Statistique et probabilités

Journée COLLÈGE, octobre 2008

Yves Ducel, Bruno Saussereau
IREM, Université de Franche-Comté

Nous proposons ici un résumé des idées directrices de notre intervention de trois heures sur le thème "Statistique et probabilités" effectuée dans le cadre de la journée de formation des professeurs de collège aux nouveaux programmes de Troisième organisée par l'Inspection pédagogique régionale de mathématiques courant octobre 2008 dans divers collèges de l'académie de Besançon.
Ce texte est une contribution à la réflexion actuelle sur l'enseignement des probabilités au collège. À ce titre, il faut le considérer comme un document de travail susceptible d'évolution.

1 Introduction

L'utilisation des probabilités est inhérente à une démarche de modélisation d'un phénomène réel. La démarche de modélisation suppose au départ une réflexion qualitative sur la situation réelle étudiée en vue de faire ressortir des "paramètres" caractéristiques de cette situation. Ces paramètres pourront être quantifiés, c'est-à-dire identifiés à des objets mathématiques sur lesquels se développera la théorie mathématique qui constituera l'étude quantitative de la situation.

Les probabilités sont introduites au collège en classe de troisième dans les programmes appliqués dès cette année. Si les élèves ont pu aborder ce type de démarche dans les sciences physiques, c'est peut-être la première fois que cette démarche intervient de façon aussi marquée dans un enseignement assuré par le professeur de mathématiques. Dans la mesure où c'est la première fois que les élèves abordent explicitement ces notions, il ne s'agit pas de penser l'enseignement des probabilités en troisième comme une simple avancée de ce qui se fait au lycée sur ce thème, mais plutôt de préparer l'élève, par une réflexion sur l'aléatoire, à ce qu'il fera au lycée. L'enseignement des probabilités en collège doit, à notre avis, mettre en premier lieu l'accent sur la démarche qualitative pour n'aborder que dans un second temps la phase (quantitative) de le modélisation qui sera surtout développée au lycée, et, pour certains, dans l'enseignement post-bac. Cette démarche qualitative devrait permettre de faire comprendre comment on en arrive à "mettre des mathématiques" dans une situation réelle.

Les échanges entre les élèves conduits par l'enseignant à l'occasion des activités proposées en classe, doivent notamment permettre

- de confronter les représentations des élèves sur l'aléatoire,
- de rappeler et structurer le vocabulaire utilisé dans la vie courante,
- d'introduire et faire fonctionner le vocabulaire standard des probabilités,

- de mettre en place quelques définitions,
- d'établir quelques règles simples sur les probabilités.

2 Statistique

Les programmes de la classe de troisième en statistique incluent désormais les notions de quartiles ainsi qu'une sensibilisation aux paramètres de dispersion. Comme l'indiquent les programmes, "l'éducation mathématique rejoint ici l'éducation du citoyen : prendre l'habitude de s'interroger sur la signification des nombres utilisés, sur l'information apportée par un résumé statistique."

2.1 Médiane et quartiles

Quelques remarques à prendre en compte à l'occasion d'activités sur les notions de moyennes, médiane et quartiles :

- Connaissant une moyenne, on peut retrouver une valeur manquante de la série. Ce ne sera en général pas le cas pour la médiane car le calcul de la médiane n'est pas "algébrique".
- Dans le calcul d'une médiane, c'est la position dans le rangement (croissant) des valeurs de la série qui intervient. La moyenne ne donne aucune information sur sa position relative dans les données ordonnées.
- Il est important de traiter des situations où la moyenne intervient sous la forme : "Somme des valeurs de la série = moyenne × effectif".
- Faire apparaître que la moyenne est la valeur commune du caractère qu'auraient tous les individus de la population s'ils avaient la même valeur du caractère.

2.2 Que retenir ?

- L'important est de retenir que la médiane est une valeur qui permet de partager la population en deux effectifs à peu près égaux. Il y a des subtilités techniques dans sa définition, ce qui explique les définitions données dans les ouvrages en fonction du degré de formalisation accepté.

