



# Bon anniversaire à vous deux !

publié le 29/01/2021 - mis à jour le 09/04/2021

## Modélisation en probabilités

---

### Descriptif :

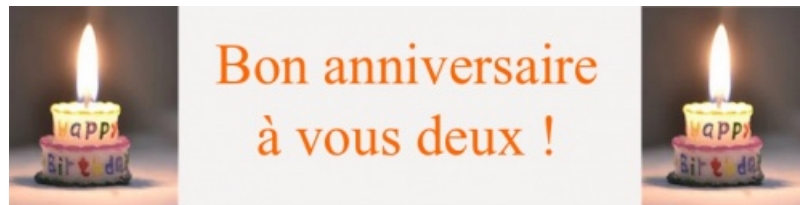
Une activité basée sur une question paradoxale par sa réponse en opposition avec l'intuition. La modélisation par les probabilités est-elle en accord avec la réalité ?

---

### Sommaire :

- [Sommaire des TraAms "modéliser"](#)
  - Description de la tâche de modélisation
  - Démarche de l'enseignant
  - Description du déroulement de la séance
  - Les élèves et la modélisation
  - Prolongement et perspectives
  - Documents de référence
  - [Retour au Sommaire des TraAms "modéliser"](#)
- 

### ● [Sommaire des TraAms "modéliser"](#)



### ● [Description de la tâche de modélisation](#)

#### ○ [Problématique](#)

Est-ce rare que deux élèves d'une classe fêtent leur anniversaire le même jour ? La modélisation par les probabilités est-elle en accord avec la réalité ?

#### ○ [Niveau concerné](#)

Classe de seconde

#### ○ [Modèle\(s\) mathématique\(s\) utilisé\(s\)](#)

Dénombrement à l'aide d'arbre, probabilités, fréquence et échantillonnage

#### ○ [Autres compétences mathématiques mises en jeu](#)

Chercher, calculer, raisonner, communiquer

#### ○ [Compétences numériques](#)

Tableur et Python

## ○ Nombre d'heures envisagées

Trois séances d'une heure

### ● Démarche de l'enseignant

Les premières questions que nous pouvons nous poser sont : pourquoi choisir une activité fondée sur une question paradoxale ? Qu'est-ce qu'un paradoxe ?

Diderot disait, « le paradoxe est une vérité opposée aux préjugés du vulgaire » et le dictionnaire mentionne : étymologiquement (para : « contre » et doxa : opinion), le paradoxe est une opinion ou une proposition à première vue surprenante ou choquante.

Apporter une réponse à une question paradoxale, c'est aller contre le sens commun et donc marquer les esprits. Cette problématique devrait donc susciter la curiosité des élèves.

Notre objectif est de répondre, de façon fiable, à la question posée qui, de surcroît, est issue du monde réel. C'est là qu'intervient l'utilisation d'un modèle. Citons David Ruelle :

*« Un modèle consiste à coller une théorie mathématique sur un morceau de réalité ».*

La solution est d'utiliser les probabilités, seul élément mathématique d'ailleurs, permettant d'étudier le hasard. Mais, il y a souvent dans l'enseignement des probabilités une confusion entre le modèle que l'on utilise et la réalité que l'on veut étudier.

Et le recours à la simulation pour faire apparaître une fréquence limite peut avoir comme conséquence de renforcer cette confusion. Les élèves n'ayant pas en effet conscience que le modèle utilisé n'est pas la réalité.

Par ailleurs, les programmes de collège abordent l'approche des probabilités cardinalistes qui suppose l'équiprobabilité des issues et l'approche fréquentiste qui assimile une fréquence observée à la « probabilité théorique » sans que l'on connaisse la précision de cette approximation.

On peut donc s'attendre à ce que les élèves soient vierges de toute réflexion à ce sujet et fassent la confusion. Mon intention est de faire émerger un questionnement sur ces approches et d'éveiller leur sens critique.

