

Algorithme 1	Algorithme 2
Variables n, i entiers et q réel	Variables i entier, p et q réels
Entrées Saisir n	Entrées Saisir p
Initialisations q prend la valeur 1	Initialisations q prend la valeur 1, i prend la valeur 0
Traitement Pour i variant de 1 à $n - 1$ q prend la valeur $q \times \frac{365 - i}{365}$	Traitement Tant que q est supérieur à $1 - p$ q prend la valeur $q \times \frac{365 - i}{365}$
Sorties Afficher $1 - q$	i prend la valeur $i + 1$ Fin du Tant que Sorties Afficher i

On peut aussi utiliser un algorithme avec boucle et condition d'arrêt (algorithme 2) pour éviter le tâtonnement et répondre rapidement à cette question ou à une question du type : déterminer la plus petite valeur de n à partir de laquelle p est supérieur à 0,99, ou à 0,999, et observer l'évolution de n .

Le calcul précédent pourrait également être réalisé à l'aide d'un tableur.

3.4. Arbres pondérés

Les situations simples à deux épreuves ont pu être travaillées au collège à l'aide de petits arbres pondérés¹⁷. Il s'agit d'entretenir, sans aucun nouveau développement ni aucune complexification, ce type de présentation et son mode opératoire, comme l'illustre l'exemple ci-dessous. Toute connaissance sur le conditionnement est hors programme.

Exemple :

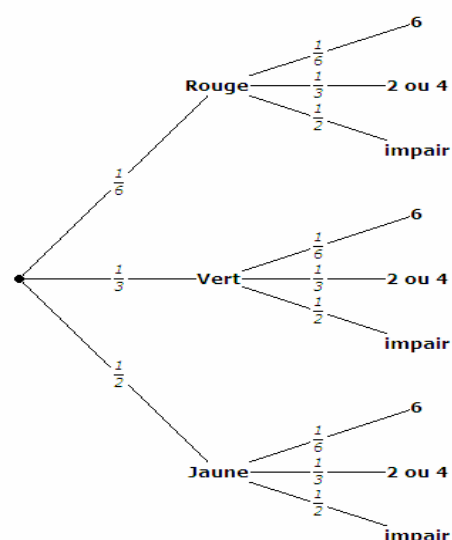
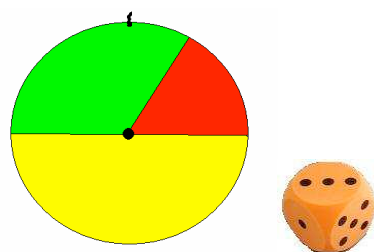
Un forain propose le jeu suivant : "À tous les coups l'on gagne"

Le joueur fait tourner une roue divisée en secteurs de mesures 60° , 120° et 180° puis il lance un dé équilibré. Il gagne un petit lot si la couleur sortie sur la roue est le vert et si le dé sort un numéro impair. Il gagne un gros lot si la couleur sortie sur la roue est le rouge et si le dé sort un six. Dans les autres cas, il gagne une pacotille.

Quelle est la probabilité de gagner un gros lot ?

Quelles sont les probabilités de gagner un lot (petit ou gros) ?

De gagner une pacotille ?



¹⁷ Cf. document ressource du collège :

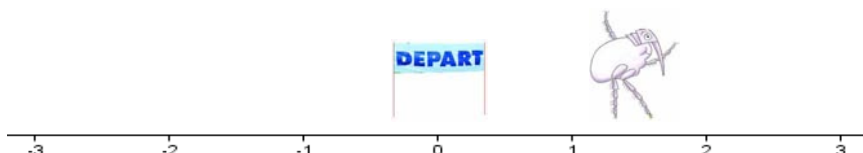
http://www.ac-grenoble.fr/maths/docresseconde/doc_ressource_clg_probabilites.pdf

L'arbre pondéré ci-dessus permet de répondre à la première question où il s'agit d'évaluer la probabilité de l'événement (R,6) : dans un arbre, la probabilité du résultat auquel conduit un chemin est égale au produit des probabilités rencontrées le long du chemin¹⁸. Cet arbre permet d'autre part de trouver facilement les formules donnant la probabilité d'une réunion ou d'un événement contraire, en leur donnant du sens.

3.5. Exemples d'algorithmes : marche aléatoire et temps moyen

Une marche aléatoire est une trajectoire constituée de pas successifs, aléatoires et indépendants. Les applications des marches aléatoires sont multiples tant en botanique (le botaniste Robert Brown, notait en 1827, le caractère apparemment erratique du déplacement de particules de pollen dans l'eau, appelé par la suite mouvement brownien), en sciences physiques (trajectoires d'une molécule dans un liquide ou un gaz), en économie (variations du cours d'une action en bourse) ou en sciences informatiques (certains moteurs de recherche utilisent des marches aléatoires pour parcourir les pages internet) etc.

