

Moyennes avec absences



Enjeu :

Ce document propose une étude d'un problème lié à l'observation suivante : lorsqu'il y a des absents à un devoir, la moyenne de classe n'est plus égale à la moyenne des moyennes des devoirs. On peut observer qu'un « mauvais » élève, lorsqu'il est absent, a tendance à rendre la moyenne de classe plus basse (que la moyenne des moyennes des devoirs), alors qu'un « bon » élève absent rend plutôt la moyenne de classe plus élevée. Cette observation, un peu contre-intuitive, mérite d'être précisée mathématiquement.

Cadre, définitions et résultat initial :

On se donne une classe ayant un nombre quelconque d'élèves ayant composé plusieurs devoirs (avec coefficients). On peut considérer plusieurs moyennes et les liens qu'elles ont entre elles.

La moyenne d'un devoir est la moyenne (arithmétique) des notes obtenues par les élèves présents pour le devoir. C'est une formule de la forme :

$$\bar{d} = \frac{g_1 + g_2 + \cdots + g_E}{E} = \frac{\sum_{j=1}^E g_j}{E}$$

Où les g_j sont les notes (grade en anglais) des élèves, et E est le nombre d'élèves ayant composé le devoir.

La moyenne d'un élève est la moyenne (arithmétique) pondérée des notes obtenues par l'élève. Elle est donnée par une formule de la forme :

$$\bar{e} = \frac{n_1 \times g_1 + n_2 \times g_2 + \cdots + n_k \times g_k}{N} = \frac{\sum_{i=1}^k n_i \times g_i}{N}$$

Où les n_i sont les coefficients des devoirs, les g_i sont les notes de l'élève, k est le nombre de notes de l'élève et $N = n_1 + n_2 + \cdots + n_k$ est la somme des coefficients des devoirs.

Nous comparons dans ce document la moyenne des moyennes des devoirs avec la moyenne des moyennes des élèves.

La moyenne des moyennes des devoirs est obtenue comme la moyenne pondérée des moyennes des devoirs. C'est comme si l'on considère un élève (fictif) moyen, dont la note à chaque devoir est exactement la moyenne de ce devoir, et que l'on calcule la moyenne de cet élève moyen. Autrement dit, la moyenne des moyennes des devoirs est **la moyenne de l'élève moyen**.

La moyenne des moyennes des élèves est la moyenne d'un devoir (fictif) moyen où chaque élève aurait eu une note qui est exactement sa moyenne. Autrement dit, la moyenne des moyennes des élèves est **la moyenne du devoir moyen**. C'est ce que l'on appelle habituellement (et que l'on utilise comme) **la moyenne de classe**.

Lemme 1

Lorsqu'il n'y a aucune absence, la moyenne des moyennes des élèves est égale à la moyenne des moyennes des devoirs. Autrement dit, la moyenne de classe est égale à la moyenne de l'élève moyen.

Démonstration : Lorsqu'il n'y a pas d'absence,

- le nombre d'élèves pour chaque devoir est toujours le même E ;
- le nombre de notes d'un élève est toujours le même k (c'est aussi le nombre de devoirs).

On note $g_{i,j}$ la note au devoir i de l'élève j (qui existe donc toujours).

Notons $\bar{d}_1, \bar{d}_2, \dots, \bar{d}_k$ les moyennes des devoirs.

$$\text{On a } \bar{d}_i = \frac{g_{i,1} + g_{i,2} + \dots + g_{i,E}}{E} = \frac{\sum_{j=1}^E g_{i,j}}{E}.$$

Notons $\bar{e}_1, \bar{e}_2, \dots, \bar{e}_E$ les moyennes des élèves.

$$\text{On a } \bar{e}_j = \frac{n_1 \times g_{1,j} + n_2 \times g_{2,j} + \dots + n_k \times g_{k,j}}{N} = \frac{\sum_{i=1}^k n_i \times g_{i,j}}{N}.$$

La moyenne de classe est :

$$\begin{aligned} \bar{d} &= \frac{\sum_{j=1}^E \bar{e}_j}{E} \\ &= \frac{\sum_{j=1}^E \frac{\sum_{i=1}^k n_i \times g_{i,j}}{N}}{E} \\ &= \frac{\sum_{j=1}^E \sum_{i=1}^k n_i \times g_{i,j}}{EN} \quad \left(\times \frac{N}{N} \right) \end{aligned}$$

La moyenne de l'élève moyen est :

$$\begin{aligned} \bar{e} &= \frac{n_1 \times \bar{d}_1 + n_2 \times \bar{d}_2 + \dots + n_k \times \bar{d}_k}{N} \\ &= \frac{\sum_{i=1}^k n_i \times \bar{d}_i}{N} \\ &= \frac{\sum_{i=1}^k n_i \times \frac{\sum_{j=1}^E g_{i,j}}{E}}{N} \\ &= \frac{\sum_{i=1}^k n_i \times \sum_{j=1}^E g_{i,j}}{EN} \quad \left(\times \frac{E}{E} \right) \\ &= \frac{\sum_{i=1}^k \sum_{j=1}^E n_i \times g_{i,j}}{EN} \quad (\text{développement}) \\ &= \frac{\sum_{j=1}^E \sum_{i=1}^k n_i \times g_{i,j}}{EN} \quad (\text{commutation}) \\ &= \bar{d} \end{aligned}$$

Propriétés sur l'absence :

Lemme 2

La moyenne de classe n'est pas changée quand on remplace une absence d'un élève par sa moyenne (sur les autres devoirs).

