

Statistique. Traiter l'adéquation à une loi équirépartie en classe terminale S ou ES : pourquoi ? à quel sujet ? comment ?

Jean-Pierre Raoult(*)

Résumé. *En s'appuyant sur une analyse critique de trois énoncés de baccalauréat relatifs à l'adéquation de données expérimentales à une loi équirépartie, cet article passe en revue les différents aspects de ce thème de la science statistique et s'interroge sur sa mise en œuvre dans l'enseignement, dans le cadre des programmes en vigueur des filières S et ES. On y prend position pour une distinction nette entre les travaux menés en classe, qui peuvent être l'occasion d'effectuer l'analyse de situations concrètes réalistes, dont on doit s'assurer qu'elles se prêtent bien à la méthodologie du test d'adéquation, et les exercices de contrôle, au registre inévitablement plus limité en matière de modélisation, mais qui doivent néanmoins respecter scrupuleusement l'énoncé des conclusions qu'autorise ce type de test statistique.*

Mots clés : Adéquation, Équirépartition, Loi du Khi-2, Quantiles, Simulation, Significatif, Terminale (lycées français), Test statistique.

1. Introduction

L'étude d'un exemple traitant de l'adéquation de données expérimentales à une loi équirépartie figure, dans ces termes, dans les programmes (mis en vigueur en 2001-2002) de Mathématiques des classes terminales des lycées français, pour les séries S (Scientifique) et ES (Économique et Sociale), au titre de contenus dénommés respectivement **Statistique et simulation** pour la série S et simplement **Simulation** pour la série ES. Dans les deux cas, il s'agit d'un « point extrémal absolu » du programme, en ceci qu'il n'a ni prolongement ni connexion latérale évidente avec une autre partie du programme de ces classes (si ce n'est, partiellement, la loi binomiale dans le cas très particulier de la loi équirépartie sur un ensemble à deux éléments). Cet item pourrait donc être retiré des programmes sans en déséquilibrer le déroulement ; ceci soulagerait peut-être nombre de professeurs qui n'avaient jamais eu de contact avec la notion de test statistique avant son introduction, par ce biais, dans le cursus de certains lycéens ; en revanche le caractère nouveau de ce thème a suscité de nombreux travaux d'enseignants qui se sont efforcés, en particulier dans le cadre des IREM (Instituts de Recherches sur l'Enseignement des Mathématiques) de mettre à la disposition de leurs collègues le fruit de leurs réflexions ou de leurs

(*) Laboratoire d'Analyse et de Mathématiques Appliquées Université de Marne-la-Vallée.
jean-pierre.raoult@univ-mlv.fr

expériences en classes de lycées⁽¹⁾, soit dans des publications, soit sur des sites Internet (citons en particulier les « classes virtuelles » d'Aromath) : [1], [2] (repris largement dans [3]), [4] (reprises et complétées numériquement et graphiquement – avec Géoplan – dans [5]), [9] (chapitre 5), [11], [16], [13] (avec fiches d'activités en classe), [17], [18] ; nous renvoyons aussi au chapitre 9 d'un ouvrage de Claudine Robert, destiné à un vaste public : [21].

Il est donc nécessaire de considérer attentivement le rôle formateur de ce chapitre du cours et, si celui-ci est reconnu, d'être particulièrement vigilant sur la correction de sa mise en œuvre et la crédibilité de son usage dans les épreuves d'examen.

Pour mener à bien ces considérations, il importe d'abord de bien noter les limites que les auteurs des programmes ont voulu donner à cette étape des cursus. On lit dans les commentaires : *L'élève devra être capable de poser le problème de l'adéquation à une loi équirépartie et de se reporter aux résultats de simulation qu'on lui fournira. Le vocabulaire des tests (hypothèse nulle, risque de première espèce) est hors programme.*

Ce cadrage a plusieurs aspects :

- volonté de mise en situation (élève « capable de poser le problème »),
- création d'occasions d'appliquer le programme de la classe de seconde (usage des « résultats de simulation »),
- simple initiation aux tests statistiques (« le vocabulaire des tests est hors programme » et donc leur problématique ne peut pas être traitée de manière générale).