Exemple : des sauts de puce¹⁹



Une puce se déplace sur un axe gradué : à chaque saut elle se déplace d'une unité, de manière aléatoire et équiprobable vers la droite ou vers la gauche. Elle part de l'origine et effectue une marche de 30 sauts.

Proposer un algorithme donnant la position d'arrivée de la puce, c'est-à-dire la position du trentième saut.

Enrichir l'algorithme précédent pour donner la liste des positions d'arrivée de N marches aléatoires.

La fonction alea fournit un nombre aléatoire dans l'intervalle $[0;1[$; l'entier x désigne la position de la puce :

```

Algorithme 1
  Variables
     $x, i$  entiers
  Initialisations
     $x$  prend la valeur 0
  Traitement
    Pour  $i$  variant de 1 à 30
      Si alea < 0,5
        alors  $x$  prend la valeur  $x + 1$ 
      sinon  $x$  prend la valeur  $x - 1$ 
      Fin du Si
    Fin du Pour
  Sorties
    Afficher  $x$ 
  
```

¹⁸ On en reste à une perception intuitive, lorsque les épreuves successives sont indépendantes (au sens où la première épreuve n'a aucune influence sur la seconde), comme cela est évoqué dans le document ressource sur les probabilités au collège en page 11 :

http://www.ac-grenoble.fr/maths/docresseconde/doc_ressource_clg_probabilites.pdf

¹⁹ Cf. une proposition de fiche élève :

http://www.ac-grenoble.fr/maths/docresseconde/fiche_eleve_marches_aleatoires.doc

On peut demander aux élèves d'examiner ce que fait l'algorithme suivant :

```

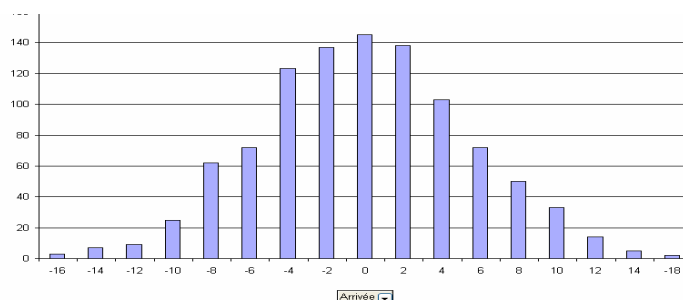
Algorithme 2
Variables
    N, S, x, i, j entiers
Entrées
    Saisir N
Initialisations
    S prend la valeur 0
Traitement
    Pour j variant de 1 à N
        x prend la valeur 0
        Pour i variant de 1 à 30
            Si alea < 0,5
                alors x prend la valeur x + 1
                sinon x prend la valeur x - 1
            Fin du Si
        Fin du Pour
        S prend la valeur S+x
    Fin du Pour
Sorties
    Afficher S/N
    
```

Enfin, on peut s'interroger sur la fréquence des différentes positions d'arrivée et pour cela simuler N marches aléatoires de 30 sauts, grâce à l'algorithme suivant :

```

Algorithme 3
Variables
    N, x, i, j entiers, L liste
Entrées
    Saisir N
Initialisations
    Vider la liste L
Traitement
    Pour j variant de 1 à N
        x prend la valeur 0
        Pour i variant de 1 à 30
            Si alea < 0,5
                alors x prend la valeur x + 1
                sinon x prend la valeur x - 1
            Fin du Si
        Fin du Pour
        L(j) prend la valeur x
    Fin du Pour
Sorties
    Afficher la liste L
    
```

Pour mieux se représenter les résultats, il peut être utile de représenter les données aléatoires ainsi obtenues, sous la forme du graphique suivant :



Exemple : nombre de lancers pour sortir tous les numéros d'un dé cubique²⁰

On souhaite estimer le nombre de lancers nécessaires pour sortir toutes les faces d'un dé cubique.

Comme pour l'exemple précédent, la réponse est aléatoire. On peut, dans un premier temps, écrire un algorithme permettant de simuler une expérience, puis faire tourner plusieurs fois cet algorithme et observer que la moyenne des résultats obtenus se stabilise²¹ pour un grand nombre d'expériences. Cette moyenne, elle aussi aléatoire, permet néanmoins de proposer une réponse au problème en termes de probabilités.

```
Algorithme
  Variables
    L liste
    S, n, x entiers
  Initialisations
    Vider la liste L
    S et n prennent la valeur 0
  Traitement
    Tant que22 n est inférieur à 6
      S prend la valeur S +1
      x prend la valeur d'un nombre aléatoire entier entre 1 et 6
      si x ne figure pas dans la liste L alors
        n prend la valeur n+1
        L(n) prend la valeur x
      Fin du Tant que
  Sorties
    Afficher S
```

On peut, pour estimer le nombre moyen de lancers permettant de sortir les six faces du dé, modifier l'algorithme précédent en s'inspirant de l'algorithme 2 ci-avant.

