Test d'équirépartition : qui a dit khi-deux ?

Louis-Marie Bonneval

Résumé : Le nouveau programme de Terminale S et ES introduit le problème de l'adéquation de données statistiques à un modèle probabiliste. L'article en évoque le contexte théorique et les enjeux didactiques.

Les programmes de Terminale S et de Terminale ES en vigueur à la rentrée 2002 comportent un paragraphe, identique dans les deux filières, intitulé *Simulation* et formulé comme suit :
Étude d'un exemple traitant de l'adéquation de données expérimentales à une loi équirépartie.
Le commentaire est le suivant :
L'élève devra être capable de poser le problème de l'adéquation à une loi équirépartie et de se reporter à des résultats de simulation qu'on lui fournit. Le vocabulaire des tests (test d'hypothèse, hypothèse nulle, risque de première espèce) est hors programme.
C'est bien entendu au test du χ^2 qu'il est fait allusion.
L'introduction de cette notion avant le baccalauréat est une nouveauté, dont on ne peut contester l'intérêt : tester la validité d'un modèle est une démarche essentielle dans toute activité scientifique, et beaucoup de nos élèves (y compris ceux qui se dirigeront vers les sciences humaines) auront à la pratiquer.
Mais elle est délicate à mettre en œuvre, car elle suppose de la part des enseignants une vision claire autant de la théorie sous-jacente que de ses enjeux didactiques.
Or beaucoup de collègues sont mal à l'aise avec les statistiques inférentielles, n'ayant pas eu dans leur cursus la formation adéquate. L'attitude volontariste du GEPS, que je soupçonne de penser « au pied du mur, ils devront bien s'y mettre », me paraît dangereuse, car on enseigne mal (ou pas du tout) ce qu'on maîtrise mal. Une formation des enseignants est indispensable.
Cet article voudrait y apporter une contribution.

1. La théorie

Pour étudier une épreuve aléatoire, on a besoin d'un modèle, qui précise d'une part les issues (on les suppose ici en nombre fini), d'autre part la distribution de probabilité entre ces issues. Or, mis à part les jeux de hasard où des hypothèses de symétrie fournissent *a priori* les probabilités des issues, c'est l'observation statistique antérieure qui permet d'estimer ces probabilités. Cette observation consiste à répéter l'épreuve le plus grand nombre de fois possible, et à relever dans cet *échantillon* les fréquences d'apparition des issues. Comme les divers échantillons fournissent des résultats différents, il est nécessaire d'étudier la *fluctuation*

d'échantillonnage, autrement dit la variabilité entre les divers échantillons possibles de même taille. Le résultat essentiel, conforme à l'intuition, tient en deux phrases :
• (1) la distribution de fréquence fluctue autour de la distribution de probabilité ;
• (2) la fluctuation est faible quand la taille de l'échantillon est grande.

Regardons de plus près ce qu'il en est.

Considérons une épreuve e ayant k issues ω_1, ω_2, …, ω_k, de probabilités p_1, p_2, …, p_k.
Appelons e^n l'épreuve qui consiste à répéter n fois l'épreuve e, ce qu'on appelle aussi[1] *tirer un échantillon de taille n* de l'épreuve e.
Notons f_1, f_2, …, f_k les fréquences d'apparition de ω_1, ω_2, …, ω_k. Ce sont des variables aléatoires relatives à e^n.
Chaque f_i a une loi de probabilité de type binomial[2], d'espérance p_i et d'écart-type $\sqrt{\dfrac{p_i(1-p_i)}{n}}$.

$(f_1, f_2, \ldots, f_k)$ est un *vecteur aléatoire* $\vec{f}$ relatif à l'épreuve e^n, qui prend ses valeurs dans $\mathbf{R}^k$.

Comme $\sum_1^k f_i = 1$, ce vecteur appartient à l'hyperplan de $\mathbf{R}^k$ d'équation $\sum_1^k x_i = 1$: on dit qu'il a k–1 *degrés de liberté*.