Une conséquence de cette modeste volonté du programme est, bien sûr, que le registre des problèmes proposables au baccalauréat est très limité. Les exemples qu'on peut en trouver en Annales (je me référerai à deux d'entre eux, que je désignerai par la suite par les professions des personnes supposées avoir eu besoin de tester l'adéquation à une loi équirépartie, *pisciculteur* ou *banquier*) ou dans le recueil proposé en 2003 par l'Inspection Générale de Mathématiques (énoncé que je désignerai sous le code *meunier*) sont tous bâtis sur la même trame :

1. Présentation d'un contexte « concret » à k modalités, et fourniture, dans ce contexte, de données issues d'un échantillon, c'est-à-dire n observations indépendantes et de même loi, modèle dit i.i.d. (indépendance et identique distribution) qui doit être assuré pour fonder toutes nos études ici.
2. Calcul d'une « statistique de test » : à un facteur multiplicatif près, variant selon nos énoncés de référence, c'est le carré de la distance euclidienne entre la suite (f_1, \dots, f_k) des fréquences observées pour les k modalités possibles et la suite constante de terme général $1/k$:

$$d^2 = \sum_{i=1}^k \left(f_i - \frac{1}{k} \right)^2.$$

(1) Expérience que je ne possède pas ; mon approche est ici celle d'un universitaire, chercheur en statistique mathématique et praticien d'études avec différents types d'utilisateurs de la Statistique, mais mes analyses rejoignent très largement les leurs.

3. Fourniture d'informations sur la distribution de la statistique de test telle qu'elle résulte de simulations effectuées à l'aide de la loi équirépartie, puis demande d'en déduire la valeur d'un quantile d'ordre élevé de cette distribution (par exemple le neuvième décile, le quatre-vingt quinzième centile, ...).

4. Interrogation sur la conclusion relative à l'adéquation des fréquences observées à l'hypothèse d'équirépartition.

La prochaine section de cet article est consacrée à situer, à l'usage des enseignants, ce thème précis de l'adéquation à une loi équirépartie au sein de l'ensemble des problématiques de la Statistique Inférentielle. Ensuite je vais suivre la trame des problèmes types, en commentant à chaque fois les intentions du programme et leurs possibilités de réalisation, en fournissant les éléments théoriques qui me paraissent indispensables pour la culture de l'enseignant (m'appuyant en particulier pour cela, en section 3, sur le cas particulier où $k = 2$) et en distinguant ce qui, à mon sens, peut se faire en classe de ce qui peut être proposé en examen, distinction fondamentale que je préciserai en conclusion.

Trois annexes complètent cet article :

A. Nos trois énoncés de référence (celui que nous appelons ici *pisciculteur* est reproduit aussi dans [18], où on en trouve une analyse qui recoupe largement la mienne).

B. Quelques propositions de contextes réalistes utilisables en classe.

C. Une proposition d'énoncé « décontextualisé ».

En bibliographie, le lecteur trouvera deux types de références : des traités de statistique où se trouve développé le point qui nous intéresse ici et quelques articles ou chapitres d'ouvrages français récents participant à la réflexion sur ce point des programmes.

2. Mise en situation de ce point des programmes

2.1. Pourquoi un point de Statistique Inférentielle dans des programmes de Terminale ?

Rappelons que **la Statistique Inférentielle peut être caractérisée comme une science qui modélise une fraction observable du réel comme résultant d'un phénomène aléatoire pour lequel on envisage non pas une mais toute une famille de lois de probabilité possibles**. Dans ce cadre les outils de la Statistique Inférentielle servent à *inférer* à partir de *statistiques* (c'est-à-dire d'observations répétées de ce phénomène) des conclusions quant à celle de ces lois qui pourrait régir ce phénomène. Cette famille de lois est présentée en général comme indexée par un (ou plusieurs) paramètre(s) réel(s) et l'inférence porte donc sur ce(s) paramètre(s).