L'attendu final est de mesurer la véracité de notre modèle : est-il acceptable ? "Colle"-t-il à la réalité ?

### ● Description du déroulement de la séance

Un travail préliminaire est effectué.

Pour introduire l'activité et susciter l'intérêt des élèves par des éléments puisés dans leur environnement, chacun est amené à effectuer un travail personnel de recherche à la maison. Je leur demande de récolter des données existantes sur des groupes de personnalités de leur choix et de les déposer sur un document collaboratif dont voici un extrait du récapitulatif ([cliquer sur l'image pour l'agrandir](#)).

Groupes	Axel	Emma et Mandy	Bastien, Jana,...	Mathis et Lou	Antoine	Simcoo et Malo	Armand	Martin
Groupes	Lauréats du Prix Nobel de physique	Joueurs de l'équipe de France	Présidents français	Joueur Lauréat du ballon d'or depuis 1956	Groupes des 93 médaillés français au Jeux Olympique de Rio en 2016.	Équipe de France de 2018 lors de la coupe du monde	Président des états unis	Pilotes de F1
Taille du groupe	215	23	25	44	93	23	36	20
Nombre de jour de l'année ayant au moins deux anniversaires	38	1	0	0	11	1	1	0
	Wilhelm Conrad Röntgen : 27 mars Hendrik Antoon Lorentz : 18 juillet Pieter Zeeman : 25 mai Antoine Henri Becquerel : 15 décembre Pierre Curie : 15 mai Marie Skłodowska-Curie : 7 novembre John William Strutt Rayleigh : 12 novembre Philipp Eduard Anton von Lenard : 7 juin Joseph John	-Sarah Boubadji: 17 octobre 1986 -Solène Durand : 20 novembre 1992 -Pauline Peyraud-Magnat : 7 mars 1992 -Julie Debever : 18 avril 1988 -Sakira Karchaviani : 26 janvier 1996 -Amel Majni : 25 janvier 1993 -Giedre Štikietis : 8 novembre 2016 -Eve Périsset : 24 décembre 1994 -Wendie Renard : 20 juillet 1990	-Napoléon Bonaparte : 20 avril 1808 -Adolphe Thiers : 15 avril 1797 -Patrice de MacMahon : 13 juin 1808 - Jules Grévy : 15 août 1807 - Sadi Carnot : 11 août 1837 - Jean Casimir-Perier : 8 novembre 1847 - Félix Faure : 30 janvier 1841 - Emile Loubet : 1929	-Stanley Matthews : 1er février 1915 -Alfredo Di Stefano : 4 juillet 1926 -Raymond Kopa : 13 octobre 1931 -Luis Suárez : 2 mai 1935 -Omar Sivori : 2 octobre 1935 -Josef Masopust : 9 février 1931 -Lev Yachine : 22 octobre 1929	Alexandre Leynaud (3 mai) Allison Farnham (17 mai) Adrien Dipanda (3 mai) Camille Aygüen-Saintina (21 mai) Kenin Mahé (22 mai) Gauthier Gormier (29 mai) Daniel Jérent (4 juin) Cédric Sorhaindo (7	Hugo Lloris : 26 décembre 1986 Steve Mandanda : 28 mars 1985 Alphonse Areola : 27 février 1993 Djibril Sidibé : 29 juillet 1992 Benjamin Pavard : 28 mars 1996 Sami Ujkani : 14 novembre 1993 Raphaël Varane : 25 avril 1993 Presnel Kimpembe : 13 août 1995 Adil Rami : 27 décembre 1985 Benjamin	15 septembre 1857 William Howard Taft : 27 octobre 1858 Theodore Roosevelt : 2 novembre 1865 Warren G. Harding : 4 juillet 1872 Calvin Coolidge : 10 août 1874 Herbert	Kimi Räikkönen : 17/10/79 Antonio Giovinazzi : 14/12/93 Pierre Gasly : 7/2/96 Yuki Tsunoda : 11/5/00 Esteban Ocon : 17/09/96 Fernando Alonso : 29/07/81 Sebastian Vettel : 3/7/87 Lance Stroll : 29/10/98

Aperçu du document collaboratif recensant le nombre de jours à deux anniversaires dans différents groupes de personnalités.