Comme chaque f_i a pour espérance p_i, l'espérance de $\vec{f}$ est le vecteur $\vec{p} = (p_1, p_2, \ldots, p_k)$, ce qui donne un sens précis à l'affirmation (1).

Pour mesurer la dispersion de $\vec{f}$ autour de son espérance $\vec{p}$, il serait naturel de former $\left\|\vec{f} - \vec{p}\right\| = \sqrt{\sum_1^k (f_i - p_i)^2}$.

Or il s'avère plus intéressant de considérer la quantité $n\sum_1^k \dfrac{(f_i - p_i)^2}{p_i}$. Elle peut jouer le même rôle de distance puisqu'elle est positive et d'autant plus petite que les f_i sont proches des p_i. Mais le fait de pondérer les termes par les $\dfrac{n}{p_i}$ a pour effet de la « normer », en ce sens que pour n assez grand, sa loi de probabilité ne dépend pratiquement plus des p_i, ni de n, mais seulement de k.

(1) Le mot « échantillon » évoque le tirage dans une urne de Bernoulli. On sait qu'une telle urne modélise toute épreuve d'univers fini (du moins quand les p_i sont rationnels). Dans le langage courant, on parle d'échantillons lors de tirages sans remise ; mais il est usuel en statistiques inférentielles d'étendre ce terme à des tirages avec remise, et donc de l'utiliser pour parler d'épreuves répétées.

(2) $f_i = \dfrac{n_i}{n}$ où n_i, nombre d'apparitions de ω_i, suit la loi binomiale $B(n,p_i)$.

On vérifie déjà que (quel que soit n) son espérance est égale à $k-1$:

$$\mathrm{E}\left(n\sum_1^k \frac{(f_i-p_i)^2}{p_i}\right)=\sum_1^k \frac{n}{p_i}\mathrm{E}\left((f_i-p_i)^2\right)$$

$$=\sum_1^k \frac{n}{p_i}\mathrm{V}(f_i)=\sum_1^k \frac{np_i(1-p_i)}{np_i}=\sum_1^k (1-p_i)=k-\sum_1^k p_i=k-1.$$

Surtout on démontre le théorème suivant :

Si n tend vers l'infini, la loi de probabilité de $n\sum_1^k \frac{(f_i-p_i)^2}{p_i}$ converge vers une loi connue, dite loi du χ^2 à $k-1$ degrés de liberté.

Je ne traiterai pas ici de la loi du χ^2, qui est définie, étudiée et tabulée dans les livres de statistiques(3). Il suffit de savoir que son quantile d'ordre 0,95, valeur Q telle que $P(\chi^2 < Q) = 0{,}95$, est fourni par les tables, ou les tableurs(4) :

k	2	3	4	5	6	7	8	9	10
degrés de liberté	1	2	3	4	5	6	7	8	9
Q	3,84	5,99	7,82	9,49	11,1	12,6	14,1	15,5	16,9

Pour les valeurs de k supérieures à 10, on peut utiliser la formule d'approximation(5) de Wilson-Hilferty :

$$\mathrm{Q}\approx(k-1)\left(1-\frac{2}{9(k-1)}+1{,}645\sqrt{\frac{2}{9(k-1)}}\right)^3.$$

Si n est assez grand pour qu'on puisse assimiler la loi de $n\sum_1^k \frac{(f_i-p_i)^2}{p_i}$ à celle du χ^2, on en déduit que :

pour 95 % des échantillons de taille n, la quantité $n\sum_1^k \frac{(f_i-p_i)^2}{p_i}$ est inférieure à $\frac{\mathrm{Q}}{n}$.

Cela donne un sens précis à l'affirmation (2) : en effet $\frac{\mathrm{Q}}{n}$ diminue quand la taille n de l'échantillon augmente.