III. Échantillonnage

1. Fluctuation d'échantillonnage

1.1. Notion d'échantillon

Le terme « échantillon » prenant différents sens, il convient de préciser ce qu'il recouvre dans le programme de seconde.

Dans le sens commun des sondages, ce terme s'apparente à un sous-ensemble obtenu par prélèvement aléatoire dans une population. Ainsi parle-t-on usuellement de résultats estimés sur échantillon.

²⁰ On peut examiner en variante le nombre moyen de tablettes de chocolat qu'il faut acheter pour obtenir la collection complète des six images qui sont jointes. Cf. une expérimentation en classe à l'adresse <http://www.statistix.fr/spip.php?article15>

²¹ On en restera à une approche intuitive de la loi des grands nombres.

²² On peut observer que cette condition n'assure pas que l'algorithme se termine en un nombre fini d'opérations. Il suffirait alors de modifier légèrement le critère de sortie de boucle, en fixant par exemple un nombre maximum pour S au-delà duquel le programme s'arrête.

En statistique, un échantillon de taille n est la liste des n résultats obtenus par n répétitions indépendantes de la même expérience²³. Par exemple, un échantillon de taille 100 du lancer d'une pièce est la liste des résultats pile ou face obtenus successivement en répétant 100 fois le lancer de la pièce. De même pour un échantillon de taille 100 relatif au lancer d'un dé dont on observe l'apparition ou non de la face 6, ou bien encore pour un échantillon obtenu par tirages successifs avec remise d'une boule dans une urne contenant 2 boules blanches et une boule verte. Ces trois exemples relèvent en fait du même modèle, celui de Bernoulli qui affecte la probabilité p au nombre 1 et la probabilité $(1 - p)$ au nombre 0, seule situation abordée en classe de seconde.

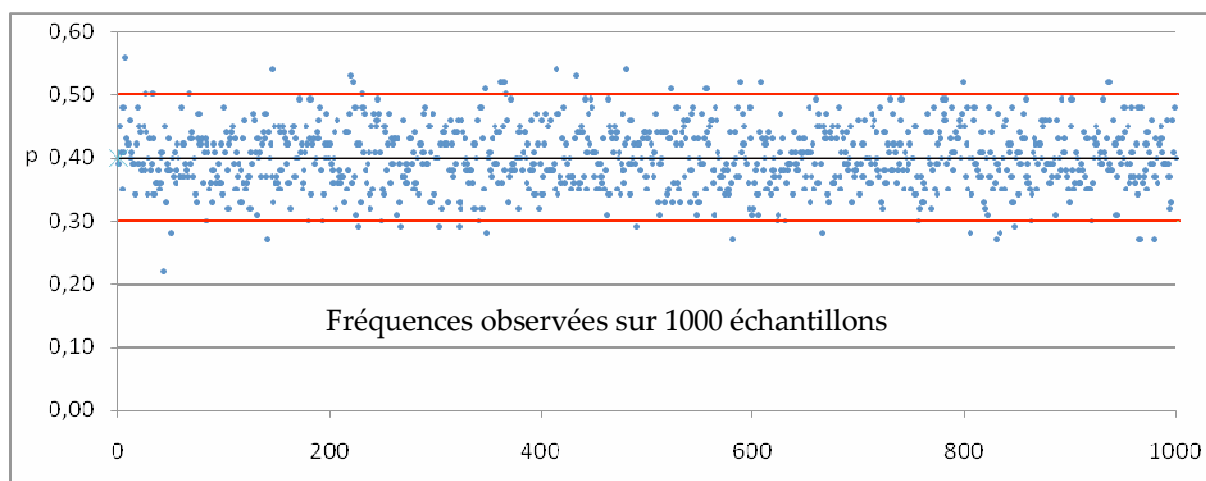
Cette notion d'échantillon fournit un cadre théorique pour démontrer les résultats énoncés ci-dessous sur la fluctuation d'échantillonnage.

En outre, ces résultats théoriques pourront s'appliquer aux sondages destinés à estimer une proportion p (par exemple le pourcentage de votes « oui » lors d'un référendum), en remarquant qu'un tirage sans remise d'un échantillon dans une population suffisamment nombreuse est assimilable à la répétition d'un tirage avec remise dans une urne.

1.2. Intervalle de fluctuation²⁴

On peut, par expérimentation et simulation, faire observer aux élèves que les échantillons de taille n obtenus à partir d'un modèle de Bernoulli ont, pour environ 95% d'entre eux, des fréquences d'apparition du nombre 1 qui fluctuent²⁵ dans un intervalle centré en p et d'amplitude $\frac{2}{\sqrt{n}}$.

On a simulé ci-dessous²⁶, 1000 échantillons de taille 100 d'un modèle de Bernoulli avec $p = 0,4$. Chaque échantillon est représenté par un point dont l'ordonnée est sa fréquence d'apparition du 1. On observe que la plupart des échantillons ont des fréquences d'apparition du 1 dans l'intervalle $[0,3 ; 0,5]$.