La séance commence par un sondage à l'aide d'un questionnaire en ligne comportant les deux questions suivantes :

1. Est-ce rare que deux élèves d'une classe fêtent leur anniversaire le même jour ?

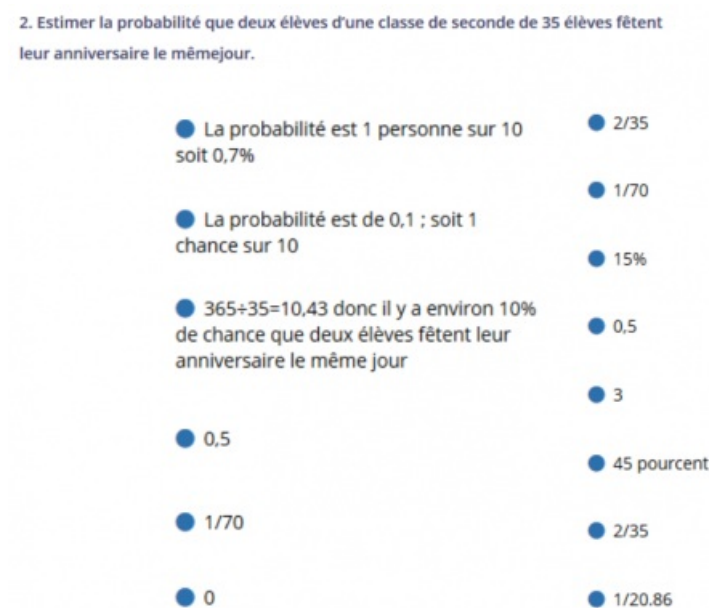
2. Estimer la probabilité que deux élèves d'une classe de seconde de 35 élèves fêtent leur anniversaire le même jour.

Voici les réponses obtenues à la question 1 :



Copie d'écran des réponses à la première question du sondage.

Et celles obtenues à la question 2 :



Copie d'écran des réponses à la deuxième question du sondage.

C'est un moyen de connaître leur opinion sur la question posée, mais aussi, d'être en capacité de confronter en fin d'activité leurs réponses aux résultats mathématiques obtenus. Nous mettrons ainsi en évidence le paradoxe.

Vient ensuite pour les élèves, un temps de réflexion personnel par écrit sur la marche à suivre, la façon de s'y

prendre pour apporter une réponse rigoureuse à la question.

Se dégagent deux types de démarches.

Un nombre conséquent cherche une valeur numérique quand d'autres assez nombreux s'orientent vers l'étude de données réelles. Enfin, pour quelques-uns, la page reste blanche. Il est donc intéressant pour moi de les accompagner vers l'obtention d'élèves virtuels et donc l'émergence d'une simulation. Nous noterons ces groupes A, B et C.

Le groupe B propose rapidement un cheminement exploitable : récolter les données au sein du lycée. Motivés, ils sont même prêts à parcourir les classes de l'établissement. Mais par commodité, je leur fournis les données préparées au préalable à l'aide de la liste que la direction du lycée m'a transmise. Ils se lancent alors dans un travail de tri, classement, et relevé des dates de naissance anonymes, rangées par ordre alphabétique et par classe, de tous leurs camarades du lycée.

Ils comptabilisent le nombre de classes de l'établissement répondant au caractère souhaité et obtiennent la fréquence de l'événement dans notre lycée. C'est un bon moyen de leur faire apprécier l'intérêt des fonctionnalités du tableur pour parvenir à leur fin plus aisément.