Ainsi un échantillon de taille 10 000 apporte une information plus précise qu'un échantillon de taille 100.

Une application importante de ce résultat est le **test du χ^2** :

(3) Voir par exemple [1], p. 254 ou [6], p. 17.

(4) Dans EXCEL utiliser la fonction KHIDEUX.INVERSE.

(5) Dans cette formule, le nombre 1,645 est le quantile 0,95 de la loi normale. Il existe d'autres formules d'approximation (voir par exemple [6], p. 173).

Pour une épreuve donnée, d'issues ω_1, ω_2, …, ω_k, on émet l'hypothèse(6) que la distribution de probabilité est un certain vecteur $(p_1, p_2, \ldots, p_k)$. Pour tester la validité de cette hypothèse, on observe un échantillon de taille n assez grande, et on calcule

$$\sum_1^k \frac{(f_i - p_i)^2}{p_i}.$$

Si la valeur obtenue dépasse le seuil critique $\frac{Q}{n}$, on rejette l'hypothèse. L'argument pour cela est le suivant : il s'est produit un événement qui, sous cette hypothèse, aurait eu moins de 5% de chances de se produire ; il paraît donc peu vraisemblable que l'hypothèse soit vraie. Bien entendu, on n'a pas une certitude : on dit que la conclusion a un *niveau de confiance* de 95 %.
On notera au passage que le test permet de rejeter une hypothèse, mais ne permet pas de la valider : tout au plus pourra-t-on dire, si $\sum_1^k \frac{(f_i - p_i)^2}{p_i}$ est inférieur à $\frac{Q}{n}$, que l'échantillon observé est compatible avec l'hypothèse.

Remarques

1) Usuellement on calcule plutôt $n\sum_1^k \frac{(f_i - p_i)^2}{p_i}$, qui s'écrit aussi $\sum_1^k \frac{(n_i - e_i)^2}{e_i}$, où $n_i = nf_i$ désigne le nombre observé d'apparitions de chaque issue, et $e_i = np_i$ leur nombre théorique d'apparitions. Puis on compare le résultat à Q.
2) Dans le cas où $k = 2$ (épreuve à deux issues), on retrouve l'intervalle de dispersion de l'alternative répétée (schéma de Bernoulli). En effet $\sum_1^2 \frac{(f_i - p_i)^2}{p_i}$ $= \frac{(f-p)^2}{p} + \frac{(1-f-1+p)^2}{1-p} = \frac{(f-p)^2}{p(1-p)}$. La condition $n\sum_1^2 \frac{(f_i - p_i)^2}{p_i} < 3{,}84$ équivaut donc à $|f - p| < 1{,}96\sqrt{\frac{p(1-p)}{n}}$.
3) On peut bien entendu choisir un autre niveau de confiance, par exemple 0,90 ou 0,99. Il faudra alors rechercher le quantile d'ordre 0,90 ou 0,99 de la loi du χ^2 à $k - 1$ degrés de liberté.
4) La loi du χ^2 n'est qu'une approximation de la loi de $n\sum_1^k \frac{(f_i - p_i)^2}{p_i}$, valable pour n assez grand. J'y reviendrai plus loin.

(6) Le mot « hypothèse » est à prendre ici au sens de « conjecture » que lui donnent les sciences expérimentales, et non pas au sens mathématique traditionnel de « prémisse ».

Cas particulier : équiprobabilité

Supposons $p_1 = p_2 = \ldots = p_k = \frac{1}{k}$. Alors $\sum_1^k \frac{(f_i - p_i)^2}{p_i} = k\sum_1^k \left(f_i - \frac{1}{k}\right)^2$.

Or la lecture du tableau (et de la formule) ci-dessus montre que pour tout k, $\frac{Q}{k}$ est compris entre 1 et 2.

Il en résulte que pour 95 % des échantillons de taille n, $\sum_1^k \left(f_i - \frac{1}{k}\right)^2$ est inférieur à $\frac{2}{n}$.

