and calculations and problem solving. Paradoxically in Cluster 2, teachers (67 per cent of total individuals) who found problem solving difficult to teach, none of them requested training in these areas. A new clustering analysis adding the notions they found difficult as students during their past student training period, show that mathematical contents, teachers found difficult as students during their student training period did not change the teacher's needs in mathematics training desired to strengthen themselves. In general in Clusters 1, 3 and 4, the attitude towards mathematics is greater than the mean. Cluster 2 differs from the others for its low average attitude towards mathematics compared to Cluster 3 which have a higher attitude average towards mathematics and also a higher teaching preparation time.

Now, let us turn to the students' database. Regarding the importance of mother tongue, we find that those whose mother tongue is French and Malagasy have a level



**Figure 5.** Clustering of the recorded teachers profiles

in maths greater than the average value recorded; otherwise, the level differ from one to another; see Figure 6. Regardless of their level in mathematics, they ask for explanations, illustrations and comprehension in their mother tongue. This shows that a mathematical level and hence language is not linked to the language of the individual. It is more than that, since it is also composed of a semiotic language in the form of symbolic writings, semiotic representations and registers, which gives its own complexity. This point of view has been developed by since by many researchers, see Barrier and Durand-Guerrier (2017); Shaftel, Belton-Kocher, Glasnapp, and Poggio (2006).

Moreover, they have language difficulties in understanding maths concepts and the assimilated maths notions is linked to their mathematics level, see Figure 7. This result is also similar to the general work of Dias (2020); Prasad (2014) and raises questions about the implementation of a system where there would be exercises to complete with instructions in french and in students' mother tongue. The mathematical concepts assimilated by students vary according to their level of mathematics. We can see that for maths levels above the mean, algorithms, statistics and geometry are mostly assimilated and contribute to an increase of the level in mathematics. Among the other concepts assimilated with an average level were algebra and analysis. Some students with a very mixed level state that geometry and maths in general are assimilated notions. Those with a good math level felt that they had no difficulty. We remark that there is a lack of motivation for students mathematics levels under the mean level. The latter do not receive any private lessons and do not have enough time to learn maths (coming back to this; note that half of them do not have any teaching hours in maths). For all students with different mathematics levels, we pointed out concentration, reading and training difficulties.

This is confirmed by the odds ratios resulting from the above described generalized Linear Model (see Fig. 1 of the **Supplementary Material**). This shows that the following variables contribute to an increase of the level of mathematics : the mathematical concepts assimilated, such as algorithm, analysis, statistics, geometry; the

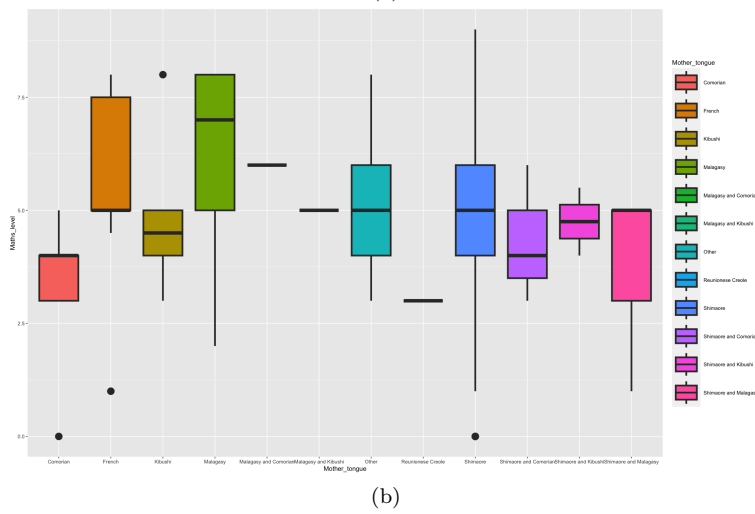
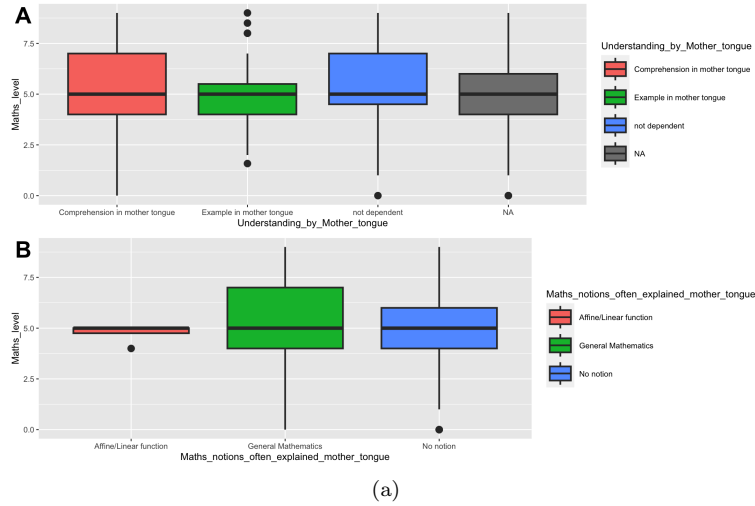


Figure 6. Observed language characteristics in student database.