Cf. document suivant :

 [Fichier tableur des dates de naissances de tous les élèves du lycée](#) (OpenDocument Spreadsheet de 42.2 ko)

Fichier tableur des dates de naissances des élèves du lycée pour la piste 1

	A	B	C	D	E	F	G	H	I	J
1	Né(e) le	Classe	Mois/Année							
2	10/01/04	1BPAM-MI	01-10							
3	23/01/04	1BPAM-MI	01-23							
4	31/01/04	1BPAM-MI	01-31	0	nombre de classes			29		
5	23/02/03	1BPAM-MI	02-23	0	nombre de classes de l'événement			11		
6	03/03/04	1BPAM-MI	03-03	0	fréquence de l'événement		0.379310345			
7	03/03/04	1BPAM-MI	03-05	0						
8	07/04/04	1BPAM-MI	04-07	0						
9	16/04/04	1BPAM-MI	04-16	0						
10	25/04/03	1BPAM-MI	04-25	0	effectif moyen			26		
11	29/05/02	1BPAM-MI	05-29	0	probabilité théorique			0.65		
12	31/05/03	1BPAM-MI	05-31	0						
13	29/06/04	1BPAM-MI	06-29	0						
14	23/07/04	1BPAM-MI	07-23	0						
15	14/08/04	1BPAM-MI	08-14	0						
16	18/08/03	1BPAM-MI	08-18	0						
17	03/09/03	1BPAM-MI	09-05	0						
18	18/09/04	1BPAM-MI	09-18	0						
19	19/09/04	1BPAM-MI	09-19	0						
20	02/11/03	1BPAM-MI	11-02	0						
21	02/11/03	1BPAM-MI	11-02	1						
22	09/11/04	1BPAM-MI	11-09	0						
23	20/11/04	1BPAM-MI	11-20	0						
24	07/12/04	1BPAM-MI	12-07	0						
25	24/12/04	1BPAM-MI	12-24	0						
26	25/12/01	1BPAM-MI	12-25	0						

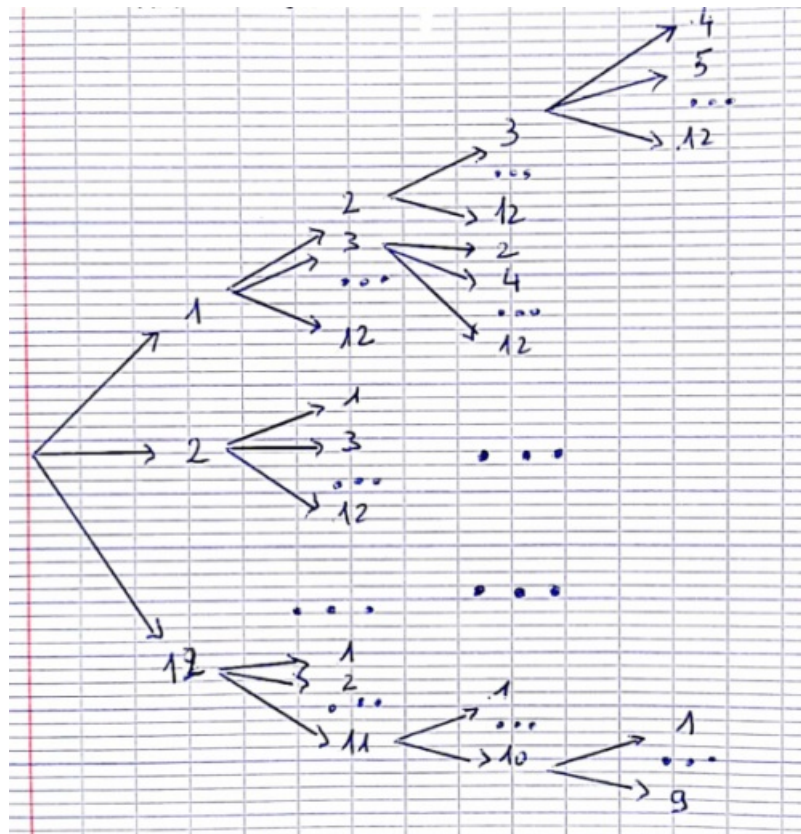
Aperçu du fichier tableur des dates de naissance des élèves du lycée, triées par ordre croissant.  
(cliquer sur l'image pour l'agrandir).