Le citoyen est constamment confronté à des affirmations qui, sans que cela soit toujours explicité, relèvent de la Statistique Inférentielle ; comme de nombreux autres auteurs (voir par exemple [10]), je considère que « renforcer les défenses » des jeunes à cet égard est un objectif primordial de l'éducation. De même les acteurs, en matière économique, industrielle, scientifique, ... , disposent de données qui, d'une manière ou d'une autre (choix d'échantillons, imprécisions des observations, ...) sont entachées d'aléatoire et à partir desquelles ils doivent prendre des décisions

pertinentes agissant sur un monde réel dont ils n'ont ainsi eu qu'une représentation partielle ; préparer les jeunes à l'exercice de telles responsabilités est aussi pour moi partie intégrante de leur formation.

Prenons l'exemple des statistiques électorales ; pour simplifier notre propos, nous allons nous intéresser à un scrutin à deux options possibles, par exemple un référendum à réponses « oui » ou « non » (nous aurions pu de manière analogue prendre l'exemple d'un ingénieur, dans un service « qualité » d'une entreprise, qui observe une machine pouvant produire des pièces bonnes ou défectueuses ou encore étudier, comme dans l'un des exercices des documents d'accompagnement du programme, la répartition des sexes des enfants à la naissance). Un sondage pré-électoral effectué sur, comme on dit, « un échantillon représentatif d'effectif n », peut, en première analyse et si on ne s'intéresse pas ici aux sondés qui se disent « indécis » ou « sans opinion », dont nous notons $n - m$ l'effectif observé dans l'échantillon, être considéré comme un tirage de m éléments dans une population constituée des électeurs qui, à la date précise de ce sondage, ont une opinion affirmée⁽²⁾. Notons p la proportion de partisans du « oui » dans cette population ; m étant petit devant l'effectif total de la population, on peut considérer qu'il s'agit d'un *tirage avec remise*, ce qui implique l'indépendance des m variables aléatoires « observations », avec, pour chacune d'entre elles, probabilité p pour le résultat « oui » et probabilité $1 - p$ pour le résultat « non ». Le nombre x de réponses « oui » suit alors une loi binomiale de taille m (CONNUE) et de paramètre p (INCONNU du sondeur) ; on tient là la *famille de lois de probabilité indexée par un paramètre* annoncée plus haut comme constitutive d'une situation relevant de la statistique inférentielle. Les affirmations, éventuellement entachées d'erreurs plus ou moins graves, que, **à partir de l'observation d'un nombre de « oui » égal à x** , on peut entendre proférer sur le paramètre inconnu p sont des types suivants :

– **p est « estimé » à telle valeur**, en pratique la fréquence observée des « oui », c'est-à-dire x/m ; par exemple on dira : avec $n = 1\ 000$, $m = 825$ et $x = 429$, p est estimé à

$$52\% \text{ (car } \frac{429}{825} = 0,52 \text{)} ;$$

– **p appartient à telle « fourchette »** et, pour être complet, on doit préciser « au niveau de confiance tant » ; par exemple on dira : avec $m = 825$ et $x = 429$, p est, au niveau de confiance 95%, compris entre 48,85% et 55,15%⁽³⁾ ;

– **l'observation faite permet (ou ne permet pas) de rejeter telle proposition qui aurait été avancée relativement à la valeur de p** et, pour être complet, on doit préciser (ce que nous développons en sous-section 3.1) : « avec un risque d'erreur de tant » (par exemple on dira : pour $m = 825$, l'observation de $x = 429$ ne permet pas (2) La réalité est en fait plus complexe ; on sait que les sondeurs procèdent par la méthode dite « des quotas » pour équilibrer la représentation de différentes fractions de la population et se livrent à des « redressements » en fonction de résultats antérieurs ; mais ce modèle grossier nous suffira ici et il permet d'avoir d'utiles ordres de grandeur sur la précision des résultats.

(3) Il s'agit de l'intervalle $\left[0,52 - \frac{1}{\sqrt{825}}, 0,52 + \frac{1}{\sqrt{825}}\right]$ dont l'usage est recommandé en page

49 du Document d'accompagnement des programmes de seconde.

de (ou « n'est pas significative pour ») rejeter, avec un risque d'erreur de 5%, la proposition selon laquelle $p \leq 50\%$).