On peut donc simplifier le test dans ce cas particulier :
Pour tester (au niveau de confiance 95 %) l'hypothèse que k issues sont équiprobables, on observe un échantillon de taille n assez grande, et on calcule $\sum_1^k \left(f_i - \frac{1}{k}\right)^2$. Si la valeur obtenue dépasse le seuil critique $\frac{2}{n}$, on rejette l'hypothèse.

Plus commodément, on calcule $n\sum_1^k \left(f_i - \frac{1}{k}\right)^2$ qu'on compare à 2, ou encore $\sum_1^k \left(n_i - \frac{n}{k}\right)^2$ qu'on compare à $2n$.

Exemple
On veut tester si un dé est régulier ; autrement dit, si $(p_1,p_2,p_3,p_4,p_5,p_6) = \left(\frac{1}{6},\frac{1}{6},\frac{1}{6},\frac{1}{6},\frac{1}{6},\frac{1}{6}\right)$. On décide de le lancer 600 fois.
On observe les effectifs n_1, n_2, … n_6 d'apparition des six valeurs. On calcule $\sum_1^6 (n_i - 100)^2$. Si ce nombre dépasse 1 200, on rejettera l'hypothèse que le dé est régulier.

2. Les enjeux didactiques

L'enjeu de formation est capital, puisqu'il concerne la démarche de modélisation.

Le programme se limite aux situations d'équiprobabilité, ce qui paraît judicieux.

D'une part c'est plus simple : la quantité $\sum_1^k \left(f_i - \frac{1}{k}\right)^2$ est plus naturelle que la quantité $\sum_1^k \frac{(f_i - p_i)^2}{p_i}$; et la majoration de $\frac{Q}{k}$ par 2 évite le recours aux tables.

D'autre part c'est le cas où l'approximation de la loi de $n\sum_1^k \frac{(f_i - p_i)^2}{p_i}$ par celle du χ^2 est la meilleure[7].

Surtout, il est essentiel de faire comprendre que l'équiprobabilité est une hypothèse forte, qui ne va pas de soi. On peut s'étonner à ce propos que dans les manuels beaucoup d'énoncés supposent une loi uniforme sans le dire : on parle d'un dé sans préciser qu'il est régulier, d'une pièce sans dire qu'elle est équilibrée[8] … : c'est comme si en géométrie on proposait un exercice sur un triangle équilatéral en omettant de dire qu'il est équilatéral ! La remarque vaut bien sûr pour les jeux de cartes (bien battus), les tirages (au hasard), etc.

Dans l'exemple ci-dessus, modéliser un dé réel consiste à le représenter par un dé régulier[9]. Il faut souligner que le dé régulier n'existe pas dans la Nature, pas plus que le triangle équilatéral. Il s'agit d'idéalités mathématiques sur lesquelles on peut raisonner. La question « ce dé est-il régulier ? » signifie en fait : « le modèle du dé régulier est-il adapté à ce dé ? ». Dès lors on conçoit que, selon la précision de la mesure, c'est-à-dire ici la taille de l'échantillon et le degré de confiance choisi, la réponse puisse être oui ou non. De même qu'à la question « Le triangle Paris-Bordeaux-Grenoble est-il équilatéral ? » on peut répondre oui ou non selon qu'on mesure les distances à 100 km près ou à 10 km près.

On peut même pousser plus loin l'analogie avec la géométrie : la question ci-dessus n'a plus de sens si on mesure les distances au km près, car le modèle ponctuel pour chaque ville n'est alors plus valide. De la même façon, on peut considérer que pour un dé réel, il n'y a pas de sens de définir les probabilités d'apparition des six faces à 10^{-4} près, car le nombre de jets qui serait nécessaire pour estimer ces probabilités[10]

(7) Voir plus loin.