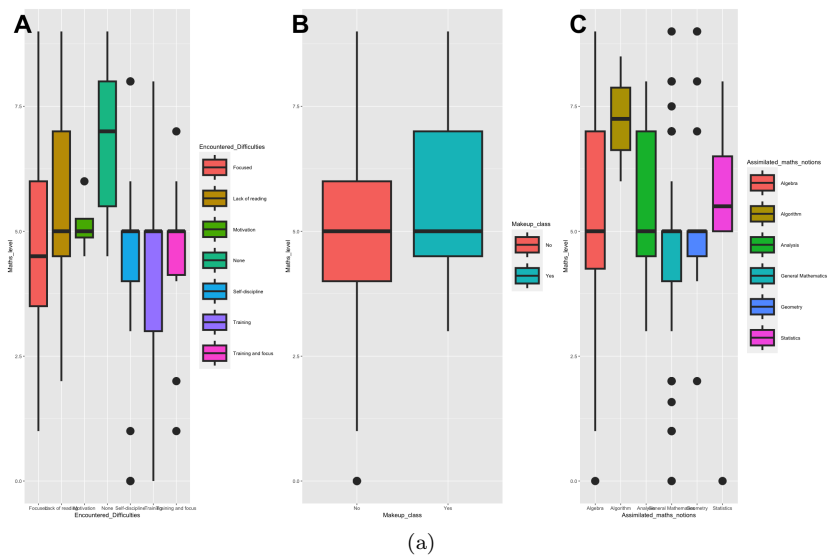


Figure 7. Observed language difficulties and assimilated maths notions of students.

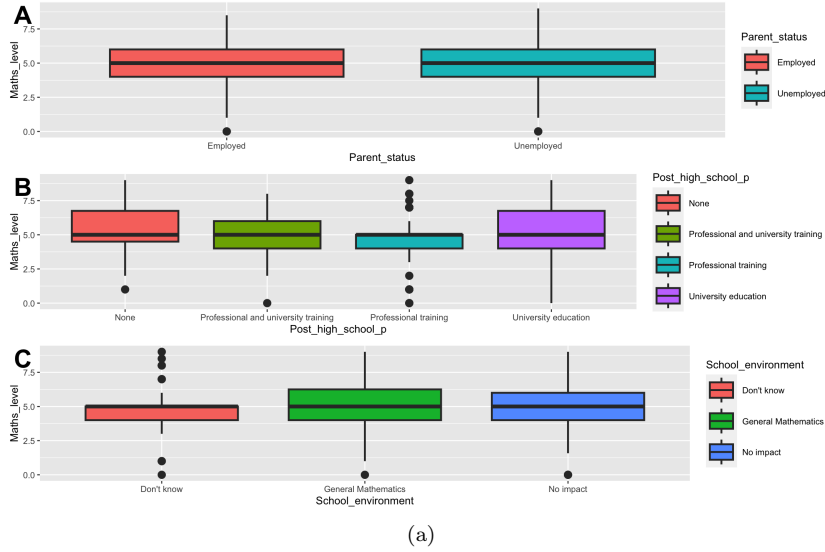


Figure 8. Observed social factors and maths level of students.

school environment, comprehension through the mother tongue, private lessons, parents' status. The other variables (self-discipline, concentration and training) entail a decreasing in maths levels. For a third of the students, we find that concepts of algebra such as linear functions and mathematics in general are often explained in their mother tongue. With regard to the profiles of the students, a common point which emerges in the mathematical concepts assimilated are in algebra and with different kind of access to Internet. We observe also that the school environment (furniture, material learning conditions excluding learning time in class) impacts their level in mathematics, see Figure 8. Otherwise, students level in mathematics seems to be independent of the Internet access, parents' status and place of birth; but unemployed parents had a negative impact on the students maths level.

The unsupervised classification reveals four type of students profiles in our database, see Figure 9. Cluster 3 (12 students) showed a very good level in maths (above 7) compared to the other profile groups and almost no difficulties were highlighted. Cluster 1 (28 students) did not record a below-average level in maths. Unlike the other clusters, Cluster 1 shows very few students wishing to go on to university, and most of them do not know where they are going. In addition, these students felt that they did not need to understand mathematics through their mother tongue. Cluster 2 (48 students) and Cluster 4 have in common the fact that not only the student average in mathematics level is around 4 but also the fact they usually need comprehensions in the mother tongue. In cluster 2, the school environment has a higher impact than the others. Among the clusters, only cluster 4 (18 students) reveals a less important impact of the school environment.