Pour le groupe A, qui cherche à calculer la probabilité, une aide s'impose. Il est nécessaire à ce stade, de guider les élèves et de leur fournir un enchaînement de questions les amenant vers la modélisation correspondante. Les élèves n'ont pas acquis les automatismes qui peuvent les conduire spontanément à modéliser. Ils ont peiné à établir un arbre qui modélise la situation y compris en restreignant l'étude, dans un premier temps, aux mois de l'année et à une classe de quatre élèves. C'est le groupe qui a nécessité le plus grand accompagnement.

Cf. activité suivante :

 [Document élève pour l'activité sur les anniversaires](#) (PDF de 68.3 ko)

Document élève pour l'activité sur les anniversaires



Arbre de probabilités envisagé pour la piste 3

En outre, par manque de temps, ils ont appliqué le programme sur Python calculant la probabilité au lieu de le mettre en place.

Quant au groupe C, obtenir des élèves virtuels n'est pas une démarche qui leur vient à l'esprit. Modéliser une classe par des nombres aléatoires n'a pas été spontané, ni dans la démarche, ni sa mise en application, l'usage des fonctions du tableur faisant là aussi réellement défaut.

Mais une fois l'idée suggérée et les outils nécessaires apportés, ils sont enthousiastes et se lancent rapidement dans la confection d'un document sur le tableur. Ils obtiennent les résultats théoriques attendus. Ayant un nombre de données important, la fréquence de l'événement est proche de la probabilité théorique.

Cf. pistes suivantes :

 [Fichier tableur pour la simulation de l'expérience aléatoire](#) (OpenDocument Spreadsheet de 229,5 ko)  
Fichier tableur pour la simulation de l'expérience aléatoire, utilisé pour la piste 2

Est venu ensuite le temps de la synthèse et de l'analyse des résultats.

Piste 4 : La réalité

- \* le travail maison est inexécuté car le nombre de groupe est insuffisant (2 groupes de personnalité et 6 groupes de A)
- \* Pour le lycée Brarly:  
Sur 23 classes, 11 sont dans A  
- fréquence :  $\frac{11}{23} \approx 0,39$

Conclusion : Dans la réalité, la fréquence est plus faible. Elle coïncide avec les avis du sondage.

Compte-rendu rédigé par le groupe ayant suivi la piste 1

Piste 2: Id 11, avec la simulation  
 on a obtenu sur le tableau un échantillon de classes.  
 → nombre de classes = 10  
 → nombre de classe de l'événement A : 42  
 → fréquence de l'événement A est :  $\frac{42}{50} = 0,84$

Cette fréquence est-elle représentative?  
 on détermine l'intervalle de fluctuation au seuil de 95%  
 la probabilité théorique obtenue avec la piste 3 est :  $p = 0,81$

$$p - \frac{1}{\sqrt{n}} = 0,81 - \frac{1}{\sqrt{50}} \quad \left| \quad p + \frac{1}{\sqrt{n}} = 0,81 + \frac{1}{\sqrt{50}} \right.$$

$$\approx 0,67 \quad \left| \quad \approx 0,95 \right.$$

L'intervalle est  $I = [0,67, 0,95]$   
 on a  $f = 0,84$ ,  $f \in I$

Donc on peut prendre en compte cette fréquence. Elle est représentative de l'échantillon de la population avec quand même un risque d'erreur d'ici au seuil de 95%.

Conclusion: Les résultats de la simulation concordent avec les réponses au quizz.

Compte-rendu rédigé par le groupe ayant suivi la piste 2

On note A l'événement : "la classe possède au moins deux éléments avec un anniversaire commun".