Démonstration : Puisque la moyenne de classe est la moyenne des moyennes des élèves, et comme l'ajout d'une note pour un élève ne modifie pas celle des autres, il suffit de vérifier que la moyenne de l'élève n'est pas modifiée.

On note $\bar{e} = \frac{\sum_{i=1}^k n_i \times g_i}{N}$ sa moyenne actuelle. Soit n le coefficient du devoir où il était absent.

Soit \bar{e}' la nouvelle moyenne en utilisant \bar{e} comme note au devoir.

Alors :

$$\bar{e}' = \frac{n \times \bar{e} + \sum_{i=1}^k n_i \times g_i}{n + N} = \frac{n \times \bar{e} + N \times \bar{e}}{n + N} = \frac{(n + N)\bar{e}}{n + N} = \bar{e}.$$

Ainsi, remplacer l'absence par la note \bar{e} ne change pas la moyenne de classe. Cependant cela change la moyenne du devoir :

Proposition

La moyenne d'un devoir est augmentée de $\frac{\bar{e} - \bar{d}}{E + 1}$ quand on remplace une absence d'un élève par sa moyenne \bar{e} sur ses autres devoirs, où \bar{d} est la moyenne du devoir avec l'absence et E est le nombre d'élèves ayant une note au devoir.

Démonstration : Notons $\bar{d} = \frac{\sum_{j=1}^E g_j}{E}$ la moyenne du devoir avec l'absence de l'élève.

En remplaçant l'absence par la note \bar{e} , la moyenne du devoir devient :

$$\bar{d}' = \frac{\bar{e} + \sum_{j=1}^E g_j}{1 + E} = \frac{\bar{e} + \bar{d} \times E}{1 + E} = \frac{\bar{e} + \bar{d} \times E}{E + 1}.$$

La différence des deux moyennes est alors :

$$\bar{d}' - \bar{d} = \frac{\bar{e} + \bar{d} \times E}{1 + E} - \bar{d} = \frac{\bar{e} + \bar{d} \times E - \bar{d} \times (E + 1)}{E + 1} = \frac{\bar{e} - \bar{d}}{E + 1}.$$

Corollaire

Avec les notations de la proposition précédente, l'impact sur la moyenne de l'élève moyen est donc un écart de : $\frac{n}{N} \times \frac{\bar{e} - \bar{d}}{E + 1}$, où :

- n est le coefficient du devoir ;
- N est la somme des coefficients des devoirs.

Démonstration : Notons $\epsilon = \frac{\bar{e} - \bar{d}}{E + 1}$, \bar{e} la moyenne de l'élève moyen avec l'absence, donc contenant

la moyenne \bar{d} et \bar{e}' la moyenne de l'élève moyen avec la note \bar{e} , donc contenant la moyenne $\bar{d}' = \bar{d} + \epsilon$.

On a quelque chose comme :

$$\bar{e}' = \frac{n \times \bar{d}' + \sum_{i=1}^k n_i \times \bar{d}_i}{N} = \frac{n \times (\bar{d} + \epsilon) + \sum_{i=1}^k n_i \times \bar{d}_i}{N} = \frac{n \times \epsilon + n \times \bar{d} + \sum_{i=1}^k n_i \times \bar{d}_i}{N} = \frac{n}{N} \epsilon + \bar{e}.$$

L'écart est donc bien $\frac{n}{N} \epsilon$.

On voit alors que si l'élève a une moyenne inférieure à celle du devoir, comme $\bar{e} - \bar{d} < 0$, la moyenne de classe est inférieure à la moyenne de l'élève moyen (avec l'absence), alors que si l'élève a une moyenne supérieure à celle du devoir, comme $\bar{e} - \bar{d} > 0$, la moyenne de classe est supérieure à la moyenne de l'élève moyen.

Une autre manière d'exprimer cela est qu'en ajoutant une note, précisément la moyenne \bar{e} de l'élève, inférieure à la moyenne \bar{d} du devoir, on diminue la moyenne de ce devoir, donc on diminue la moyenne de l'élève moyen, alors que la moyenne de classe ne change pas. Ainsi, on retrouve bien qu'avec l'absence, la moyenne de l'élève moyen est supérieure à la moyenne de classe.