Now let us discuss about the parameter estimation strategy developed above. The following result is the EM estimation based on a two-component mixture model, the mixture component parameter is confirmed by the use of k-means clustering method Sinaga and Yang (2020). We summarize the estimation results of the mixture models in Figures 10 and 11 and in Table 6 and 7 of the **Supplementary Material**. We conclude that the best-fit mixture is given by a Weibull mixture distributions.

As concerned the proposed new estimation method based on the tail regression



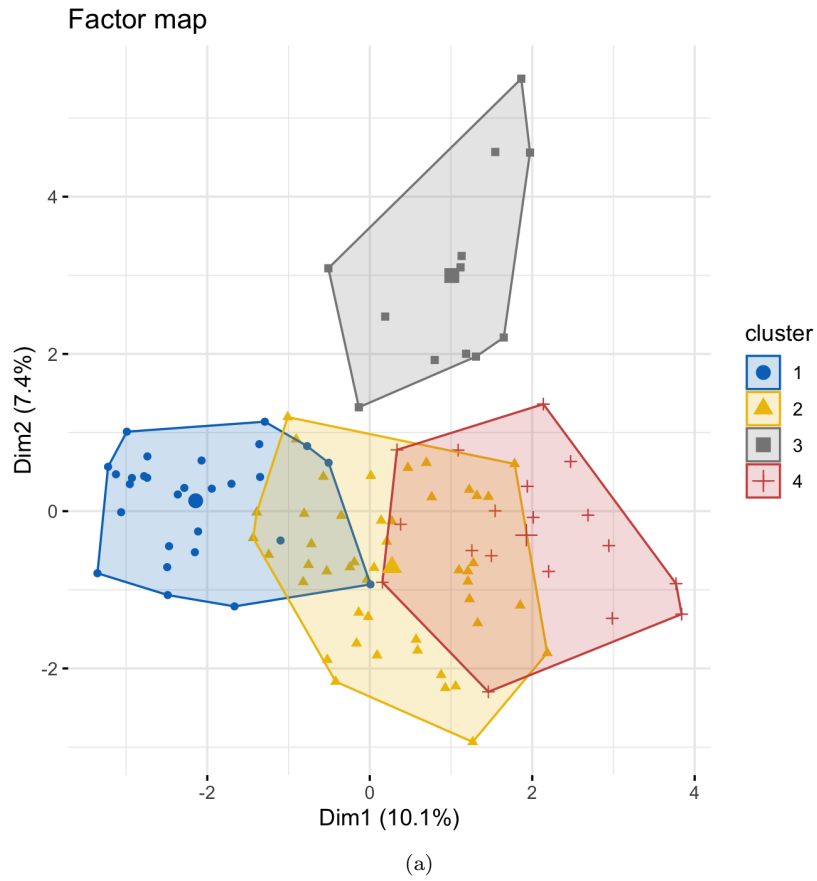


Figure 9. Clustering of student profiles.

