Journée régionale APMEP de Bretagne Occidentale
Brest, le 13 juin 2009

L'introduction de la notion de probabilités en collège et en seconde

Questionnements didactiques

Michel Henry, CII *Statistique et probabilités*,
IREM et Université de Franche-Comté

1 - Commentaire du programme de probabilités de 3ème, rentrée 2008 (BO du 21 avril 2007) :

La notion de probabilité est abordée à partir de situations familières (pièces de monnaie, dés, roues de loteries, urnes). Certaines de ces situations permettent de rencontrer des cas pour lesquels les probabilités ne sont pas définies à partir de considérations intuitives de symétrie ou de comparaisons mais sont approximativement évaluées par les fréquences observées expérimentalement (approche fréquentiste des probabilités).

2 - La nouvelle rédaction pour la rentrée 2009

(BO du 28 août 2008) :

La notion de probabilité est abordée à partir d'expérimentations qui permettent d'observer les fréquences des issues dans des situations familières (pièces de monnaie, dés, roues de loteries, urnes, etc.).

La notion de probabilité est utilisée pour modéliser des situations simples de la vie courante. Les situations étudiées concernent les expériences à une ou **à deux épreuves.**

La dualité de la notion semble être passée dans l'implicite ?

Mais le point de vue de la modélisation est affirmé

3 - Document d'accompagnement du programme de troisième

Le projet publié en mars 2008 rappelle dans son introduction :

« ***Le langage élémentaire de la statistique (avec ses mots tels que moyenne, dispersion, estimation, fourchette de sondage, différence significative, corrections saisonnières, espérance de vie, risque, etc.) est, dans tous les pays, nécessaire à la participation aux débats publics : il convient donc d'apprendre ce langage, ses règles, sa syntaxe, sa sémantique ; l'enseignement de la statistique étant, par nature, associé à celui des probabilités, il s'agit en fait d'une 'formation à l'aléatoire'*** ».

(CREM, Rapport au Ministre, 2002).

Le document indique d'abord les « ***choix du programme*** », les différentes interprétations de la probabilité, considérations de symétries et approche fréquentiste :

« ***Les justifications solliciteront l'une quelconque des interprétations de la probabilité : interprétation fréquentiste dans sa variante 'propension' ; mais certains élèves feront certainement appel à l'interprétation épistémique, dans sa variante personnelle ou interpersonnelle ; la variante logique conduisant à faire appel au principe d'indifférence (ou de raison insuffisante)*** ».

4 - Impasses didactiques d’une définition fréquentiste de la probabilité

La formulation du programme de troisième, l’approche « fréquentiste » du programme de première des années 90 pourraient suggérer de définir la probabilité comme fréquence « stabilisée ».

En se limitant à la « *description d’expériences aléatoires simples* » sans proposer d’interprétations par des modèles théoriques, ce programme laissait la place à la confusion entre fréquence expérimentale et probabilité, conçue alors comme une fréquence limite.

Mais quel statut peut-on donner à une telle limite faisant un lien entre des relevés d’observations expérimentales et un nombre théorique ?

Pousse au crime. Termes utilisés dans certaines classes de seconde (et certains manuels de troisième !) : *Fréquence théorique*, *probabilité expérimentale* !!!

En fait la confusion entre modèle et réalité a été omniprésente dans cet enseignement des probabilités des années 90 et à l’origine de difficultés didactiques essentielles.

5 - Le point de vue de la modélisation

Ni classique, ni fréquentiste, quelle réponse donner à la demande d'une définition ?

Il faut dépasser le « langage des chances » ainsi que le débat « philosophique » entre objectivistes et subjectivistes.

Le point de vue de la modélisation réalise cet enjeu, donne des clés didactiques et contribue à la formation de la démarche scientifique : observation de la réalité - description - hypothèses - modèle abstrait - développement théorique - résolution de problèmes - interprétation dans le contexte réel - validation expérimentale.

La probabilité est axiomatiquement définie comme un objet théorique, quantifiant idéalement la possibilité d'un événement calculée a priori ou estimée expérimentalement.

Tout en proposant une approche fréquentiste, le projet de programme de première L de 1993 plaçait la modélisation comme objectif, anticipant sur la réforme des années 2000 :

« ***Il s'agit, comme dans les autres programmes, d'aborder la notion de probabilité à partir de la fréquence, mais*** **on a choisi dans cette série d'affiner l'explicitation du processus de modélisation.**

L'objet de cette partie de la formation est donc de faire découvrir, en s'appuyant sur l'expérimentation numérique, quelques notions qualitatives et quantitatives liées à la modélisation mathématique des phénomènes aléatoires….