- Comme définition de référence l'enseignant peut prendre celle utilisée dans certaines filières de lycée :
  On appelle **médiane** *d'une série statistique un nombre, noté M_e, tel que :*
  - *Au moins* 50% *des valeurs de la série sont inférieures ou égales à M_e,*
  - *Au moins* 50% *des valeurs de la série sont supérieures ou égales à M_e.*

  Remarque : Dans ce cas la médiane n'est pas nécessairement une valeur de la série. Ce n'est pas nécessairement une définition à apprendre par les élèves, mais c'est à notre avis la définition qui doit servir de référence à l'enseignant pour accepter ou refuser les diverses valeurs de la médiane que les élèves sont susceptibles de proposer. Il semble surtout important que les élèves comprennent le rôle joué par la médiane par rapport à la population.

– Mêmes difficultés pour la définition des quartiles. De plus la définition des manuels n'est pas toujours cohérente avec celle retenue dans le même manuel pour la médiane :

  On appelle **premier quartile** *d'une série statistique la plus petite valeur de la série, notée Q_1, telle qu'au moins* 25% *des valeurs de la série sont inférieures ou égales à Q_1.*
  On appelle **troisième quartile** *d'une série statistique la plus petite valeur de la série, notée Q_3, telle qu'au moins* 75% *des valeurs de la série sont inférieures ou égales à Q_3.*

  Remarque : On peut accepter ces définitions en première approche compte tenu que les définitions statistiques ne font pas l'objet d'un consensus aussi marqué que pour les définitions des mathématiques classiques. Dans ce cas la population est partagée en quatre sous-populations d'effectifs pratiquement égaux par le premier quartile, la médiane (qui joue le rôle d'un second quartile) et le troisième quartile.

Le choix de la définition pour un indicateur est étroitement lié à l'utilisation pratique qui va en être faite, ceci peut expliquer les différents énoncés (pas tous équivalents) proposés dans les manuels scolaires et les changements observés lors des renouvellements de programmes comme entre les différents niveaux d'enseignements.

3 Probabilités

3.1 Quels objectifs ?

À travers les activités proposées par l'enseignant, l'enseignement des probabilités au collège doit notamment permettre aux élèves de s'interroger et de se positionner par rapport à des questions du type :

– Qu'est-ce que le hasard ? Quelles sont les situations où intervient le hasard que je connais ? Quand est-ce que j'affirme que le hasard intervient dans une situation ? Qu'est-ce qui me permet d'affirmer cela ? Qu'est-ce qui distingue une situation aléatoire d'une situation qui ne l'est pas ? Comment est généré le hasard dans une situation où il intervient ?
– Quel vocabulaire courant j'utilise quand je parle d'une situation aléatoire ? Quelle démarche me permet d'avoir des "informations" sur la situation aléatoire qui m'intéresse ?
– Peut-on "quantifier" le hasard ? Comment le faire ? Dans quelles conditions les nombres, le calcul, interviennent dans l'étude d'une situation aléatoire ? Quel lien entre les nombres et le vocabulaire que j'utilise ? Quel sens donner à la phrase "J'ai une chance sur trois d'observer A" ? Comment puis-je valider (ou vérifier) les affirmations que j'énonce ? Est-ce que je peux énoncer des règles élémentaires sur les nombres qui interviennent dans une situation aléatoire ? Est-ce que je peux trouver des ressemblances ou des analogies entre des situations aléatoires différentes ?

3.2 Quelle méthodologie ?

Un travail ou un exercice portant sur une situation aléatoire doit être l'occasion de pointer quatre étapes. Suivant l'objectif visé on peut être amené à développer davantage l'une ou l'autre de ces étapes, mais il semble utile de les identifier toutes les quatre à chaque fois.

1) **Étape 1 : L'expérience aléatoire.**
   Cette étape doit permettre à l'élève de :
   - se familiariser avec la situation aléatoire étudiée,
   - d'énoncer précisément les conditions de l'expérience aléatoire,
   - d'identifier toutes les issues de l'expérience et de convenir d'une façon de les noter.

   Cette étape est donc l'occasion notamment de se poser les questions suivantes :
   - En quoi consiste la situation aléatoire ?
   - Où intervient le hasard dans cette situation ? Comment est-il généré ?
   - Quelles sont les issues de cette expérience ?
   - Est-ce que je connais toutes les issues de l'expérience ?