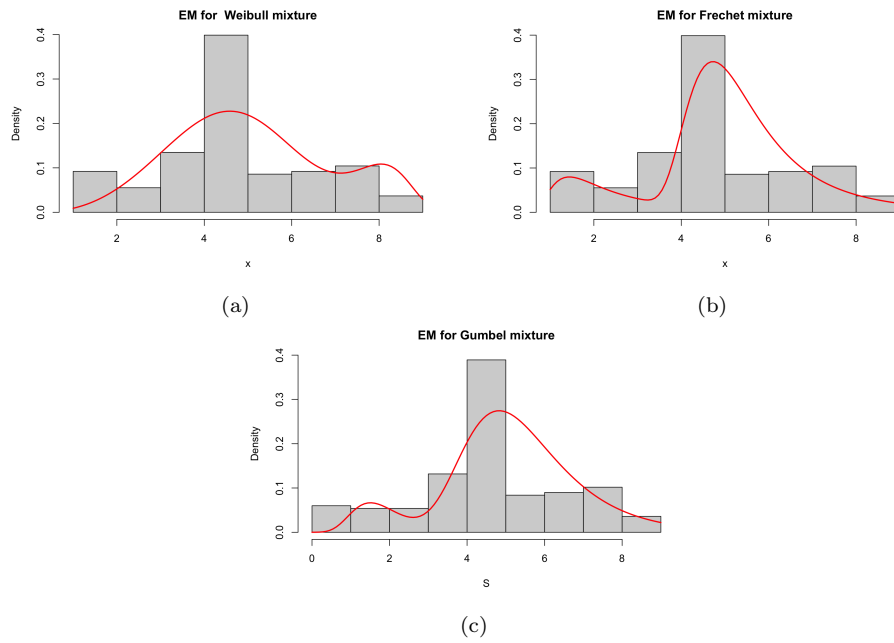
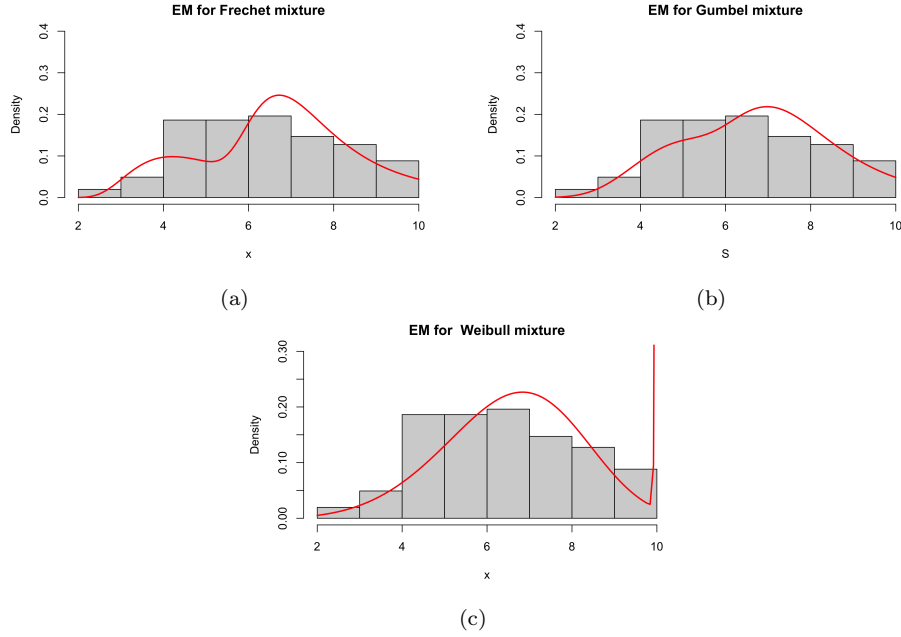
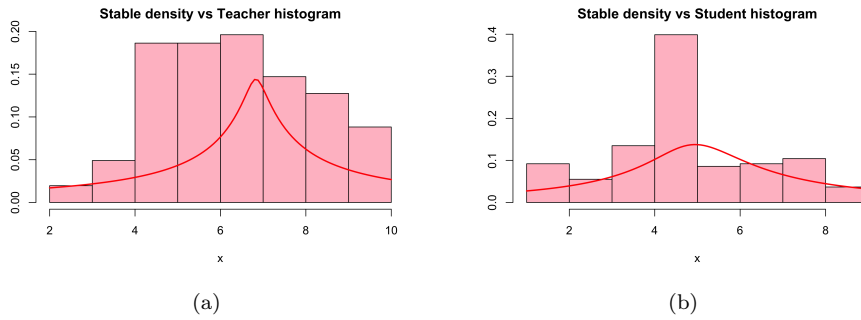


Figure 10. Mixture estimation using EM algorithm for students' Maths level distribution.



**Figure 11.** Mixture estimation using EM algorithm for teachers' Maths attitude level distribution.



**Figure 12.** Stable density Estimation.

estimation method and log-moments estimation method for stable distributions, we firstly give a simulation study result based on the performance of the underlined estimators of  $\alpha$  (see Table 5 of the **Supplementary Material**) for various sample size  $n$  and some choice of the threshold  $k$  (or largest positive values  $x$ ). Note that we are concerned with positive and independent sample data of student maths level and teachers' maths attitude. The others parameters of the strictly stable distribution was done using the log-moments estimation method, see Table 8 of the **Supplementary Material**. The estimated stable density is shown in Figure 12. We can see that  $\alpha$ -stable distributions,  $\alpha \in (0, 2)$ , are not suitable for teachers' maths attitude toward mathematics distributions.

## *Conclusion and Perspectives*

In this study, we considered two aspects of the attitude towards mathematics both for students (perception, usefulness, interest, social factors, school environment) and teachers (perception, experience). Didactic approaches through semiotic representation and contextualisation offers a modeling study from a survey data where languages issues are considered. Statistical modeling tools such as the Generalized linear model allow preliminary to check for the variable selection and importance. We examined the different cluster profiles of teachers and students obtained by the unsupervised classification method. We consider some new parameter estimation methods to study the distribution of the students' maths level and teachers' attitude level from stable distributions. We also deal with classical parameter estimation for extreme value distributions. This work provide a first view the attitudes toward mathematics modeling and allow some fruitful discussions in order to point out a teaching-learning process for a constructive engagement and the future development of the teaching profession taking into account students attitudes' toward mathematics. We are currently experiencing the use of traditional games for student to better learn some mathematical contents and for teachers to perform the teaching of mathematics in the form of Lesson study; thus improving the attitudes toward mathematics.

