

# Test d'équivalence

*Lucas Chabeau*

*23 avril 2019*

Jusqu'à présent nous n'avons pas pu établir de lien entre la pédagogie suivie par les enfants de notre étude et l'avancement de leurs capacités cognitives liées aux mathématiques. Si nous n'avons pu établir de lien, nous allons maintenant voir si les deux pédagogies sont significativement équivalentes sur l'assimilation de notions mathématiques (pour les élèves de moyenne section).

## 1. Calcul d'un score total pour chaque élève.

Nous avons choisi de comparer nos deux échantillons sur le score total de chaque élève au test. Pour calculer ce score, nous avons simplement fait la somme des résultats de l'enfant à chaque question. (Rappelons qu'à part la Q1, les questions ont pour résultat 0 ou 1.)

$$Score = \sum_{i=1}^{29} Q_i$$

Où  $Q_i$  est le résultat de l'élève à la question  $i$  (Donc bien-sûr dans cette formule, la question 4b par exemple ne correspond pas à  $Q_{4b}$  mais  $Q_5$ ).

Nous devons vous préciser que nous avons utilisé le même traitement pour la question 1 que lors de la création de la variable **Objet** pour éviter qu'elle n'ait trop de poids par rapport aux autres. C'est à dire que la question 1 peut prendre les scores suivants : 0, 0.3, 0.6, 0.9 et 1.2 . (0 si  $T1 \leq 3$ ; 0.3 si  $T1 \in [4; 7]$ ; 0.6 si  $T1 \in [8; 10]$ ; 0.9 si  $T1 \in [11; 16]$ ; 1.2 si  $T1 \geq 17$  ).

Le score d'un enfant peut donc aller de 0 à 29.2, leur répartition est représentée sur le graphique ci-dessous.

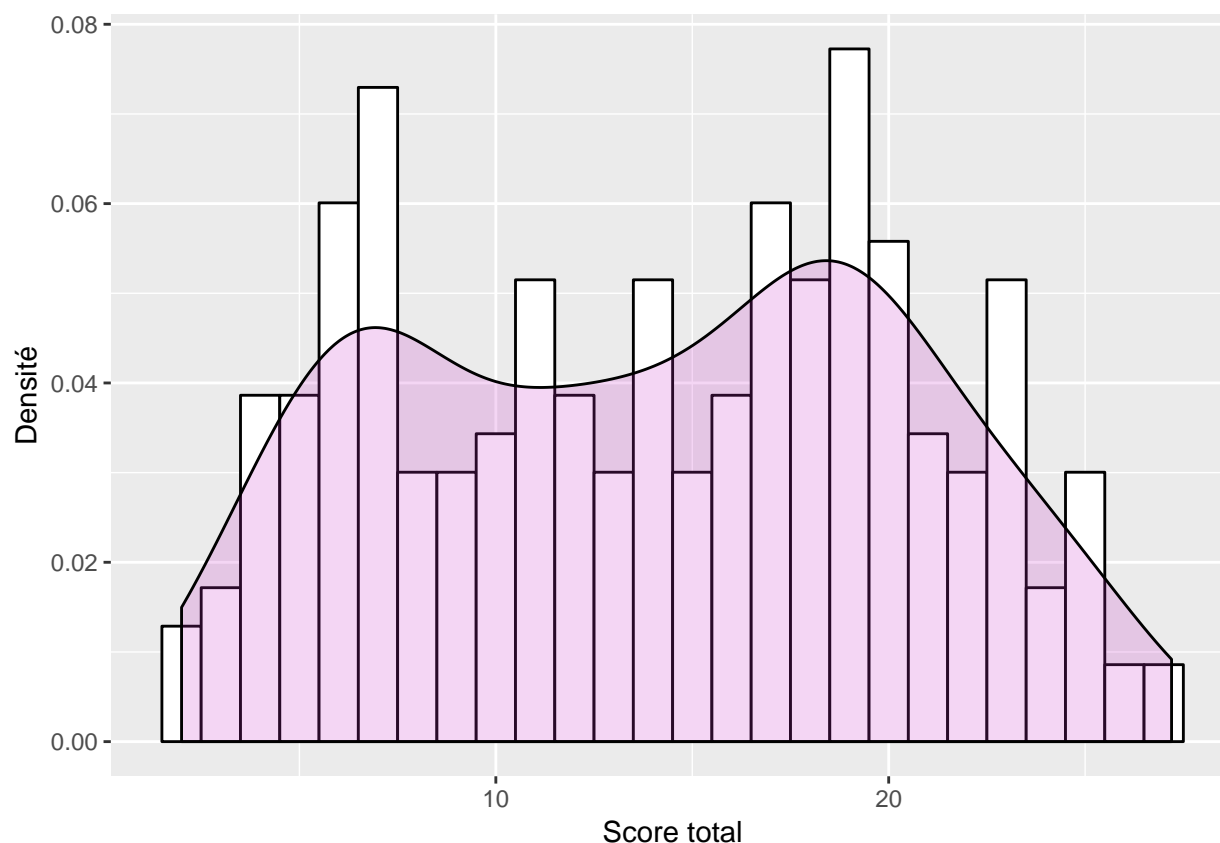


Figure 1: Répartition des scores des élèves au test

Vous pouvez voir deux pics de concentration autour des scores de 7 et de 19. Les résultats de nos élèves ne suivent pas une distribution normale. En voyant ce graphique, ceux qui cherchent absolument une différence pourraient espérer que ces deux pics correspondent chacun à une pédagogie pour en voir une plus efficace pour ce test. Ce n'est pas du tout le cas.