   Cette étape doit se conclure par la rédaction d'un court texte décrivant les conditions de l'expérience aléatoire, la liste de toutes les issues possibles et les notations adoptées.

2) **Étape 2 : Les événements attachés à l'expérience aléatoire.**
   Cette étape a pour objectif d'approfondir la compréhension de l'expérience aléatoire. Elle doit notamment permettre à l'élève de :
   - d'imaginer des événements en relation avec l'expérience aléatoire telle qu'elle est décrite à l'étape 1,
   - de repérer et expliciter les événements qui font partie des données connues de l'exercice,
   - d'identifier les issues qui réalisent ces événements en référence à la liste établie dans l'étape 1,
   - de repérer et expliciter les événements qui sont sous-jacents aux questions proposées dans l'exercice,
   - de mettre en évidence des relations entre ces événements, notamment les événements élémentaires, incompatibles, contraires à un autre événement.
   - de convenir d'une notation pour les événement qui seront utilisés.

3) **Étape 3 : Le choix des probabilités.**
   À partir des données explicites de l'exercice ou de considérations sur les conditions de l'expérience aléatoire définie à l'étape 1 (considérations statistiques, considérations d'équiprobabilité), cette étape doit permettre à l'élève, en explicitant les raisons de

son choix, d'affecter une probabilité aux événements énoncés dans l'étape 2 susceptibles de nous intéresser pour la poursuite du travail demandé.

4) **Étape 4 : Le calcul des probabilités.**
Cette étape doit permettre à l'élève de calculer les probabilités d'autres événements, à partir des probabilités introduites à l'étape 3 et des règles élémentaires sur les probabilités introduites dans le cours.

Remarque : Les étapes 1, 2 et 3 correspondent à la mise en place du modèle probabiliste qui est choisi pour décrire la situation aléatoire étudiée. Ce modèle est ensuite mis en oeuvre dans les calculs de l'étape 4.

3.3 Une ébauche de "définitions" et de règles de calculs

Le vocabulaire et les règles de calcul élémentaires à retenir sur les probabilités doivent être mis en place et utilisés lors des échanges entre élèves pendant les activités de classe.

- Une expérience **aléatoire** est une expérience qui, quand on la répète dans les mêmes conditions, ne donne pas toujours le même résultat . Les résultats qu'on peut observer en réalisant une expérience aléatoire sont appelées les **issues** de l'expérience aléatoire. On doit pouvoir faire la liste de toutes les issues possibles d'une expérience aléatoire.
- Un **événement** désigne un énoncé (une phrase) concernant les issues de l'expérience aléatoire dont on peut dire, seulement après avoir réalisé l'expérience, s'il est vrai ou s'il est faux. Un événement se représente par la liste des issues qui rendent l'énoncé vrai. Ces issues sont dites **favorables** à l'événement.
- Un événement dont l'énoncé n'est vrai que pour une seule issue de l'expérience est appelé un événement **élémentaire**.
- Un événement qui est réalisé quelle que soit l'issue de l'expérience aléatoire est dit **certain**. Un événement qui n'est jamais réalisé quelle que soit l'issue de l'expérience aléatoire est dit **impossible**.
- Deux événements pour lesquels on ne peut pas trouver une même issue de l'expérience aléatoire qui les réalisent sont dits **incompatibles**.
- La **probabilité** d'un événement est un nombre compris entre 0 et 1 associé à l'événement. Ce nombre est choisi à partir de considérations sur l'expérience aléatoire étudiée et sur l'événement étudié.
- La somme des probabilités de tous les événements élémentaires est égale à 1.
- La probabilité d'un événement est la somme des probabilités des événements élémentaires qui le composent.
- Si la réalisation d'un événement entraîne la réalisation d'un autre événement, la probabilité du premier événement est inférieure ou égale à la probabilité du second événement.

Références

[1] BAKKER Arthur., GRAVEMEIER Koeno P.E. : An historical phenomenology of mean and median, *Educational Studies in Mathematics* (2006) 62 : 149-168, Springer

[2] BROUSSEAU Guy, BROUSSEAU Nadine, GINGER Warfield : Une expérience de premier enseignement des statistiques et de probabilités, article publié en anglais dans *The Journal of mathematical behavior of children* (2001)