## *Declarations*

The authors declare that they have no competing interests. The two co-authors have seen and agree with the contents of the manuscript. We certify that the submission is an original work and is not under review at any other publication.

## SUPPLEMENTARY MATERIAL

- Table 1:** Full description of the student survey database variables.
- Table 2:** Full description of the teacher survey database variables.
- Table 3:** Results description in French and Maths for students.
- Table 4:** Teacher's choice for training in french and maths lessons.
- Table 5:** Performance of the tail exponent parameter estimator with simulated data.
- Table 6:** Explicit parameters and Log-likelihood for teachers' maths attitude level fitted distribution.
- Table 7:** Explicit parameters and Log-likelihood for students' maths level fitted distribution.
- Table 8:** Explicit parameters and Log-likelihood with a strictly stable density.
- Fig 1:** Summary of variable importance outputs for teachers' database.
- Fig 2:** Summary of variable importance outputs for students' database.
- R-codes and Data set:** The R codes and Data are available from the corresponding author on reasonable request.

## References

- Aiken Jr, L. R. (1970). Attitudes toward mathematics. *Review of educational research* 40(4), 551–596.
- Andreu, S. et al. (2021). Evaluations 2021 point d'étape cp: premiers résultats. *Ministère de l'Éducation nationale, de la jeunesse et des sports*.

- Ashkenazi, S. and H. Eisner (2022). Problem-solving among math anxious individuals: The role of advanced strategy and testing of online anxiety. *European Journal of Education and Pedagogy* 3(4), 14–21.
- Barrier, T. and V. Durand-Guerrier (2017). La quantification au cœur des relations entre langage, raisonnement et apprentissages mathématiques. In *Actes du 22e colloque de la CORFEM*. CORFEM.
- Bernier, J. R. (2020). *Understanding social factors in small group work in undergraduate mathematics classrooms*. The University of Maine.
- Bingham, N. H. and J. M. Fry (2010). *Regression: Linear models in statistics*. Springer Science, Business Media.
- Borasi, R. (2014). The invisible hand operating in mathematics instruction: Students conceptions and expectations. In *Problem Posing*, pp. 83–91. Psychology Press.
- Brousseau, G. (2006). *Theory of didactical situations in mathematics: Didactique des mathématiques, 1970–1990*, Volume 19. Springer Science & Business Media.
- De Saussure, F. (1989). *Cours de linguistique générale*, Volume 1. Otto Harrassowitz Verlag.
- Delcroix, A., T. Forissier, and F. Anciaux (2013). Vers un cadre d’analyse opérationnel des phénomènes de contextualisation didactique.
- Dempster, A. P., N. M. Laird, and D. B. Rubin (1977). Maximum likelihood from incomplete data via the em algorithm. *Journal of the royal statistical society: series B (methodological)* 39(1), 1–22.
- Dias, C. M. (2020). Implications didactiques de l’appropriation du français sur une année scolaire, par les élèves allophones. *Allophonie. Inclusion et langues des enfants migrants à l’école*, 187–201.
- Duval, R. (1993). Registres de représentation sémiotique. In *Annales de didactique et de sciences cognitives*, Volume 5, pp. 37–65.
- Faure, K. M. (2020). Analyse quantitative et qualitative des difficultés rencontrées par les élèves allophones dans leurs apprentissages mathématiques.
- Hache, C. (2019). *Questions langagières dans l’enseignement et l’apprentissage des mathématiques*. Ph. D. thesis, Université Paris Diderot.
- Hernandez-Martinez, P. and P. Vos (2018). “why do i have to learn this?” a case study on students’ experiences of the relevance of mathematical modelling activities. *ZDM* 50, 245–257.
- Houdement, C. (2017). Résolution de problèmes arithmétiques à l’école. *Grand N, Revue de mathématiques, de sciences et technologie pour les maîtres de l’enseignement primaire* 100.
- Houdement, C. (2022). Promenade sémiotique et mathématique.
- Huet, S., A. Bouvier, M.-A. Poursat, E. Jolivet, and A. Bouvier (2004). *Statistical tools for nonlinear regression: a practical guide with S-PLUS and R examples*. Springer.
- Jolliffe, I. (2005). Principal component analysis. *Encyclopedia of statistics in behavioral science*.
- Julo, J. (1995). *Représentation des problèmes et réussite en mathématiques: un apport de la psychologie cognitive à l’enseignement*. FeniXX.
- Julo, J. (2002). Des apprentissages spécifiques pour la résolution de problèmes? *Grand N n* 69.
- Kansanen, P. and M. Meri (1999). The didactic relation in the teaching-studying-learning process. *Didaktik/Fachdidaktik as Science (-s) of the Teaching profession* 2(1), 107–116.
- Kotz, S. and S. Nadarajah (2000). *Extreme value distributions: theory and applications*. world scientific.
- Kuruoglu, E. E. (2001). Density parameter estimation of skewed/spl alpha/-stable distributions. *IEEE Transactions on signal processing* 49(10), 2192–2201.
- Laborde, C. (1982). *Langue naturelle et écriture symbolique: deux codes en interaction dans l’enseignement mathématique*. Université scientifique et médicale, Institut national polytechnique de Grenoble.
- Laborde, C. (1991). Lecture de textes mathématiques par des élèves (14-15 ans): une expérimentation. *Petit x* 28, 57–90.
- Li, J. and A. Barron (1999). Mixture density estimation. *Advances in neural information*