²³ C'est-à-dire relative au même modèle.

²⁴ Voir les simulations avec un tableur :

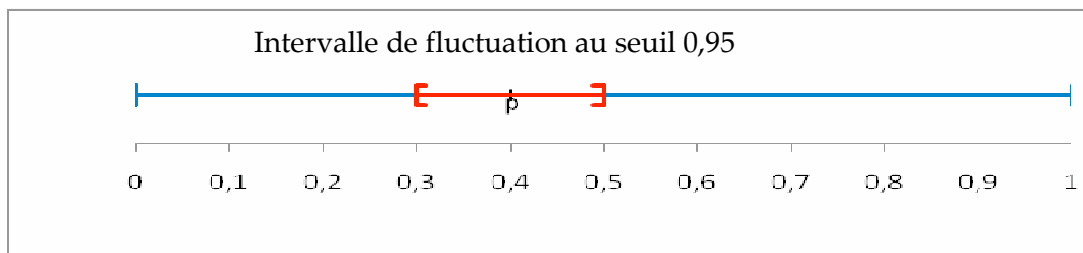
http://www.ac-grenoble.fr/maths/docresseconde/intervalles_fluctuation_confiance.xlsx

ou http://www.ac-grenoble.fr/maths/docresseconde/intervalles_fluctuation_confiance.ods

²⁵ Pour n assez grand, on observera que $\frac{1}{\sqrt{n}}$ est l'ordre de grandeur de cette fluctuation autour de p au seuil 95%.

²⁶ Ici la simulation a été effectuée au tableur à l'aide de la formule ENT(ALEA()+0,4)

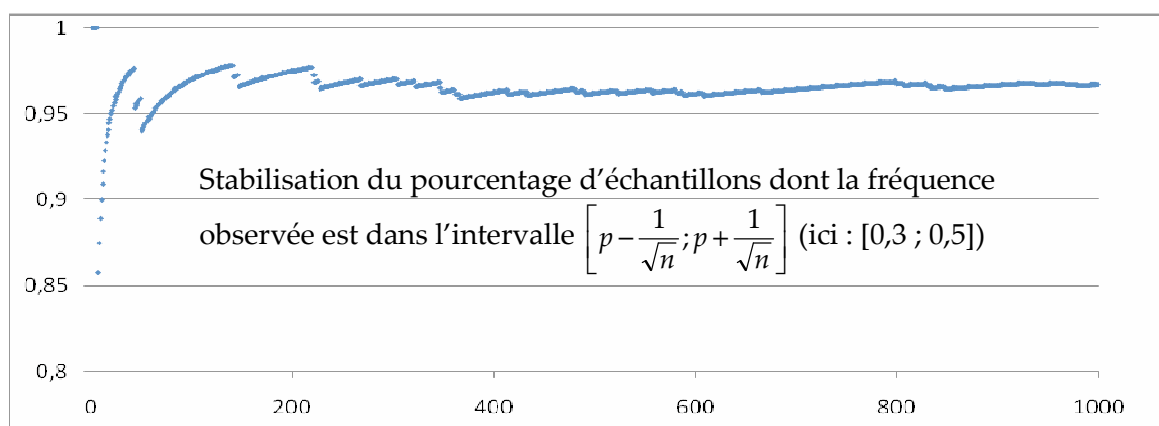
On peut alors définir²⁷ l'intervalle de fluctuation au seuil de 95%, relatif aux échantillons de taille n , comme l'intervalle centré autour de p , où se situe, avec une probabilité égale à 0,95, la fréquence observée dans un échantillon de taille n .



Dans la pratique, on utilise l'intervalle²⁸ $\left[p - \frac{1}{\sqrt{n}}; p + \frac{1}{\sqrt{n}} \right]$, pour des probabilités p comprises entre 0,2 et 0,8, et des échantillons de taille n supérieure ou égale à 25.

Pour p donné, on peut faire calculer les bornes de cet intervalle pour quelques valeurs de n , et remarquer qu'il faut multiplier la taille de l'échantillon par k^2 pour diviser par k l'amplitude de l'intervalle. On pourra calculer l'amplitude correspondant aux échantillons de taille 1000, taille souvent retenue dans les sondages.

Il est possible de visualiser²⁹ le pourcentage d'échantillons dont les fréquences d'apparition du 1 sont situées dans l'intervalle $\left[p - \frac{1}{\sqrt{n}}; p + \frac{1}{\sqrt{n}} \right]$. Le graphique³⁰ suivant est obtenu à partir des 1000 échantillons simulés et représentés ci-avant.



²⁷ Il faudrait en fait considérer le plus petit intervalle où se situent les fréquences observées avec une probabilité au moins égale à 0,95. Mais pour une première approche de cette notion, on s'est limité à l'énoncé ci-dessus.