En effet, le graphique juste en dessous montre qu'il y a bien une forte densité autour de 7 et 19 chez les élèves conventionnels et Montessoriens. La seule différence est que nous observons un petit creux de densité entre 7 et 19 chez les élèves conventionnels quand la densité augmente légèrement chez les Montessoriens.

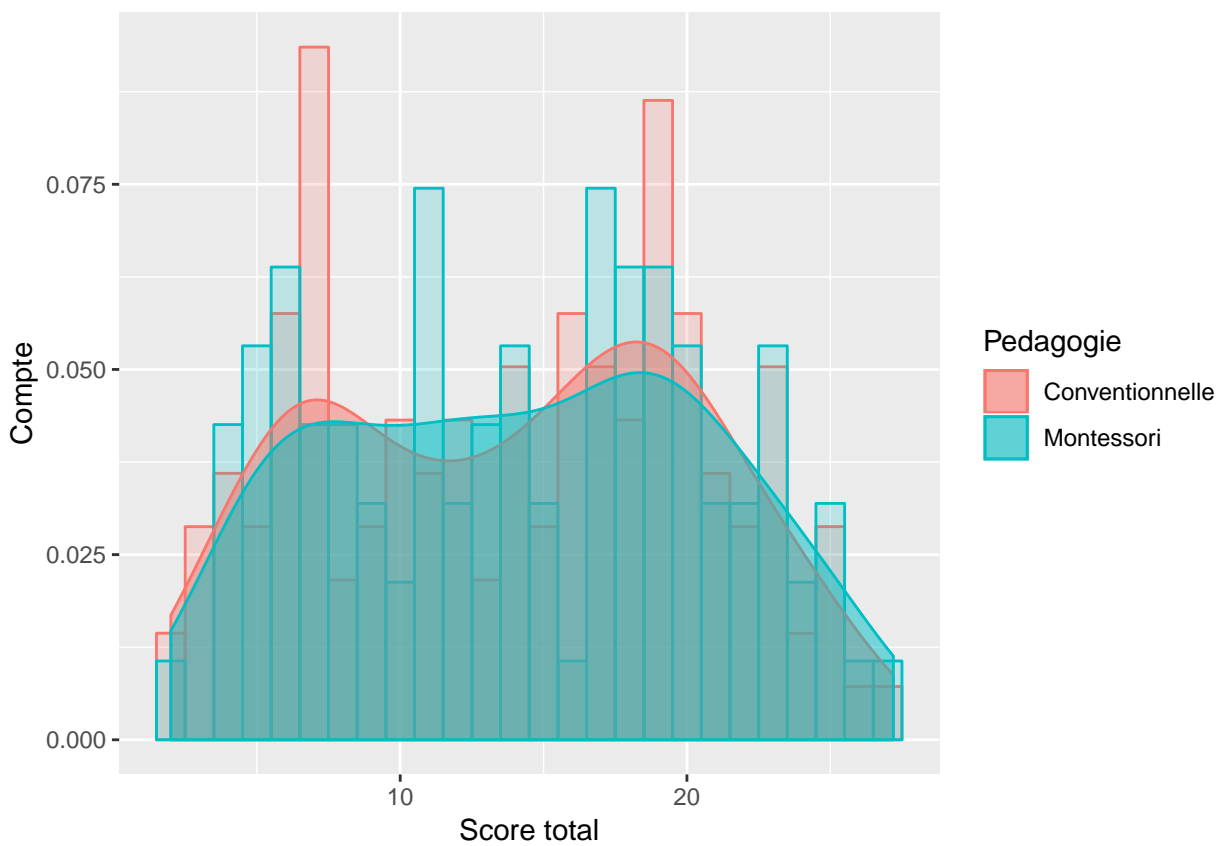


Figure 2: Répartition par pédagogie des scores des élèves au test

## 2. Réalisation du test

Notre test d'équivalence se base sur l'écart entre la moyenne des résultats des élèves conventionnels au test et celle des élèves suivants un enseignement Montessori.

Pour réaliser ce test d'équivalence, nous avons besoin de définir deux bornes (une inférieure et une supérieure). Si l'intervalle de confiance (à 95%) de l'écart entre les deux moyennes franchit au moins l'une des deux bornes, alors nous ne pourrions pas conclure à l'équivalence des deux échantillons (donc l'équivalence de résultat entre les deux méthodes d'éducation). En revanche, nous pourrions conclure que les deux méthodes sont aussi efficaces l'une que l'autre pour réussir le test cognitif qui a été soumis si l'intervalle ne franchit pas l'une des bornes.