C'est le contraire si la note ajoutée est supérieure à la moyenne du devoir.

Comme on s'attend généralement à ce que la moyenne de classe soit égale à la moyenne du devoir moyen (puisque c'est le cas quand il n'y a pas d'absences), et comme, surtout, on estime souvent (à tort, donc) la moyenne de classe par la moyenne du devoir moyen (plus facile à estimer de tête quand il y a peu de notes en comparaison du nombre d'élèves) on peut donc dire que, d'une certaine manière, l'absence d'un mauvais élève défavorise la moyenne de classe, alors que l'absence d'un bon élève la favorise.

On peut vouloir s'intéresser au remplacement de toutes les absences par les moyennes. Une formule globale permettant de calculer l'écart entre la moyenne de classe est donnée tout à la fin du document.

On peut penser effectuer des remplacements absence après absence, de manière incrémentale. Les moyennes de devoirs \bar{d} évoluent au fur et à mesure que l'on remplace l'absence des élèves par leur moyenne, ainsi que le nombre d'élèves E .

Voyons le cas de deux absences sur un même devoir :

on a déjà $\bar{d}' = \bar{d} + \epsilon$ avec $\epsilon = \frac{\bar{e} - \bar{d}}{E + 1}$ pour le premier absent.

On note $E' = E + 1$. Pour le second élève absent, on a $\epsilon' = \frac{\bar{e}' - \bar{d}'}{E' + 1}$ et $\bar{d}'' = \bar{d}' + \epsilon'$.

Au total, pour deux élèves, la différence de moyenne du devoir est $\epsilon + \epsilon'$.

On a :

$$\begin{aligned}
 \epsilon + \epsilon' &= \frac{\bar{e} - \bar{d}}{E + 1} + \frac{\bar{e}' - \bar{d}'}{E + 2} \\
 &= \frac{(E + 2)(\bar{e} - \bar{d})}{(E + 1)(E + 2)} + \frac{(E + 1)(\bar{e}' - \bar{d}')}{(E + 1)(E + 2)} \\
 &= \frac{(E + 2)(\bar{e} - \bar{d})}{(E + 1)(E + 2)} + \frac{(E + 1)(\bar{e}' - (\bar{d} + \epsilon))}{(E + 1)(E + 2)} \\
 &= \frac{(E + 2)(\bar{e} - \bar{d}) + (E + 1) \left(\bar{e}' - \left(\bar{d} + \frac{\bar{e} - \bar{d}}{E + 1} \right) \right)}{(E + 1)(E + 2)} \\
 &= \frac{(E + 2)(\bar{e} - \bar{d}) + (E + 1)(\bar{e}' - \bar{d}) - (\bar{e} - \bar{d})}{(E + 1)(E + 2)} \\
 &= \frac{E(\bar{e} - \bar{d} + \bar{e}' - \bar{d}) + 2(\bar{e} - \bar{d}) + \bar{e}' - \bar{d} - (\bar{e} - \bar{d})}{(E + 1)(E + 2)} \\
 &= \frac{E(\bar{e} + \bar{e}' - 2\bar{d}) + \bar{e} + \bar{e}' - 2\bar{d}}{(E + 1)(E + 2)} \\
 &= \frac{(E + 1)(\bar{e} + \bar{e}' - 2\bar{d})}{(E + 1)(E + 2)} \\
 &= \frac{\bar{e} + \bar{e}' - 2\bar{d}}{E + 2}
 \end{aligned}$$

Ce qui n'est pas vraiment une surprise : c'est ce que l'on obtiendrait en remplaçant les deux absences simultanément. Ainsi, autant faire simultanément les remplacements de toutes les absences à un devoir.

L'écart est alors : $\frac{\left(\sum_{p=1}^m \bar{e}_p \right) - m\bar{d}}{E + m} = \frac{\sum_{p=1}^m (\bar{e}_p - \bar{d})}{E + m}$

où m est le nombre d'absents du devoir et les \bar{e}_p sont les moyennes des élèves absents.

En faisant le remplacement de toutes les absences à tous les devoirs, la formule de l'écart entre la moyenne de l'élève moyen (initial) et la moyenne de classe (qui est la moyenne de l'élève moyen une fois tous les remplacements effectués) est :

$$\frac{\sum_{i=1}^k n_i \frac{\left(\sum_{p=1}^{m_i} \bar{e}_{p,i} \right) - m_i \bar{d}_i}{E_i + m_i}}{N} = \frac{\sum_{i=1}^k n_i \frac{\sum_{p=1}^{m_i} (\bar{e}_{p,i} - \bar{d}_i)}{E_i + m_i}}{N}$$

où :

- m_i est le nombre d'absents au devoir i ;
- E_i est le nombre d'élèves présents (notés) au devoir i
- $\bar{e}_{p,i}$ est la moyenne du p^e élève absent au devoir i .