²⁸ Il s'agit d'un résultat asymptotique, résultant de la convergence en loi de la variable aléatoire f_n correspondant à la fréquence d'un échantillon de taille n vers la loi normale de moyenne p et d'écart type $\sigma = \sqrt{p(1-p)}$. Ainsi,

pour n assez grand, f_n appartient avec une probabilité d'environ 0,95 à l'intervalle $\left[p - \frac{1,96\sigma}{\sqrt{n}}; p + \frac{1,96\sigma}{\sqrt{n}} \right]$ qui est

inclus dans $\left[p - \frac{1}{\sqrt{n}}; p + \frac{1}{\sqrt{n}} \right]$ (car $\sigma = \sqrt{p(1-p)} \leq \frac{1}{2}$). On pourra se référer au document d'accompagnement du

programme 2001 de la classe de seconde à l'adresse :

http://www.ac-grenoble.fr/maths/docresseconde/doc_proba_stat_seconde2001.pdf

²⁹ cf. feuille de calcul : http://www.ac-grenoble.fr/maths/docresseconde/intervalles_fluctuation_confiance.xlsx

ou http://www.ac-grenoble.fr/maths/docresseconde/intervalles_fluctuation_confiance.ods

³⁰ Lecture : parmi les 200 premiers échantillons de taille 100 obtenus, environ 97% ont une fréquence d'apparition du 1 dans l'intervalle [0,3 ; 0,5]

2. Applications de la fluctuation d'échantillonnage

2.1. Prise de décision à partir d'un échantillon

On pourra développer quelques exemples de prises de décision sur échantillon fondées sur la connaissance de l'intervalle de fluctuation.

Exemple : la parité, c'est quoi ?

Deux entreprises A et B recrutent dans un bassin d'emploi où il y a autant de femmes que d'hommes, avec la contrainte du respect de la parité. Dans l'entreprise A, il y a 100 employés dont 43 femmes ; dans l'entreprise B, il y a 2500 employés dont 1150 femmes (soit 46%). Or 46% est plus proche de 50% que 43% : les chiffres parlent d'eux-mêmes, pourrait-on dire, et l'entreprise B respecte mieux la parité que l'entreprise A. Si on admet que la parité, c'est exactement 50% de femmes, il est vrai que B en est plus proche que A. Mais une telle définition, à l'unité près, de la parité n'aurait ici pas de sens.

La parité, cela signifie que l'identité sexuelle n'intervient pas au niveau du recrutement, c'est-à-dire qu'au niveau du caractère homme ou femme, les résultats observés pourraient être obtenus par choix au hasard des individus dans la population. Dans ce cadre, l'entreprise A est assimilable à un échantillon de taille 100 du modèle de Bernoulli (avec $p = 0,5$) dont l'intervalle de fluctuation est $[0,4 ; 0,6]$ et l'entreprise B à un échantillon de taille 2500 dont l'intervalle de fluctuation est $[0,48 ; 0,52]$. La valeur 43% est donc dans l'intervalle de fluctuation, alors que 46% ne l'est pas. Autrement dit, pour l'entreprise B, la proportion de 46% s'observe dans moins de 5% des échantillons obtenus selon le modèle accordant une probabilité égale d'obtenir un homme et une femme. On peut alors rejeter l'hypothèse que cette entreprise respecte la parité.

En résumé, le raisonnement pour apprécier si une fréquence observée f sur un échantillon de taille n est compatible ou non avec un modèle de Bernoulli de probabilité p , est le suivant : on regarde si cette fréquence est dans l'intervalle de fluctuation à 0,95³¹ relatif aux échantillons de taille n du modèle, c'est-à-dire si l'écart entre f et p est probable, au sens où le hasard produirait un tel écart dans 95% des échantillons envisageables. Si f est en dehors de l'intervalle de fluctuation, on considère que l'observation n'est pas compatible avec le modèle, en ce sens que dans un tel modèle elle ne s'observerait que dans 5% des échantillons de taille n (avec le risque de prendre la mauvaise décision dans 5% des cas). On s'interrogera alors, par exemple dans le contrôle de qualité industrielle, sur le réglage d'une machine lorsque dans un lot de pièces produites la fréquence de défauts observés est peu probable au regard du modèle indiquant une probabilité p de défauts.

Ce type de raisonnement, qui intègre la connaissance de l'intervalle de fluctuation, est à la base de ce que l'on appelle parfois une « preuve statistique ».