(8) On peut remarquer que l'hypothèse d'équiprobabilité est plus naturelle pour un dé, conçu en principe pour être régulier, que pour une pièce de monnaie, qui n'est pas fabriquée pour cela, et dont il existe de multiples variétés !

(9) Les issues ω_1, ω_2, …, ω_k ayant été définies, un modèle est caractérisé par une distribution de probabilité $(p_1, p_2, \ldots, p_k)$. Le programme se limite à l'équiprobabilité, mais bien entendu on pourrait envisager pour un dé tout autre modèle, par exemple $\left(\frac{1}{8}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8}, \frac{3}{8}\right)$. Notons à ce propos que le modèle intègre non seulement les caractéristiques physiques du dé mais aussi la façon dont il est lancé.

(10) La théorie conduit à $n > 10^8$: peu de dés résisteraient à plus de 100 millions de jets (ce qui d'ailleurs, à raison d'un jet par seconde, durerait plus de trois ans) !

risquerait d'altérer ses caractéristiques physiques, rendant impossible la répétition de l'épreuve dans les mêmes conditions.

Dernier argument, mais non le moindre, en faveur de l'équiprobabilité : toute épreuve d'univers fini peut se ramener à une situation d'équiprobabilité(11). Par exemple, jeter un dé truqué où le 6 a trois fois plus de chances d'apparaître que chacune des 5 autres valeurs équivaut à tirer une boule dans une urne de Bernoulli contenant 8 boules marquées 1, 2, 3, 4, 5, 6, 6, 6. C'est d'ailleurs cette propriété qui permet les simulations faites en Seconde, qui doivent amener à distinguer ce qui est équiprobable et ce qui ne l'est pas(12). De ce point de vue, il me paraîtrait très utile de faire tester les générateurs de « nombres au hasard » fournis par les tableurs ou les calculatrices, de façon à faire comprendre qu'ils simulent bien une loi uniforme. Faute de quoi on risque de gros malentendus lors des simulations (par exemple simuler la somme de deux dés réguliers par des nombres au hasard entre 2 et 12).

Cela dit, la théorie ci-dessus n'est pas accessible à un élève de Terminale.
C'est pourquoi le commentaire du programme(13) propose la démarche suivante :

- Dans un premier temps, faire **simuler** au tableur une épreuve à k issues équiprobables, et calculer $n\sum_{1}^{k}\left(f_i-\frac{1}{k}\right)^2$ pour un grand nombre d'échantillons de taille n donnée. Écarter les 5% d'échantillons qui fournissent les valeurs les plus élevées, et conserver comme seuil critique la plus grande valeur des échantillons restants.
- Dans un deuxième temps, appliquer le test pour une situation d'équiprobabilité supposée.

Ainsi, pour tester la régularité d'un dé, on proposera comme ci-dessus de le lancer 600 fois, puis de calculer $600\sum_{1}^{6}\left(f_i-\frac{1}{6}\right)^2$, pour rejeter l'hypothèse de régularité si ce nombre est trop grand.
Mais pour savoir quel est le seuil critique qui permettra de trancher, on simulera au tableur 600 jets d'un dé régulier. On itérera par exemple 1 000 fois un tel échantillon de taille 600. On écartera les 5 % d'échantillons qui fournissent pour

(11) Cela suppose les p_i rationnels, ce qui ne restreint pas vraiment le champ d'application de l'urne de Bernoulli à des situations concrètes : tout irrationnel pouvant être approché d'aussi près qu'on veut par un rationnel, et tout modèle étant une approximation de la réalité, on peut toujours envisager un modèle où les p_i sont rationnels. On trouvera dans [5] une réflexion approfondie sur l'utilisation des urnes pour l'enseignement des probabilités au lycée.
(12) Par exemple quand on jette deux dés réguliers indépendants, c'est parce que les couples sont équiprobables que les sommes ne le sont pas.
(13) disponible (en version provisoire du 26/02/02) sur http://mathematiques.scola.ac-paris.fr/telechar/bo/2002/probastat.pdf

$600\sum_{1}^{6}\left(f_i-\frac{1}{6}\right)^2$ les valeurs les plus élevées, et on prendra comme seuil critique la plus grande valeur des échantillons restants.