- processing systems 12.*
- Ma, X. and N. Kishor (1997). Assessing the relationship between attitude toward mathematics and achievement in mathematics: A meta-analysis. *Journal for research in mathematics education*, 26–47.
- Mazana, Y. M., C. Suero Montero, and C. R. Olifage (2019). Investigating students’ attitude towards learning mathematics.
- Michael, I. (2015). *Factors leading to poor performance in mathematics subject in Kibaha secondary schools*. Ph. D. thesis, The Open University Of Tanzania.
- Nguala, J.-B. (2005). La multireprésentation, un dispositif d’aide à la résolution de problèmes. *Grand N, Revue de mathématiques, de sciences et technologie pour les maîtres de l’enseignement primaire*.
- Nguala, J.-B. (2023). Contextes sémantiques et dispositifs de multireprésentation de problèmes: enjeux, structuration du milieu didactique, portée et limites. *Les pratiques d’enseignement et de formation à l’épreuve des contextes: Questions épistémologiques et méthodologiques. Téraèdre*, 179–199.
- Nolan, J. P. (2012). *Stable distributions*.
- Patlolla, C. R. (2020). Understanding the concept of hierarchical clustering technique.
- Peteros, E., A. Gamboa, J. O. Etcuban, A. Dinauanao, R. Sitoy, and R. Arcadio (2019). Factors affecting mathematics performance of junior high school students. *International Electronic Journal of Mathematics Education* 15(1), em0556.
- Pino-Fan, L. R., A. Assis, and W. F. Castro (2015). Towards a methodology for the characterization of teachers’ didactic-mathematical knowledge. *EURASIA Journal of Mathematics, Science and Technology Education* 11(6), 1429–1456.
- Praene, J. P., B. Malet-Damour, M. H. Radanielina, L. Fontaine, and G. Riviere (2019). Gis-based approach to identify climatic zoning: A hierarchical clustering on principal component analysis. *Building and Environment* 164, 106330.
- Prasad, G. L. (2014). Portraits of plurilingualism in a french international school in toronto: Exploring the role of visual methods to access students’ representations of their linguistically diverse identities. *Canadian Journal of Applied Linguistics* 17(1), 51–77.
- Redner, R. A. and H. F. Walker (1984). Mixture densities, maximum likelihood and the em algorithm. *SIAM review* 26(2), 195–239.
- Salone, J.-J. (2022). Contexte et contextualisation à mayotte, une approche systémique. *Contextes et didactiques. Revue semestrielle en sciences de l’éducation* (20).
- Samoradnitsky, G. (2017). *Stable non-Gaussian random processes: stochastic models with infinite variance*. Routledge.
- Shaftel, J., E. Belton-Kocher, D. Glasnapp, and J. Poggio (2006). The impact of language characteristics in mathematics test items on the performance of english language learners and students with disabilities. *Educational Assessment* 11(2), 105–126.
- Sinaga, K. P. and M.-S. Yang (2020). Unsupervised k-means clustering algorithm. *IEEE access* 8, 80716–80727.
- Wakhata, R., V. Mutarutinya, and S. Balimuttajjo (2022). Secondary school students’ attitude towards mathematics word problems. *Humanities and Social Sciences Communications* 9(1), 1–11.