On abordera ensuite une analyse plus quantitative permettant de dégager à partir de l'étude des fréquences et en relation avec les travaux effectués en statistique, la notion

Qu’est-ce que modéliser ?

Le programme de première de 2001 réalisait un pas didactique décisif, introduisant le point de vue de la modélisation dans l’enseignement secondaire :

« Modéliser une expérience aléatoire, c'est lui associer une loi de probabilité… Une fréquence est empirique : elle est calculée à partir de données expérimentales, alors que la probabilité d'un événement est un nombre théorique… Les distributions de fréquences issues de la répétition d'expériences identiques et indépendantes varient (fluctuent), alors que la loi de probabilité est un invariant associé à l'expérience ».

(Document d’accompagnement des programmes de première de 2001).

Définition donnée par John Von Neumann, précurseur en la matière :

« Les sciences n’essayent pas d'expliquer, c'est tout juste si elles tentent d'interpréter, elles font essentiellement des modèles. Par modèle, on entend une construction mathématique qui, à l'aide de certaines interprétations verbales, décrit les phénomènes observés. La justification d'une telle construction mathématique réside uniquement et précisément dans le fait qu’elle est censée fonctionner ».

A propos de la notion de modèle, citons David Ruelle :

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

6 - Simulations

Les programmes des années 2000 font largement appel à la simulation informatique.
Le document d'accompagnement des programmes de première précise :

> ***« Modéliser consiste à associer un modèle à des données expérimentales, alors que simuler consiste à produire des données à partir d'un modèle prédéfini. Pour simuler une expérience, on associe d'abord un modèle à l'expérience en cours, puis on simule la loi du modèle »*.**

On trouve une définition de la simulation dans l'Encyclopédie Universalis :

> ***« La simulation est l'expérimentation sur un modèle. C'est une procédure de recherche scientifique qui consiste à réaliser une reproduction artificielle (modèle) du phénomène que l'on désire étudier, à observer le comportement de cette reproduction lorsque l'on fait varier expérimentalement les actions que l'on peut exercer sur celle-ci, et à en induire ce qui se passerait dans la réalité sous l'influence d'actions analogues ».***

Il convient donc de faire d'abord le choix d'un modèle.

Les situations aléatoires considérées en classe sont en principe toutes à base d'équiprobabilité.

Les modèles utilisés pour les simulations proposées sont donc souvent implicites quand ils se réduisent à la loi uniforme discrète, produite par le générateur aléatoire de l'ordinateur ou de la calculette qui fournissent des chiffres pseudo-aléatoires supposés équirépartis.

7 - Définition mathématique moderne de la probabilité

La probabilité est un ***concept mathématique*** dont la définition a du sens au sein d'un modèle théorique.

Le modèle contemporain est le modèle de Kolmogorov (1933) :

E est une ensemble représentant les issues possibles d'une expérience aléatoire.

Les parties de E représentent les événements associés à cette expérience

On distingue une famille ***B*** de parties de E représentant les événements susceptibles d'une description, fermée pour les opérations ensemblistes de base (tribu).

Une probabilité P est une mesure (au sens de Borel) sur ***B*** comprise entre 0 et 1, vérifiant : $P(E) = 1$, $P(\emptyset) = 0$, $P(E \setminus A) = 1 - P(A)$ pour tout A de ***B***,
et pour toute famille dénombrable d'événements A_i incompatibles deux à deux
($A_i \cap A_j = \emptyset$), $P(\cup A_i) = \sum P(A_i)$.

Cette définition suppose une bonne dose de théorie de la mesure…

Alors, comment fait-on dans l'enseignement secondaire ?

8 - Les programmes des lycées à la rentrée 2008

(pour 2009, voir après)

Classe de seconde : expérimentations numériques, observations des fluctuations d'échantillonnage, simulation et distributions de fréquences

Classes de première : expérience aléatoire, vocabulaire des événements, loi de probabilité sur un ensemble fini d'issues, probabilité d'un événement, loi des grands nombres, modèle d'équiprobabilité.

Classe de terminale S : probabilités conditionnelles, indépendance, formule des probabilités totales, lois discrètes (Bernoulli, binomiale), lois continues (uniforme, exponentielle), adéquation de données à une loi équirépartie.

Dans cette progression, la définition (scolaire) de la probabilité synthétise les deux approches dans le cadre de la modélisation :

- Une **expérience aléatoire** donne lieu à n **issues possibles** notées x_i. Un **événement** est représenté par un ensemble de ces issues.

