

Atelier Modélisation-Simulations en Lycée

Maurice Ocquidant

IUFM de Montpellier

Résumé :

L'atelier proposera un contenu de formation permettant de mener une réflexion sur les problèmes de la modélisation, de la simulation et de l'échantillonnage qui sous-tendent les programmes de statistiques du lycée et conduisent à la notion de probabilité. Cela nous conduira naturellement à l'utilisation du tableur pour aborder ces notions à travers des exercices et activités utilisables dans les classes du lycée.

Présentation de l'atelier

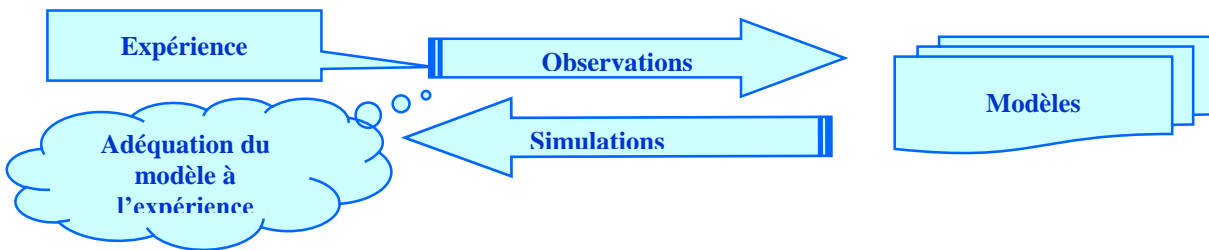
1. Les programmes

- Le chapitre des *probabilités-statistiques* a été profondément remanié dans les nouveaux programmes des Lycées (2000 en seconde...), non seulement dans ses contenus mais aussi dans son esprit.
- Dans les anciens programmes les probabilités apparaissaient comme un sous-produit de la combinatoire, avec une formule magique et presque unique : $p(A) = \text{card}(A) / \text{card}(\Omega)$. On faisait des probabilités pascaliennes : la probabilité était donnée « **a priori** » sans essayer de justifier son émergence.
- Actuellement l'esprit c'est celui des **statistiques inférentielles** où se rejoignent les statistiques et les probabilités. On fait donc du Bernoulli. La probabilité et la loi de probabilité sont données « **a posteriori** » comme limite d'une fréquence et d'une distribution de fréquence.
- C'est à l'articulation Seconde-Première que se situe la difficile tâche de faire émerger le concept de loi de probabilité à partir des distributions de fréquences de seconde.
- Il est recommandé dans les commentaires de traiter des données en nombre suffisant pour justifier une étude statistique.

2. Modélisation-Simulation

- La simulation est l'**expérimentation sur un modèle** : c'est une procédure de recherche scientifique qui consiste à réaliser une reproduction artificielle (le modèle) du phénomène à étudier puis à étudier le comportement de cette reproduction en faisant varier certains paramètres puis induire ce qui se passerait dans la réalité.
- On peut rappeler l'évolution des programmes de **Seconde** qui introduisent après des statistiques descriptives une partie permettant une première sensibilisation à la modélisation, mais sans vraiment mettre le modèle en place.

- Les notions de fluctuation d'échantillonnage et de simulation ne doivent pas faire l'objet d'un cours mais l'objectif de ce programme est de faire entrer dans l'enseignement l'observation, la modélisation et la simulation pour effectuer la démarche scientifique complète rappelée ci-dessus.
- Dans un deuxième temps, pour poursuivre la conceptualisation, il nous faudra **construire le ou les modèles** susceptibles de correspondre au problème étudié (le plus souvent il sera question de choisir une loi de probabilité mais seulement en Première).
- Dans un troisième temps, nous aurons à simuler les modèles pour tester ou vérifier leur bien-fondé. C'est le moment de la **confrontation au réel**.
- Simulations avec la touche « **Random** » de la calculatrice ou la fonction «=alea() » du tableur ou utilisation d'une **table de nombres aléatoires**.
- On va donc réintégrer dans ce chapitre une partie importante de la démarche scientifique qui est celle de l'expérimentation, des essais, des erreurs et des observations qui conduisent au modèle selon le schéma suivant.



3. Quelques illustrations

J'ai choisi de vous présenter quelques activités afin d'offrir une base et des illustrations à la critique de l'utilisation des TICE. Les situations pourront être utilisées soit dans le cadre de l'enseignement secondaire, soit dans le cadre de la future épreuve expérimentale du baccalauréat, soit dans le cadre des formations PLC1-PLC2.