Pour choisir ces deux bornes, nous nous sommes fiés à une "règle" qui dirait que deux élèves ont à priori un niveau différent sur une notion particulière s'ils ont deux points sur 20 d'écart à un contrôle portant sur cette notion.

Nous adaptons donc cet écart au test soumis aux élèves : avec la transformation réalisée sur la question 1, ce test peut donner une note maximale de 29. Un écart de 2.9 serait donc révélateur d'une différence de niveau sur les notions abordées.

Une fois le test réalisé (résultats juste après), nous obtenons une IC 95% de notre écart à égal à  $[-2.025; 1.417]$  et une p-valeur  $< 0.2\%$  c'est à dire que nous pouvons rejeter l'hypothèse de différence entre les deux échantillons. Nous concluons donc de ce test **que suivre la pédagogie Montessori ou la pédagogie conventionnelle n'aura pas d'effet en moyenne section de maternelle sur l'assimilation des thèmes abordés dans le test soumis par l'équipe du projet Cogmont**

Les résultats du test statistique :

```
## TOST results:
## t-value lower bound: 2.97      p-value lower bound: 0.002
## t-value upper bound: -3.67     p-value upper bound: 0.0002
## degrees of freedom : 231
##
## Equivalence bounds (raw scores):
## low eqbound: -2.9
## high eqbound: 2.9
##
## TOST confidence interval:
## lower bound 95% CI: -2.025
## upper bound 95% CI: 1.417
##
## NHST confidence interval:
## lower bound 97.5% CI: -2.275
## upper bound 97.5% CI: 1.666
##
## Equivalence Test Result:
## The equivalence test was significant,  $t(231) = 2.971$ ,  $p = 0.00164$ , given equivalence bounds of  $[-2.900; 2.900]$ 
##
##
## Null Hypothesis Test Result:
## The null hypothesis test was non-significant,  $t(231) = -0.348$ ,  $p = 0.728$ , given an alpha of 0.025.
##
##
## Based on the equivalence test and the null-hypothesis test combined, we can conclude that the observed difference between the two groups is not statistically significant.
```

##

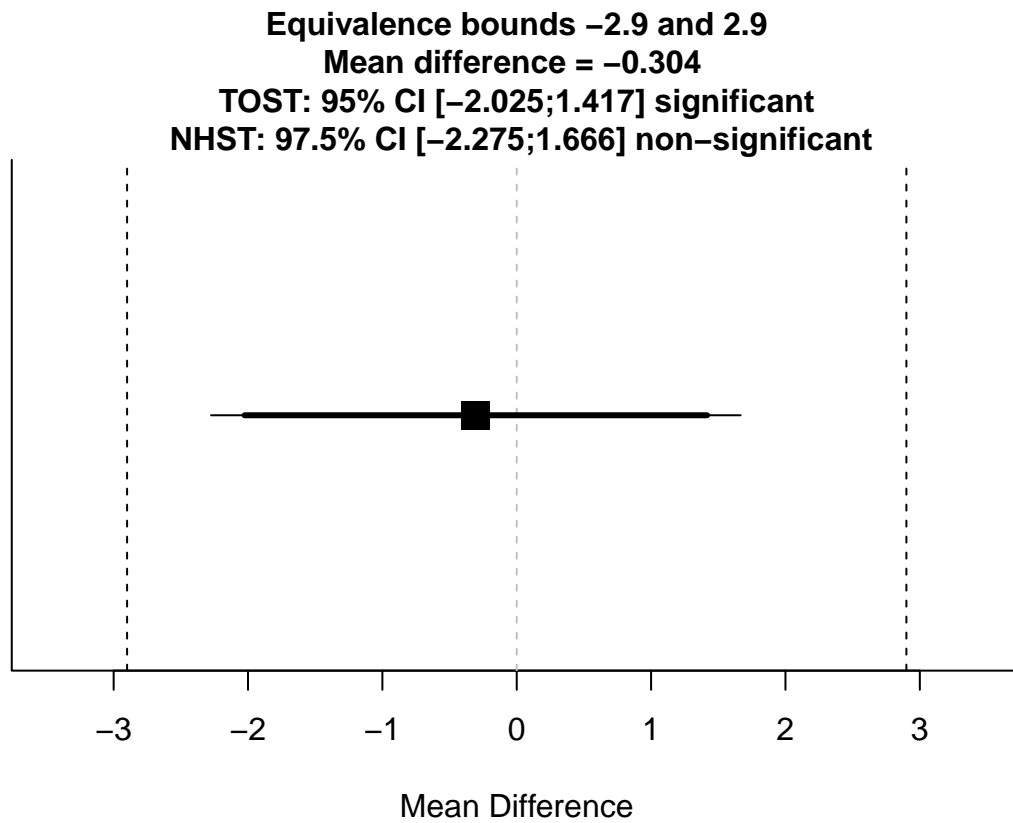


Figure 3: Résultat du test d'équivalence des moyennes des scores au test soumis par l'équipe