Parmi les trois problématiques classiques de la statistique inférentielle que nous venons d'évoquer successivement sur cet exemple usuel (estimation ponctuelle, intervalle de confiance, test d'hypothèse), les deux premières ne figurent pas actuellement explicitement aux programmes de S ou ES des lycées. Elles sont cependant sous-jacentes au programme de la classe de Seconde sur la simulation et la fluctuation d'échantillonnage, mais non formulables à ce moment-là puisqu'on n'y dispose pas du vocabulaire des probabilités.

Pendant imaginons qu'un enseignant fasse faire à ses élèves la succession d'opérations suivantes :

- simuler avec la touche RANDOM des calculatrices n expériences résultant chacune en un nombre entier compris entre 0 et 99 ;
- faire affecter le code P aux nombres inférieurs ou égaux à 50 et le code F à ceux strictement supérieurs à 50 ;
- observer l'évolution, en fonction de n , de f , fréquence observée de résultats P.

Tout est en place pour :

- se convaincre que, plus n est grand, meilleure est l'approximation que cette proportion donne de 0,51 (aux termes du programme de la classe de seconde, notre procédure se décrit en disant que dans cette expérience il y a à chaque fois 51 chances sur 100 d'obtenir un P) ;

- construire autour de f l'intervalle $\left[f - \frac{1}{\sqrt{n}}, f + \frac{1}{\sqrt{n}} \right]$;

– interpréter cet intervalle comme cela est fait dans l'exercice intitulé *Sondages* dans le *Document d'accompagnement des programmes de classe de seconde* (CNDP, octobre 2000, pages 46 à 51) où l'on trouve les expressions *fourchette de sondage au*

niveau 0,95 et estimation de p avec une précision de $\frac{1}{\sqrt{n}}$ au niveau de confiance

0,95 (ici p est le « rapport de chances » égal à 0,51) ;

- constater que, si on avait cru que cette expérience simulait un jeu de Pile ou Face avec une pièce correcte, on serait amené, au fil de l'augmentation du nombre de simulations, à mettre de plus en plus en doute cette croyance (on constaterait, en renouvelant M fois cette batterie de simulations, que c'est environ pour $M = 10\,000$ que 0,50 n'appartiendrait plus qu'à environ un intervalle de confiance sur 20).

Et nous voici déjà à l'orée du test d'adéquation à la répartition équirépartie sur un ensemble à deux éléments !

Mais il ne suffit pas, quoique ce soit essentiel pour la formation, de se familiariser avec l'esprit de ces problématiques et avec les techniques calculatoires qui leur sont adaptées ; le rôle de la Statistique Mathématique est d'élaborer pour ces techniques des critères de qualité qui soient accessibles au calcul. Or l'introduction dans les cursus des classes terminales des critères relatifs à l'estimation ponctuelle et aux intervalles de confiance aurait été assez délicate à ce niveau de culture probabiliste (même en se limitant aux modèles simples fondés sur la loi binomiale qu'on vient

d'évoquer) et aurait en tout cas considérablement alourdi ces programmes⁽⁴⁾.

2.2. Pourquoi s'intéresser aux situations de tests d'adéquation ?

Face aux obstacles pédagogiques que nous venons de soulever à propos de l'estimation ou des intervalles de confiance, la difficulté conceptuelle est moindre (sans être négligeable !) quand on s'intéresse aux tests dits *d'hypothèse nulle simple*, c'est-à-dire ceux pour lesquels la proposition mise en jeu relativement à la loi de probabilité régissant les observations consiste en la spécification d'une seule loi particulière : on la teste « contre » (*versus*, disent les anglophones) toutes les autres mises en jeu dans le modèle.