## Appendix

**Table 1.** Results description in French and Maths for students

<b>Variable</b>	<b>Observation</b>
Understanding phrases read by the teacher	[32.10%; 33.60%]
Read aloud words	[31%; 42%]
Read aloud a text	[31%; 41%]
Understand sentences read alone	[39.5%; 44%]
Problem solving	[20%; 25.5%]

**Table 2.** Teacher's choice for training in french and maths lessons

<b>Variable</b>	<b>Observation</b>
Learning to read	25%
Written production training	21%
Developing oral skills	16%
Reading comprehension	11%
Developing the lexicon	7%
Phonology	11%
Mental calculation	2%
Problem solving	44%
Number's construction	26%
Geometry	19.5%
Measure and object	5%
Explicit teaching	3.5%

**Table 3.** Full description of the student survey database variables

<b>Recoded Variable</b>	<b>Description</b>
Math_level	Student mean math level recorded
Makeup_class	Question related to homework reinforcement
Age	Age of the respondent
Encountered_Difficulties	Question about the difficulties encountered in maths course
Nbh_maths	Total number of teaching math hours courses received
Nbh_French	Total number of teaching french hours courses received
Parent_status	Question related to parent professional status
Maths_notions_often_explained_mother_tongue	Type of mathematical concepts explained in local languages by classmates
Understanding_by_Mother_tongue	Type of mathematical explanations received in local language
Assimilated_maths_notions	Mathematical concepts assimilated
Birth_place	Place of birth
Commune	Commune of residence
Post_high_school_p	Post baccalaureate study plans
Mother_tongue	Type of the local mother tongue
Internet_access	Does the student have access to the internet?

**Table 4.** Full description of the teacher survey database variables

Recoded Variable	Description
Experience	Experience
Math_level_and_report_as_teacher	Teacher's attitude level towards mathematics
liked_math_notion_as_student	Mathematical contents, teachers enjoyed as students
Difficulty_math_notion_as_student	Mathematical concepts, teachers found difficult as students
Math_notions_like_to_teach	Mathematics concepts teachers like to teach
Math_notions_difficult_to_teach	Mathematics concepts, teachers found difficult to teach
Professional_training_desired_as_teacher	Mathematics training desired by teachers
Teaching_preparation_time	Teaching preparation time

**Table 5.** Performance of the tail exponent parameter estimator with simulated data

n	k= x% n	$\alpha$	$\hat{\alpha}$
100	x=15	0.5	<b>0.55</b>
150	x=15	0.5	<b>0.52</b>
250	x=20	0.5	<b>0.49</b>
1000	x=20	0.5	<b>0.50</b>
100	x=15%	1.5	<b>1.53</b>
250	x=25	1.5	<b>1.49</b>
1000	x=20	1.5	<b>1.58</b>

**Table 6.** Explicit parameters and Log-likelihood Estimation (LLE) for teachers' maths attitude level fitted distribution

Law	Parameters	LLE
Weibull	p=0.8242; $scale_1=5.008906$ ; $shape_1= 3.597334$ ; $scale_2=8.1768$ ; $shape_2= 12.7306$	-177.761
Frechet	p=0.8675; $scale_1=3.3841$ ; $shape_1= 1.6743$ ; $scale_2=7.2974$ ; $shape_2= 15.9227$	-219.784
Gumbel	p=0.9074; $location_1=3.9914$ ; $scale_1= 1.7145$ ; $location_2=7.4265$ ; $scale_2= 0.5208$	-208.447

**Table 7.** Table S7: Explicit parameters and Log-likelihood (LLE) for students' maths level fitted distribution

Law	Parameters	LLE
Weibull	p=0.8241; $scale_1=5.0087$ ; $shape_1= 3.5976$ ; $scale_2=8.1766$ ; $shape_2= 12.72805$	-325.7509
Frechet	p=0.8675; $scale_1=3.3841$ ; $shape_1= 1.6743$ ; $scale_2=7.2974$ ; $shape_2= 15.92271$	-394.4327
Gumbel	p=0.9073; $location_1=3.9914$ ; $scale_1= 1.7144$ ; $location_2=7.4265$ ; $scale_2= 0.5208$	-340.5867

**Table 8.** Explicit parameters and Log-likelihood (LLE) with a strictly stable density

Target distribution	Parameters	LLE
Students' maths level	$\alpha=0.7730$ ; $\beta= 0.08$ ; $\sigma=2.6823$ ; $\mu= 0$	-401.5666
Teachers' attitude level	$\alpha=0.4821$ ; $\beta= 0.008$ ; $\sigma=4.73$ ; $\mu= 0$	-482.4655