Piste 3: on a calculé la probabilité

Complémentaire  $P(\bar{A})$   
 le nombre d'issues de  $\bar{A}$  est :  $365 \times 364 \dots \times 331$   
 le nombre total d'issues est :  $365 \times 365 \dots 365 = 365^{35}$

Donc  $P(\bar{A}) = 1 - P(A) = 1 - \frac{365 \times 364 \dots \times 331}{365^{35}}$

Avec le programme Python on obtient  $P(\bar{A}) \approx 0,81$

Nb d'éléments de la classe	10	20	25	30	35	40	45	50	55
Probabilité de l'événement A	0,12	0,41	0,57	0,71	0,81	0,89	0,94	0,97	0,99

Conclusion:  
 Ces valeurs sont en accord avec les réponses au quizz : avec le sondage 75% des réponses sont "oui, c'est un mythe".

Compte-rendu rédigé par le groupe ayant suivi la piste 3

Bilan:

Probabilité et simulation	Réalité
→ théorique → basé sur le hasard → dépendant d'un modèle mathématique	→ ce qui est vrai → se trouve dans la vie → basé sur des faits réels
Éloigné de l'opinion	En accord avec l'opinion

Bilan rédigé en classe

● Les élèves et la modélisation

Les sollicitations préalables à la résolution du problème par le travail maison de recherche et le quiz ont pleinement

joué le rôle attendu, à savoir, immerger les élèves dans le problème et provoquer de l'enthousiasme pour le résoudre. Il est issu de la réalité, tous les élèves se sentent concernés. Ils doivent trouver un moyen d'y répondre, utiliser un modèle. Il est toujours souhaitable qu'un modèle soit issu de la réflexion des élèves. Dans notre activité, il est indéniable que le travail préliminaire a influencé certains d'entre eux dans leur réflexion, dans l'élaboration d'une démarche et dans la recherche d'un modèle.

Mais tout l'intérêt pour nous est d'obtenir les deux types d'études souhaitées pour les comparer. Et tel est bien le cas : des élèves ont travaillé dans le domaine de la réalité et d'autres ont utilisé le modèle de Bernoulli que ce soit pour calculer la probabilité de l'événement, ou, pour obtenir une fréquence s'en rapprochant par simulation avec l'instruction « ALEA.ENTRE.BORNES(1 ;365) ».

Une première réaction s'est produite avec les probabilités, issues du modèle mathématique, qui sont en opposition avec les réponses au quiz de quasiment tous les élèves,

Nombre d'élèves de la classe	10	20	25	30	35	40	45	50	55
Probabilité de l'événement A	0,12	0,41	0,57	0,71	0,81	0,89	0,94	0,97	0,99

Tableau présentant les probabilités de succès selon l'effectif du groupe

La fréquence de l'échantillon par simulation a renforcé cette opposition à leurs préjugés.

Occasion ici de soulever la question : la fréquence de notre échantillon obtenue par simulation est-elle représentative de l'ensemble ? Ce qui nous amène à l'intervalle de fluctuation.

nombre de classes n	50			
nombre de classes de l'événement A	42			
fréquence de l'événement A	0,84			
probabilité théorique p	0,81			
Intervalle de fluctuation	min	0,668578644	max	0,951421356

Copie d'écran de tableur présentant les bornes de l'intervalle de fluctuation.

Puis arrive la comparaison de ces résultats du modèle mathématiques à ceux de la réalité.