Pour ma part, je suis heureux que les auteurs des programmes aient sélectionné une telle situation d'hypothèse simple, celle d'adéquation à une loi équirépartie. J'y vois plusieurs avantages que je vais détailler ici et que je m'efforcerai de préciser dans la suite de cet article.

a. On donne ainsi aux élèves **au moins une occasion de mener à bien le traitement maîtrisé, au vu des observations, d'une situation aléatoire à loi non totalement connue a priori**.

b. Le test d'adéquation à la loi équirépartie fournit, me semble-t-il, un **bon compromis pédagogique** entre des facilités de traitement mathématique, indispensables à ce niveau scolaire (j'y reviendrai aux sections 4 et 5) et le réalisme expérimental, car il s'agit de situations que l'on est effectivement assez souvent amené à envisager dans la pratique et qui dépassent heureusement le cadre trop restrictif des modèles à un paramètre réel auxquels nous nous étions limités jusqu'ici dans cet article.

c. Au regard de ma préoccupation principale s'agissant de l'enseignement de la statistique, qui est celle de la **formation critique du citoyen**, ce choix-là me paraît d'autant plus pertinent que les média nous abreuvent d'usages souvent inconsiderés de l'adjectif « significatif » et qu'on trouve ici un cadre idéal pour contrôler cet usage (à telle enseigne que, tout en comprenant que les commentaires du programme excluent l'emploi systématique du vocabulaire théorique relatif aux tests, je regrette un peu que ne soit pas maintenue la réflexion sur ce vocable clef de « significatif »).

d. S'agissant de la **formation à l'esprit scientifique**, on a ici un bel exemple de **modestie** dans les conclusions autorisées ; certes il faut avoir élaboré un modèle du phénomène aléatoire considéré, mais, aux termes de la théorie classique des tests, aucune des conclusions possibles n'est péremptoire⁽⁵⁾ :

(4) Pour l'estimation ponctuelle, il aurait fallu au moins présenter la notion d'estimateur sans biais et de comparaison des estimateurs à l'aide du risque quadratique moyen (certes occasions intéressantes de réflexion sur l'espérance mathématique et la variance) ; quant à la formule empirique de calcul des intervalles de confiance à 95% déjà évoquée en note 3 ci-dessus, il faudrait pour la justifier (et aller au delà en faisant varier le niveau de signification) tout un travail, difficilement envisageable à ce niveau d'études, sur l'approximation normale des lois binomiales, et ce en faisant varier le paramètre de celles-ci. La faisabilité de cette tâche est cependant étudiée dans [15].

(5) Le lecteur non familier avec la théorie des tests pourra revenir à ces considérations après avoir lu la sous-section 3.1 ci-dessous, où j'en rappelle les fondements en m'appuyant sur un exemple bien connu de test d'adéquation pour deux modalités.

- ou bien on rejette l’hypothèse testée (ici celle de la répartition équirépartie), mais d’une part on n’est pas certain de cette conclusion (on contrôle seulement sa probabilité d’erreur, sans pouvoir l’annuler) et d’autre part on n’est pas pour autant armé pour proposer une autre répartition ;
- ou bien on ne rejette pas l’hypothèse testée, mais cela ne veut pas dire pour autant qu’on est assuré de sa validité : on va seulement continuer à la tenir pour utilisable, faute d’avoir su l’infirmier.

3. Où aller choisir des contextes « expérimentaux » ?

3.1. Rappel sur la théorie classique des tests

Pour bien comprendre les précautions que l’on doit prendre pour présenter aux élèves des situations de tests d’adéquation qui ne soient pas par trop artificielles, il nous faut nous placer un instant dans le cadre le plus général de la théorie des tests (avec ses notations traditionnelles). Je vais en dessiner ici les principes (brièvement introduits aussi dans [16]) ; le lecteur qui se serait déjà familiarisé avec cette théorie lors de ses études pourra se reporter à un traité connu de lui ; j’en fournis ici en bibliographie quelques-uns, publiés en France et dans les quinze dernières années : [6], [7], [12] [19], [22].