*A dessin je me suis servi de deux types de logiciels différents dont l'utilisation est gravée dans le marbre des programmes officiels : le **tableur** bien entendu (Excel ici mais on peut faire l'équivalent avec Open Calc) et de **Géoplance** qui permet de donner du sens et surtout de rentrer dans les situations par les interactions ou l'aspect dynamique des situations géométriques utilisées (la fourmi qui marche sur l'arête d'un cube ou bien l'ivrogne aléatoire marchant dans New York ...). Cela a au moins l'intérêt de montrer la nécessité de la modélisation.*

a) Pile ou Face :

On commence par une situation simple avec une modélisation classique mais utilisant [Géoplan](#). Les intentions [didactiques](#) sont claires :

- Modélisation citée dans les programmes
- Répétition de la simulation facilement
- Vision simultanée de 3 traductions imagées de l'expérimentation
- Approche simplissime et très visuelle de la stabilisation des fréquences
- On peut compléter, comparer avec un [fichier tableur](#)

b) **Mendel Gregor** :

Voici un [Sujet](#) que j'ai proposé en préparation à l'épreuve sur dossier du Capes. Les objectifs :

- [Modéliser](#) une situation expérimentale liée à l'hérédité et s'appuyant sur des données recueillies il y a presque un siècle et demi
- Activité pluridisciplinaire avec une référence historique
- Un sujet à vocation formative qui tente de faire un bon tour d'horizon sur la problématique de l'adéquation qui va un peu au-delà des programmes(ici la loi n'est pas équirépartie comme le souhaite le programme)...
- Construire une distribution de χ^2 ([Voir Mendel \$\chi^2\$](#)) à l'aide du tableur et qui permet de mener l'enquête sur l'honnêteté scientifique de Mendel.

c) **La cible** :

Situation basée sur un exercice de première et terminale (voir [ici](#) les objectifs pédagogiques) qui se retrouve dans de nombreux manuels. On pourra lui donner un caractère plus expérimental en faisant retrouver le modèle qui se cache derrière la simulation.

d) **Sondages** :

J'ai essayé dans ce fichier de construire une situation expérimentale sur l'échantillonnage.

- ✓ On modélise une population (oui-non) avec choix possible de la taille ($\leq 20\,000$) et d'une probabilité théorique que l'on dissimule aux élèves.
- ✓ Il est question alors de faire des échantillonnages : nous aurons, par l'intermédiaire d'un curseur, le choix de la taille des **échantillons** ($k \leq 1500$). Ainsi nous pourrions simuler de nombreux phénomènes relevant de cette modélisation
- ✓ Pour bien visualiser, chaque expérience sera représentée par une cellule d'une plage $100 \times (n/100)$ et cette cellule aura un motif (vert=F ou ocre=P)
- ✓ On va ensuite procéder à des sondages dans cette population (échantillons de taille k) de 2 façons :
 - Par tirage global d'un bloc aléatoire de k cellules (tirages sans remise) : c'est **l'échantillon 2**
 - Par tirage aléatoire de k cellules (tirages avec remise) : c'est **l'échantillon 1**
 - Il est possible de multiplier les échantillonnages à l'aide de deux boutons.

e) **Pluie aléatoire** :

Une animation traitée avec Géoplan pour montrer que la géométrie peut aussi permettre de modéliser et de faire des statistiques.

- ✓ Pour commencer on peut [mettre en évidence](#) les pièges lorsqu'on veut faire une simulation s'appuyant sur des points aléatoires dans un disque : [exemple 1](#) ; [exemple 2](#) ;....
- ✓ Ce [premier fichier](#) permet de visualiser la situation du paradoxe de Bertrand d'une part, de retrouver la stabilisation d'une fréquence et de traiter le problème de la pluie aléatoire en offrant une approximation de π . On peut comparer ([ici](#)) avec le traitement par tableur.
- ✓ [Bertrand3](#). Il s'agit de faire comprendre le célèbre paradoxe de Bertrand et d'illustrer l'importance du choix du modèle. Voir ([ici](#)) les explications

f) **Surbooking**:

Pour finir un fichier qui essaye de traiter expérimentalement la problématique de la [surallocation de réservations](#). On peut voir les intentions didactiques ([ici](#)). C'est un fichier très technique, il faudra l'utiliser comme une boîte noire et développer une activité s'appuyant sur ce fichier informatique soit comme objectif de formation en PLC2 soit pour une expérimentation en classe, ce qui n'a jamais été fait.

g) Quelques outils

- ✓ [Loi normale](#) : permet de retrouver rapidement les paramètres des intervalles de confiance avec une bonne visualisation de la signification (remplace finalement les tables)
- ✓ [Loi binomiale](#) : c'est un peu le même esprit que l'outil précédent avec, en plus, une bonne visualisation des rapports entre loi normale et binomiale.
- ✓ [Intervalles de confiance](#) : ayant eu beaucoup de mal à comprendre les rapports entre les 3 paramètres (taille de l'échantillon, degré de confiance et fourchette), la mise au point de cet outil m'a aidé (voir des explications [ici](#))
- ✓ [Boîte à moustache](#) : je vous laisse le soin de juger de l'intérêt de cet outil
- ✓ [Générateurs de nombres \(pseudo\)aléatoires](#) : pour essayer de comprendre comment fonctionnent ces générateurs et pourquoi les nombres engendrés sont pseudo-aléatoires.

4. Conclusion

J'espère que les quelques pistes ouvertes ici vous aideront à avancer dans les différentes tâches nouvelles qui vous attendent dans vos exaltantes fonctions.

Je laisse à votre disposition les fichiers que vous avez visualisés et d'autres. Ce ne sont que des productions personnelles que vous pouvez modifier/adapter à vos besoins...il n'y a pas de copyright. Je ne souhaite qu'une chose, c'est que vous en fassiez profiter vos élèves, les étudiants, les stagiaires.