On trouvera de nombreuses activités pour la classe dans le document ressource relatif au nouveau programme de mathématiques de la voie professionnelle, disponible à l'adresse suivante : http://www.ac-grenoble.fr/maths/docresseconde/Proba_stat_LP.doc. On s'y interroge par exemple, à l'aune de la fluctuation d'échantillonnage, sur la fréquence des cas de leucémie chez les garçons de moins de 15 ans dans une petite ville des États-Unis (9 cas sur 5969, soit une fréquence de 0,0015 alors que la proportion constatée sur l'ensemble du territoire est de 0,00052). De même on questionne les 16 naissances masculines sur les 20

³¹ Le seuil 0,95 est conventionnel. D'autres seuils sont retenus selon les questions ou domaines étudiés.

naissances d'un petit village de Chine ou encore les 46 naissances de garçons parmi les 132 naissances dans une petite ville du Canada. Dans ces différents exemples, le raisonnement statistique permet de mettre en évidence des observations « rares » au regard de la fluctuation d'échantillonnage, qui ont déclenché des investigations complémentaires, extérieures au champ des mathématiques, pour rechercher et analyser les causes des phénomènes observés.

Par ailleurs, une activité étudiée, au regard des probabilités de défections constatées à l'embarquement, le risque de surréservation que prend une compagnie aérienne lorsqu'elle vend plus de billets que de places.

On y aborde aussi un exemple de contestation d'un jugement au Texas au motif que la désignation des jurés était discriminante à l'égard des américains d'origine mexicaine. Cette possibilité de quantifier, en termes de probabilité, si la fréquence observée (ici la proportion de jurés d'origine mexicaine) est rare ou fréquente dans un modèle considéré, fournit au juge des éléments qui, agrégés à d'autres arguments, fondent son jugement.

Exemple : élection annulée

Dans un entretien retranscrit dans le livre : "le hasard aujourd'hui ", Jean-Louis Boursin relate ce qui s'est passé lors d'une élection dans la région parisienne :

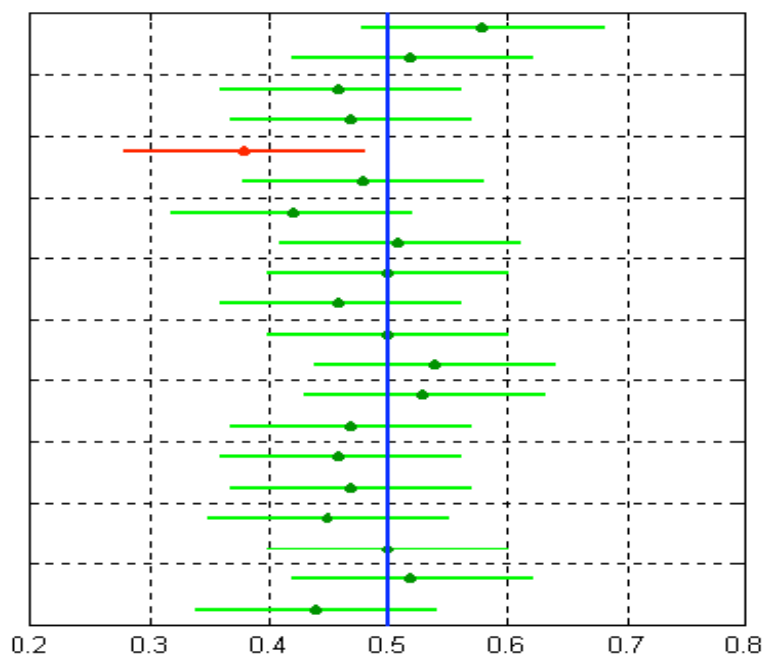
[... Un candidat aux élections législatives soupçonnait de fraude un certain nombre de bureaux de vote. Il pensait que dans ces bureaux là, il y avait un risque parce qu'il n'avait pas confiance en ceux qui tenaient les bureaux. Il a fait faire des sondages extrêmement précis, il a fait une étude sur les élections précédentes et, muni de ces chiffres et des résultats, centaine par centaine, il est allé au tribunal administratif et a affirmé que le hasard ne pouvait pas produire cela : " voilà une centaine de bulletins qui donnent 98 bulletins à mon adversaire et 2 à moi, cela n'est pas possible. Le calcul des probabilités montre que cela a une chance sur dix ou quinze millions de se produire ". Le tribunal lui a donné raison sur un simple raisonnement probabiliste...].

2.2. Estimation d'une proportion

On se place dans la situation de référence suivante, d'une urne contenant plusieurs centaines de petites billes de couleur blanche ou verte dans une proportion p inconnue de billes vertes. On cherche à estimer p à partir d'un échantillon de taille n (on imagine aisément des applications de cette situation, par exemple à l'estimation du résultat d'un référendum).

On considère alors un échantillon de taille n (par répétition de n tirages aléatoires avec remise dans l'urne) et on calcule la fréquence des billes vertes dans cet échantillon. On dispose ainsi d'un échantillon parmi tous ceux qu'on aurait pu obtenir et, d'après le paragraphe 1.2, on sait qu'environ 95% des fréquences observées sont dans l'intervalle $\left[p - \frac{1}{\sqrt{n}}; p + \frac{1}{\sqrt{n}} \right]$. Or l'appartenance de f à $\left[p - \frac{1}{\sqrt{n}}; p + \frac{1}{\sqrt{n}} \right]$ équivaut à celle de p à l'intervalle $\left[f - \frac{1}{\sqrt{n}}; f + \frac{1}{\sqrt{n}} \right]$ (appelé aussi fourchette de sondage), ce qui permet de dire que, parmi tous les échantillons de taille n possibles, 95% des intervalles associés $\left[f - \frac{1}{\sqrt{n}}; f + \frac{1}{\sqrt{n}} \right]$ contiennent le nombre p .