Cette démarche expérimentale, intéressante en soi, n'est pas si facile à mettre en œuvre :

1) Elle suppose une bonne maîtrise de la simulation sur tableur ou calculatrice. Cette compétence est en principe acquise en Seconde, ce qui rend d'ailleurs cohérente la succession des programmes Seconde-Première-Terminale en statistique et probabilités. Il y a cependant quelques difficultés, dues entre autres à la mauvaise compréhension de ce qu'est un générateur de « nombres au hasard ». Est-ce pour cela que le programme parle de « résultats de simulation qu'on lui fournit » ? Si ce n'est pas l'élève qui fait lui-même la simulation, l'intérêt s'en trouve réduit.

2) Elle privilégie la quantité $n\sum_{1}^{k}\left(f_i-\frac{1}{k}\right)^2$, ce qui s'explique par le désir de faire apparaître un invariant, mais rend plus difficile la compréhension. Bien sûr il est équivalent de dire que $n\sum_{1}^{k}\left(f_i-\frac{1}{k}\right)^2$ est inférieur à 2, ou que $\sum_{1}^{k}\left(f_i-\frac{1}{k}\right)^2$ est inférieur à $\frac{2}{n}$. Mais c'est la deuxième écriture qui fait sens, puisqu'elle indique que la précision de l'approximation augmente avec n. Prendre conscience de cela suppose d'ailleurs de varier la taille des échantillons, ce qui rend très lourd ce travail expérimental.

Surtout, cette approche soulève des questions difficiles :

1) Elle suppose qu'on observe un nombre suffisant d'échantillons (combien ?) pour que la proportion de 95 % ait un sens statistique. Ce double niveau d'échantillonnage (de l'épreuve e pour constituer un échantillon, puis on répète l'épreuve e^n pour comparer les échantillons) constitue d'ailleurs un obstacle important à la compréhension.
2) Elle suppose que le générateur de nombres aléatoires du tableur est parfait. Or précisément d'après la théorie ci-dessus, cela ne va pas de soi. Tout au plus peut-on le tester, notamment à l'aide du χ^2, à un niveau de confiance donné. Il y a là comme un cercle vicieux.
3) Elle risque de faire croire que le seuil critique dépend de l'expérimentateur. Il est déjà difficile de faire comprendre qu'un même dé puisse être jugé régulier au vu d'un échantillon, non régulier au vu d'un autre échantillon. Il ne faudrait pas obscurcir cette compréhension en faisant croire que le seuil critique n'est pas le même en Terminale S1 et en Terminale S2.

C'est pourquoi il me semble qu'il faudrait ramener cette démarche expérimentale à ce qu'elle doit être en mathématiques : une activité d'introduction, indispensable pour motiver les notions à introduire, mais qui nécessite une institutionnalisation, à savoir l'énoncé d'un théorème.
Que dirait-on d'un professeur de mathématiques qui, ayant fait mesurer les côtés de divers triangles rectangles, et constater que la somme des carrés des côtés est à peu près égale au carré de l'hypoténuse, ne jugerait pas utile d'énoncer le théorème de Pythagore ?

En énonçant un théorème (admis bien entendu), on dépasse le questionnement ci-dessus. L'activité ne prétend rien démontrer, elle a un rôle heuristique, fondamental mais limité : faire manipuler et comprendre les résultats théoriques qui viennent ensuite.

Mais quel théorème énoncer ?

Pour ne pas trop compliquer une notion déjà délicate, on pourrait se limiter à un seul niveau de confiance, par exemple 95 %. Cela permettrait d'appeler *rare* ou *exceptionnel*, par définition, un événement qui a une probabilité inférieure à 0,05.

