

Cadre : activité préparée à la maison, travaillée en classe.

Durée indicative en classe : au moins une heure.

Thème : des statistiques aux probabilités.

Objectifs : étudier rapidement un relevé statistique fait par les élèves, et découvrir comment statistiques et probabilités sont liées.

Prérequis : fréquences ; notion de probabilité.

Commentaires : attention ! Cette loi de Benford est très spectaculaire mais elle a des détracteurs, et les élèves doivent le savoir. Plus exactement, c'est une loi quantitative basée sur des faits expérimentaux observés par messieurs Newcomb (en 1881) et Benford (en 1938).

Beaucoup de mathématiciens s'en sont passionnés, les financiers par exemple l'emploient pour la détection de fraude, mais fort heureusement les comptes sont absolument vérifiés lorsqu'on a une suspicion de fraude grâce à cette loi de Benford ! En effet, comme le rapporte M. Testard dans une de ses conférences (voir diapositive 113 du document :

http://www.ac-clermont.fr/disciplines/fileadmin/user_upload/Mathematiques/pages/conferences/loiDeBenford.pdf), les séries de nombres qui répondent à la loi de Benford ne sont pas si quelconques ...

Enoncé :

1° étape : un relevé statistique commun

a) On demande aux élèves de récolter soit des pages de journaux, ..., comportant beaucoup de nombres (au moins quarante chacun).

b) En classe, on donne la définition suivante :

Définition :

Le premier chiffre significatif de 4853,746 est **4**.

Le premier chiffre significatif de 0,003911 est **3**.

On demande à chaque élève de remplir le tableau ci-dessous pour le plus grand nombre possible de nombres récoltés (la colonne centrale sert à écrire une petite barre pour chaque chiffre significatif rencontré).

| Chiffre significatif | | total |
|----------------------|--|-------|
| 1 | | |
| 2 | | |
| 3 | | |
| 4 | | |
| 5 | | |
| 6 | | |
| 7 | | |

| | | |
|---|--|--|
| | | |
| 8 | | |
| 9 | | |

c) On rassemble les quantités de chiffres significatifs relevés dans toute la classe, pour remplir un tableau semblable au précédent.

2° étape : des fréquences aux probabilités.

Questions : quelle est la probabilité que dans notre liste de nombres rencontrés sur un ticket de caisse, une page de journal ..., le premier chiffre significatif soit le 1 ? Qu'il soit le 2 ? ...

Remarque : quand on a récolté beaucoup de nombres, il est encore plus intéressant de calculer ces probabilités (les fréquences d'apparition) d'un côté pour les nombres issus de tickets de caisse, d'un autre côté pour les nombres issus de pages de journaux, ou les populations des communes d'un pays ...

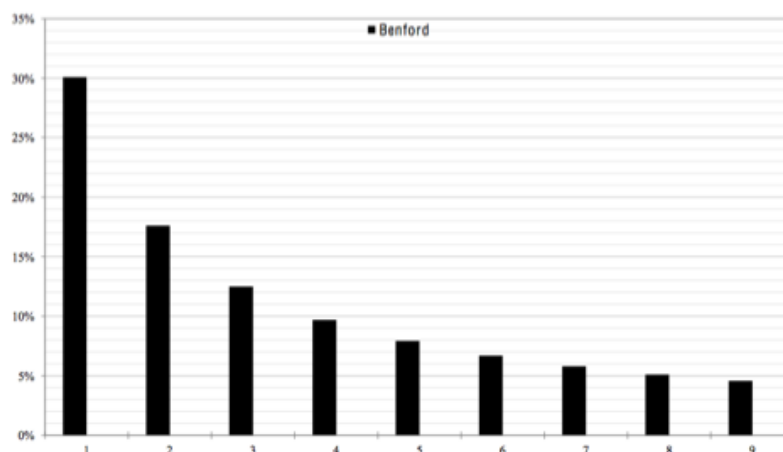
3° étape : la loi mathématique .

On fait alors découvrir aux élèves la loi de Benford (voir ci-dessous), qui donne une répartition très particulière et stable chiffres comme premier chiffre significatif des séries de nombres que l'on étudie avec les élèves, et en expliquant son utilisation dans la recherche de fraudes.

Loi de Benford

Expériences:

- Nombres extraits de coupures de journaux
- Prix relevés au hasard dans un magasin:
prix sur un assez long ticket de caisse (ou assemblage de plusieurs),
prix figurant sur une publicité, ...
- Résultats des élections présidentielles françaises 2012
- Résultats sportifs
- Nombre d'habitants de communes
- Altitudes de montagnes,
longueurs de fleuves, ...
- PIB d'un ensemble de pays
- Cours de la bourse
- Nombres extraits de comptabilité d'entreprises
- ...



Exemple de fraude détectée notamment grâce au non-respect de la loi de Benford

En 1993, Wayne J. Nelson, employé du Trésor de l'état d'Arizona, est reconnu coupable d'avoir détourné près de 2 millions de dollars en versant à des personnes fictives 23 chèques dont voici les montants:

| Date d'émission | Montants en dollars | Date d'émission | Montants en dollars |
|--------------------|---------------------|--------------------|---------------------|
| 9 octobre 1992 | 1'927.48 | 19 octobre 1992 | 96' 879.27 |
| | 27'902.31 | | 91' 806.47 |
| 14 octobre 1992 | 86' 241.90 | | 84' 991.67 |
| | 72'117.46 | | 90' 831.83 |
| | 81'321.75 | | 93' 766.67 |
| | 97'473.96 | | 88' 338.72 |
| | | | 94' 639.49 |
| 19 octobre 1992 | 93' 249.11 | | 83' 709.28 |
| | 89'658.17 | | 96' 412.21 |
| | 87'776.89 | | 88' 432.86 |
| | 92'105.83 | 71' 552.16 | |
| | 79'949.16 | | |
| | 87'602.93 | Total | 1'878'687.58 |

Indices de fraude

- 1) **Les chiffres significatifs sont à l'opposé de la loi de Benford (plus de 90 % commencent par 7, 8 ou 9).**
- 2) Valeurs d'abord petites, puis les montants et leurs fréquences ont augmenté.
- 3) Tous les montants restent inférieurs à 100'000 dollars. (Des montants supérieurs auraient sans doute dû être visés par un supérieur hiérarchique.)
- 4) Les paires de premiers chiffres 87, 88, 93 et 96 ont été utilisées deux fois dans les 23 montants.

D'après une conférence de M. Duchene pour l'APMEP dont un diaporama beaucoup plus complet est disponible sur internet.

Conclusion : comme nous l'avons vu plus haut dans les commentaires, cette loi devrait être maniée avec précaution, et pourtant elle est très utilisée et généralement imprudemment !