On a représenté ci-dessous³², les fourchettes associées à 20 échantillons de taille 100 dans le cas $p = 0,5$. Un échantillon (en rouge) donne une fourchette ne contenant pas p , les autres (en vert) contiennent p .



Pour exprimer l'idée qu'avant tirage³³ de l'échantillon on avait 95% de chances d'obtenir une fourchette $\left[f - \frac{1}{\sqrt{n}} ; f + \frac{1}{\sqrt{n}} \right]$ qui contienne p , on dira que la fourchette obtenue, une fois l'échantillon tiré, est un intervalle de confiance au niveau 0,95 de p . On notera que l'intervalle de confiance associé à un échantillon au niveau 0,95 ne dépend que de la taille n de l'échantillon et non pas de la taille de la population³⁴.

Exemple : estimation par sondage du pourcentage d'artisans parmi les entreprises d'un département.

À partir de données réelles³⁵ indiquant, pour les entreprises d'un département, si elles ont ou non le statut d'artisan, on peut estimer grâce à un échantillon de taille 1000 la proportion d'artisans. Un tirage aléatoire de 1000 entreprises³⁶ a donné 36,5% d'artisans. Ainsi,

³² D'après le document d'accompagnement du programme de seconde publié en 2001 consultable à l'adresse :

http://www.ac-grenoble.fr/math/docresseconde/doc_proba_stat_seconde2001.pdf

Voir aussi des simulations :

http://www.ac-grenoble.fr/math/docresseconde/intervalles_fluctuation_confiance.xlsx

ou http://www.ac-grenoble.fr/math/docresseconde/intervalles_fluctuation_confiance.ods

³³ Il est clair qu'une fois l'échantillon tiré, l'intervalle de confiance associé est entièrement fixé et qu'il n'y a alors que deux possibilités : p appartient ou n'appartient pas à l'intervalle. C'est pourquoi il est incorrect de dire que p a une probabilité 0,95 d'appartenir à cet intervalle.

³⁴ Dans son livre " Les structures du hasard", Jean Louis Boursin explique ceci à l'aide d'une métaphore : étudier un échantillon c'est comme prendre une louche dans une grande marmite pour goûter la soupe, peu importe la taille de la marmite, on aura déjà beaucoup de renseignements sur la soupe à partir de cet échantillon.

³⁵ D'après l'INSEE, fichier à l'adresse :

http://www.ac-grenoble.fr/math/docresseconde/Exemple_de_progression/activites_artisans.doc

³⁶ On trouvera un exemple d'algorithme qui permet d'extraire un échantillon d'une liste, utilisable aussi sur calculatrice : http://www.ac-grenoble.fr/math/docresseconde/Annexe_algorithme_echantillons.doc

l'intervalle [33,3% ; 39,7%] est un intervalle de confiance au niveau 0,95 du pourcentage d'artisans dans ce département (la vraie valeur de 37,34% pouvant en l'occurrence être calculée). Cette situation réelle permet aux élèves de pratiquer une estimation par sondage statistique.

IV. Repères pour l'évaluation

La diversité des objectifs visés par l'enseignement des statistiques et des probabilités, tels qu'ils sont précisés dans le programme, invite à proposer des formes d'évaluation variées, prenant davantage en compte l'usage des TIC ou l'expression orale.

Bien sûr, de nombreuses capacités attendues peuvent aisément s'évaluer sous la forme habituelle d'un devoir en temps limité, comme par exemple tout ce qui a trait aux calculs de probabilités, aux représentations graphiques ou aux résumés statistiques.

En revanche, s'agissant de la fluctuation d'échantillonnage, l'objectif est de faire réfléchir les élèves à la conception et à la mise en œuvre d'une simulation et de les sensibiliser aux notions d'intervalle de fluctuation, d'intervalles de confiance et à l'utilisation qui peut en être faite. Aussi, semble-t-il prématuré d'exiger dans des contrôles écrits une autonomie totale des élèves pour conduire les raisonnements qui sont attachés à ces notions : on prendrait en effet le risque de restitutions par cœur pour compenser une assimilation naissante et encore fragile.