Mais on bute ici sur une difficulté, déjà rencontrée en Seconde avec la fourchette de sondage. Énoncer un théorème suppose pouvoir donner un résultat rigoureux et démontré, donnant la valeur de n à partir de laquelle la probabilité dépasse toujours 0,95.

Or à ma connaissance la théorie ne fournit pas actuellement un tel théorème. Elle indique seulement que la vitesse de convergence est meilleure quand les p_i ne sont pas trop proches de 0. On prend souvent comme critères[14] $n \geq 30$, $np_i \geq 1\ \forall\ i$, et au moins 80% des np_i au moins égaux à 5. Dans le cas de l'équiprobabilité, ces conditions se ramènent à $n \geq 30$ et $n \geq 5k$. Le commentaire des programmes stipule $n \geq 100$. Mais il s'agit simplement d'habitudes des statisticiens : aucun livre ne donne la validité de l'approximation quand ces conditions sont remplies. En fait il faudrait étudier la loi de $n\sum_{1}^{k}\frac{(f_i-p_i)^2}{p_i}$ pour pouvoir énoncer un véritable théorème[15], garantissant un niveau de confiance 0,95.

Notons que cette absence de précision justifie, dans le cas de l'équiprobabilité, de prendre $\frac{2}{n}$ comme seuil critique pour $\sum_{1}^{k}\left(f_i-\frac{1}{k}\right)^2$: raffiner en utilisant les tables du χ^2 constituerait un gain de précision illusoire, puisque la variable $kn\sum_{1}^{k}\left(f_i-\frac{1}{k}\right)^2$

(14) Voir par exemple [1], p. 354.
(15) Dans le cas $k = 2$, cette étude a été faite tout récemment : voir le Bulletin APMEP n° 436, pages 732-733.

ne suit qu'approximativement une loi du χ^2 .

En attendant que les chercheurs nous fournissent un théorème plus précis, je propose l'énoncé suivant, accessible en Terminale :
Soit une épreuve ayant *k* issues équiprobables. On la répète *n* fois et on note f_i les fréquences d'apparition des différentes issues. Si *n* est assez grand, il est exceptionnel que $\sum_1^k \left(f_i - \frac{1}{k}\right)^2$ dépasse $\frac{2}{n}$.

Volontairement les expressions « *n* est assez grand » et « il est exceptionnel » n'ont pas été précisées. On pourra signaler en commentaire que si $n \geq 30$ et $n \geq 5k$, il y a environ un échantillon sur 20 pour lequel $\sum_1^k \left(f_i - \frac{1}{k}\right)^2$ dépasse $\frac{2}{n}$.

Armés de ce théorème, les élèves pourront dès lors construire des tests d'équiprobabilité, sans devoir, chaque fois qu'une situation se présente, commencer par en simuler de multiples échantillons. Notamment, ils pourront tester les générateurs de nombres au hasard qui sont l'outil des simulations. Ils pourront avoir une véritable activité mathématique, c'est-à-dire raisonner. Surtout, et c'est là l'objectif principal, on peut espérer qu'ils auront une première idée de ce que peut être un test statistique.

Bibliographie :

- [1] Droesbeke J.-J., *Éléments de statistique* (Ellipses), 1997 (3e édition).
- [2] Spiegel M.-R., *Théorie et applications de la statistique* (McGraw-Hill), 1993 (2e édition).
- [3] Engel A., *Les certitudes du hasard* (Aléas), 1990.
- [4] Saporta G., *Probabilités, analyse des données et statistiques* (Technip), 1990.
- [5[Thiénard J.-C., *À propos de l'enseignement du calcul des probabilités* (IREM de Poitiers), 1993.
- [6] Veysseyre R., *Statistique et probabilités pour l'ingénieur* (Dunod), 2000.