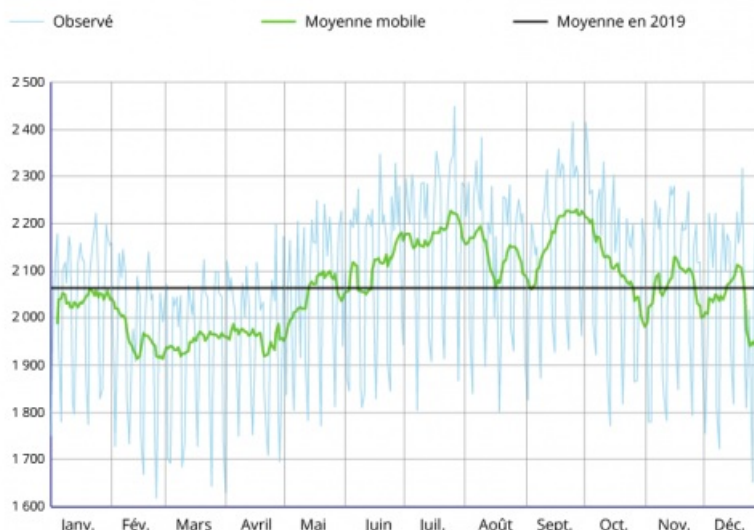
Il est regrettable de ne pas avoir pu exploiter les récoltes de données sur les personnalités à cause de la taille trop faible de l'échantillon obtenu.

En revanche, les données du lycée sont exploitables et nous obtenons la différence attendue.

nombre de classes	29
nombre de classes de l'événement	11
fréquence de l'événement	0,379310345

Copie d'écran de tableur donnant la fréquence de l'événement pour les élèves du lycée Branly

Nous pouvons également illustrer notre propos avec [les données de l'INSEE](#) donnant le nombre de naissances par jour.



Répartition des naissances en France selon les jours de l'année 2019

Le sens critique est éveillé et les questions émergent spontanément des élèves :

Pourquoi cette différence ? Qu'est-ce qui pourrait expliquer que le modèle ne « colle » pas à la réalité ?

Les élèves viennent écrire leurs propositions de réponses au tableau.

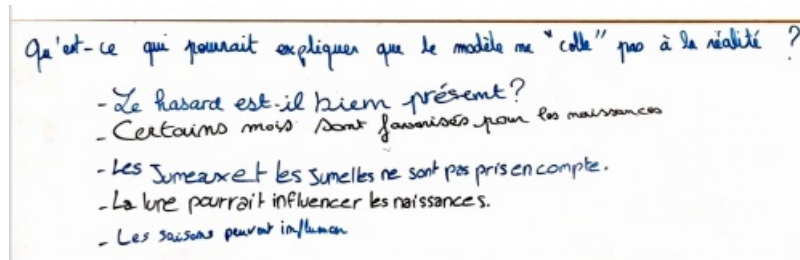


Photo du tableau de classe où sont notées les propositions d'élèves.

Pour conclure, cette activité met bien en lumière le paradoxe. De plus, elle permet de constater que les modèles mathématiques nous donnent des résultats qui ne concordent pas toujours avec les résultats réels.


En conséquence, la question de la véracité des modèles mathématiques est sous-jacente, chose rarement remise en cause dans l'enseignement, les élèves faisant généralement confiance à leur professeur.

### ● Prolongement et perspectives


Il serait intéressant de faire réfléchir les élèves sur la différence entre la stabilisation de la fréquence pour une expérience et la variation de la fréquence d'une expérience à l'autre dans le cadre d'un travail sur les fluctuations d'échantillonnage.


### ● Documents de référence

 [Document élève pour l'activité sur les anniversaires](#) (PDF de 68.3 ko)  
Document élève pour l'activité sur les anniversaires

 [Travail préparatoire de l'activité](#) (PDF de 51 ko)  
Fiche de consignes pour le travail à faire à la maison afin de préparer l'activité

 [Fichier tableur récapitulatif sur les groupes de personnalités](#) (OpenDocument Spreadsheet de 20.6 ko)  
Fichier tableur récapitulatif sur les groupes de personnalités

 [Fichier tableur des dates de naissances de tous les élèves du lycée](#) (OpenDocument Spreadsheet de 42.2 ko)  
Fichier tableur des dates de naissances des élèves du lycée pour la piste 1

 [Fichier tableur pour la simulation de l'expérience aléatoire](#) (OpenDocument Spreadsheet de 229.5 ko)  
Fichier tableur pour la simulation de l'expérience aléatoire, utilisé pour la piste 2

### ● [Retour au Sommaire des TraAms "modéliser"](#)