- On **modélise** cette expérience par une **loi de probabilité P**: aux x_i on fait correspondre les p_i tels que $0 \leq p_i \leq 1$ et $\sum p_i = 1$. Si $p_i = 1/n$, la loi est **équirépartie**.

- La **probabilité d'un événement** est la somme des p_i associées aux issues constituant cet événement (deuxième principe de Laplace).

9 - Pour une progression en collège

- Qu'est-ce que le hasard ? Hasards sauvages et hasards bénins reproductibles

- Peut-on le quantifier ? Observations familières (dés…) et données statistiques empiriques. Échantillons pris au hasard dans une population et fluctuations

- Des générateurs de hasard familiers : pièces, dés, roulettes, urnes… Distinction entre issues possibles et « chances » de les obtenir (dés pipés, sommes de deux dés…)

- La probabilité pour évaluer le poids des cas favorables par rapport à tous les cas possibles.

- L'équiprobabilité comme hypothèse de modèle dans les situations simples des jeux (pièces parfaites, dés équilibrés, boules indiscernables…)

- Fluctuations des fréquences et stabilisation. Usage du tableur (Exemple de l'activité « planche de Galton »)

- Caractère théorique de la probabilité : un nombre pour évaluer les « chances », approximativement mesuré par une fréquence stabilisée . Distinction entre modèle et réalité.

- Mesure expérimentale approximative d'une probabilité par la fréquence stabilisée. Punaises…

10 - Épilogue : quel programme en seconde pour 2009 ?

Retour de consultation. Fin de la partie statistique :

Échantillonnage Notion d'échantillon. Intervalle de fluctuation au seuil de 95%* pour la proportion d'un caractère dans une population. Réalisation d'une simulation.	_ Concevoir, mettre en oeuvre et exploiter des simulations de situations concrètes à l'aide du tableur. _ Exploiter et faire une analyse critique d'un résultat d'échantillonnage.	Par définition, un échantillon s'obtient par tirage avec remise. À l'occasion de la mise en place d'une simulation, on peut : _ utiliser les fonctions logiques d'un tableur ou d'une calculatrice, _ mettre en place des instructions conditionnelles dans un algorithme. L'objectif est d'amener les élèves à un questionnement lors des activités suivantes : _ l'estimation d'une proportion inconnue à partir d'un échantillon ; _ la prise de décision à partir d'un échantillon.

Suite : définition de l'intervalle de fluctuation

- L'intervalle de fluctuation au seuil de 95% pour un échantillon de taille n est l'intervalle centré autour de p où se situe, avec une probabilité égale à 0, 95, la proportion observée dans un échantillon de taille n. Cet intervalle peut être obtenu, de façon approchée, par simulation.
- Le professeur peut aussi dire aux élèves, sans que ce résultat soit exigible, que, dans la pratique, lors de l'étude d'échantillons de taille $n > 25$, si f désigne la fréquence dans l'échantillon d'un caractère dont la proportion p dans la population est comprise entre 0, 2 et 0, 8, alors f appartient à l'intervalle

  - $[p-1/\sqrt{n}\ ;\ p+1/\sqrt{n}]$

avec une probabilité d'au moins 0, 95 ; il peut, dans ce cas, faire percevoir expérimentalement la propriété.

Partie probabilités, objectifs :

Objectifs visés par l'enseignement des statistiques et probabilités à l'occasion de résolutions de problèmes, dans le cadre des probabilités, rendre les élèves capables :

_ d'étudier et modéliser des expériences relevant de l'équiprobabilité (par exemple, lancers de pièces ou de dés, tirage de cartes) ;

_ de proposer un modèle probabiliste à partir de l'observation de fréquences dans des situations simples.

_ d'interpréter des événements de manière ensembliste,

_ de mener à bien des calculs de probabilité.

Les situations étudiées concernent des expériences à une ou plusieurs épreuves.
La répétition d'expériences aléatoires peut donner lieu à l'écriture d'algorithmes (marches aléatoires).

Programme pour 2009, en attendant la suite …

CONTENUS	CAPACITÉS ATTENDUES	COMMENTAIRES
Probabilité sur un ensemble fini Probabilité d'un Événement Réunion et intersection de deux événements, formule $p(A \cup B) + p(A \cap B) = p(A) + p(B)$.	_ Déterminer la probabilité d'événements dans des situations d'équiprobabilité. _ Utiliser des modèles définis à partir de fréquences observées. _ Connaître et exploiter cette formule	La probabilité d'un événement est définie comme la somme des probabilités des événements élémentaires qui le constituent. Pour les calculs de probabilités, on utilise des arbres, des diagrammes ou des tableaux.