C'est pourquoi, l'évaluation des capacités attendues suivantes :

- concevoir, exploiter et mettre en œuvre des simulations de situations concrètes à partir d'un tableur ou d'une calculatrice,
- exploiter et faire une analyse critique d'un résultat d'échantillonnage,
- utiliser un logiciel ou une calculatrice pour étudier une série statistique,

devrait majoritairement être réalisée sous forme de comptes-rendus de travaux pratiques ou de devoirs à la maison. Ces modalités d'évaluations donnent accès à des situations statistiques ou probabilistes appelant l'usage d'un logiciel ou d'une calculatrice, et permettent, en raison de l'interaction possible entre le professeur et les élèves, d'ouvrir le champ de l'évaluation à des situations plus riches et plus ouvertes, qui mobilisent davantage les capacités de recherche, d'expérimentation et d'initiative.

Par ailleurs, concernant tout particulièrement ce chapitre, la place de l'oral gagnerait à être développée, tant cette forme de communication facilite, par le questionnement interactif qu'elle permet, l'explicitation de certains raisonnements statistiques délicats à consigner à l'écrit. Dans ce cadre, on peut envisager de proposer des situations dont l'étude est réalisée en classe et dont le compte-rendu, rédigé à la maison, est suivi d'un exposé en classe ou bien d'échanges avec le professeur permettant d'approfondir certaines argumentations ou démarches imparfaitement restituées à l'écrit afin de les améliorer.

L'expérience acquise lors de l'expérimentation de l'épreuve pratique de mathématiques³⁷ et les critères d'évaluation qui y ont été explicités constitueront de précieux points d'appui, transférables à ces nouvelles modalités d'évaluation.

³⁷ Cf. le site du groupe des mathématiques de l'inspection générale : <http://igmaths.net/>

Exemple : discrimination et statistique³⁸

À partir de l'activité « contester un jugement » présentée en page 15 dans le document ressource relatif aux nouveaux programmes de lycée professionnel³⁹, on peut proposer un travail pratique en classe amenant les élèves à reformuler la question sur la discrimination éventuelle des jurés d'origine mexicaine (339 sur 870 jurés) dans un contexte statistique prenant en compte la fluctuation d'échantillonnage. Il s'agit alors de concevoir l'ensemble des jurys désignés à ce jour comme un des échantillons de taille 870 de la loi de Bernoulli de paramètre p , où p désigne la proportion au Texas d'américains d'origine mexicaine ($p \approx 0,791$). Cette reformulation étant effectuée, les élèves sont amenés ensuite à simuler des échantillons de taille 870 pour voir si la fréquence observée est dans la zone de fluctuation au seuil 95% ou bien ils peuvent utiliser l'intervalle $\left[p - \frac{1}{\sqrt{n}} ; p + \frac{1}{\sqrt{n}} \right]$. Il leur faut enfin conclure au regard du problème étudié.

Il est clair que ces différentes étapes peuvent donner lieu à des échanges oraux avec le professeur et à des productions écrites partielles qui peuvent être évaluées. L'évaluation peut porter, par exemple, sur la capacité des élèves à tirer profit du questionnement oral du professeur ou de ses indications pour reformuler ce problème de discrimination dans le cadre statistique de l'échantillonnage, ou aussi sur la capacité à mettre en œuvre une simulation sur tableur ou sur calculatrice, ou bien encore sur la capacité à interpréter les résultats de la simulation pour conclure. Tous ces éléments peuvent ensuite être intégrés dans l'évaluation globale de la situation étudiée, soit sous la forme d'une notation intermédiaire, soit par des appréciations qui seront prises en compte lors de la notation finale à la suite de la production écrite. D'une façon générale la notation ne peut être dissociée de l'évaluation des compétences acquises.

Vouloir évaluer sous une forme traditionnelle les capacités mobilisées dans de telles activités amènerait le professeur à proposer des questions intermédiaires qui limiteraient la prise d'initiative et l'intérêt que les élèves peuvent porter à l'étude de situations en lien avec la vie courante, avec des faits de société ou des questions scientifiques.

Signalons, pour terminer, l'idée originale qui a amené les professeurs de mathématiques du groupe « statistique et citoyenneté » de l'IREM de Paris Nord⁴⁰ à travailler avec des professeurs de français sur le même thème et à comparer la preuve statistique ainsi apportée aux critères d'un texte argumentatif en français⁴¹.

Une bibliographie et des compléments pour le professeur sont disponibles à la fin du document ressource pour les lycées professionnels.

³⁸ Activité du groupe « statistique et citoyenneté » de l'IREM de Paris Nord présentée sur le site statistix (<http://www.statistix.fr/spip.php?article27>) et disponible aussi à l'adresse <http://www.ac-grenoble.fr/maths/docresseconde/discrimination.doc>.

³⁹ Voir document ressource pour les lycées professionnels : http://www.ac-grenoble.fr/maths/docresseconde/Proba_stat_LP.doc

⁴⁰ Les travaux de ce groupe sont disponibles à l'adresse : <http://www-irem.univ-paris13.fr/spip/spip.php?rubrique15>

⁴¹ Cf. <http://www.ac-grenoble.fr/maths/docresseconde/discrimination.doc>