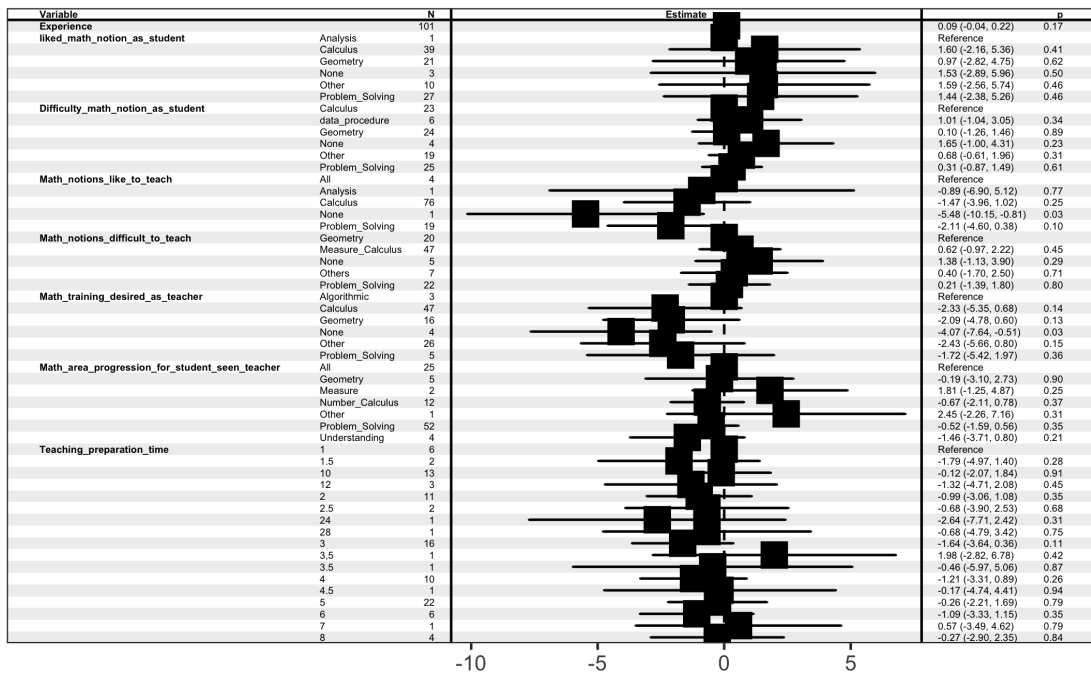


Figure 13. A view of the variable importance in the teachers' database.

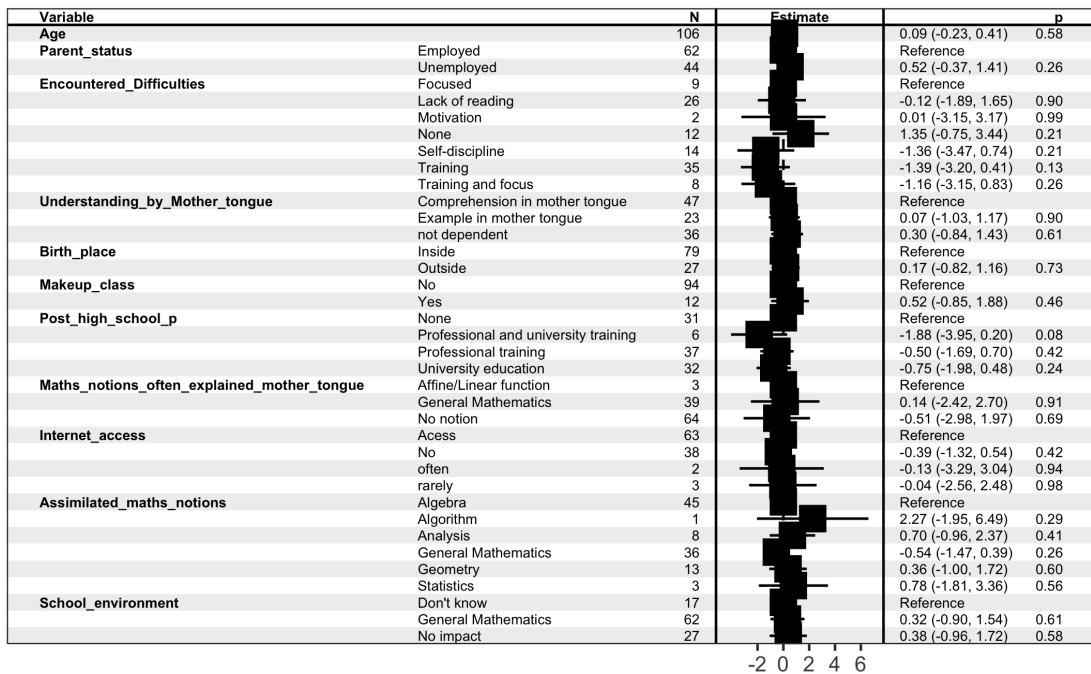


Figure 14. A view of the variable importance in the students' database.