On dispose d’une observation ω appartenant à un ensemble Ω (par exemple $\Omega = \mathbf{R}^n$ s’il s’agit d’échantillons de taille n d’observations numériques, ou $\Omega = A^n$ s’il s’agit de n tirages dans une population « matérielle » $A^{(6)}$) et on a adopté un modèle selon lequel la loi de probabilité, sur Ω , régissant ces observations est l’une des lois P_θ , où le paramètre θ appartient à un ensemble Θ . Étant fixée une partie Θ_0 de Θ , il est demandé de *tester l’hypothèse (dite « nulle »), selon laquelle le paramètre inconnu θ appartient à Θ_0 , contre l’hypothèse alternative (aussi dite « contre-hypothèse »), selon laquelle θ appartient à Θ_1 , complémentaire de Θ_0 dans $\Theta^{(7)}$.*

Dans ce cadre, une « méthode de test » (on dira en bref « un test ») va consister en la détermination d’une partition (Ω_0, Ω_1) de Ω telle que, grossièrement dit (on le raffinerait un peu plus loin), on annonce :

(6) Certains auteurs dans ce cas notent Ω la population ; ceci peut se révéler dangereux car il est traditionnel en modélisation probabiliste que Ω renvoie à l’espace mesurable sur lequel sont définies toutes les variables aléatoires (v.a.) en jeu dans l’étude et, les n observations devant impérativement être identifiées comme autant de v.a. distinctes, ceci n’est possible que si elles sont définies sur A^n ; de plus c’est sur A^n que, quand le modèle le justifie, vont être formulées les hypothèses d’indépendance.

(7) Attention à l’emploi du mot « hypothèse » : adopter un modèle (caractérisé ici par la famille de lois $(P_\theta)_{\theta \in \Theta}$) est bien sûr faire une certaine « hypothèse » sur le phénomène considéré ; on dira ainsi, par exemple, qu’on « suppose » (ou qu’on « admet ») avoir affaire à des observations i.i.d. ; l’étude envisagée n’a pas pour objet de remettre en cause ce modèle. En revanche l’hypothèse nulle caractérise un « sous-modèle » sur la validité duquel on s’interroge. La langue anglaise est ici plus riche, qui peut distinguer entre *assumptions* pour les hypothèses qui sous-tendent le modèle et *hypotheses* pour les hypothèses en jeu dans le test (nulle et alternative) ; on sait que le verbe *to assume* a un sens plus fort que *to suppose*.

- une conclusion en faveur de l'appartenance du paramètre inconnu θ à Θ_0 si l'observation ω appartient à Ω_0 ,
- une conclusion en faveur de l'appartenance de θ à Θ_1 si ω appartient à Ω_1 .

À tout tel choix de la partition (Ω_0, Ω_1) sont inévitablement associées des possibilités d'erreurs : on se trouve proclamer une conclusion erronée dans les deux types de situations suivantes :

- si ω « tombe » dans Ω_1 alors qu'en fait θ appartient à Θ_0 (erreur dite de *première espèce*, consistant à rejeter, à tort, l'hypothèse d'appartenance de θ à Θ_0) ;
- si ω « tombe » dans Ω_0 alors qu'en fait θ appartient à Θ_1 (erreur dite de *seconde espèce*, consistant à ne pas rejeter, à tort, l'hypothèse d'appartenance de θ à Θ_0).

Il en résulte toute une famille de valeurs de probabilités d'erreur :

- les $P_\theta(\Omega_1)$ pour tous les θ appartenant à Ω_0 (valeurs dites aussi *risques de première espèce*),
- les $P_\theta(\Omega_0)$ pour tous les θ appartenant à Ω_1 (valeurs dites aussi *risques de seconde espèce*).

Les comportements de ces probabilités d'erreur sont antinomiques : par exemple, grossir Ω_0 au détriment de Ω_1 a pour effet salutaire de diminuer les probabilités d'erreur de première espèce mais, hélas, en revanche, augmente celles des erreurs de seconde espèce ; en d'autres termes, au regard de l'ensemble des probabilités d'erreur, il n'y a pas de test (i.e. de partition (Ω_0, Ω_1)) optimal.

C'est pour sortir de ce dilemme que J. Neymann et E. S. Pearson (fils de K. Pearson dont nous évoquons en 4.2 les résultats sur la loi du χ^2) ont, dans un célèbre article paru en 1928 (voir le chapitre 5 dans la monographie d'histoire de la statistique [8] ou encore le chapitre 3 (*Grande et petite histoire de la statistique*) dans [9]), systématisé la hiérarchisation des deux espèces d'erreurs (dont la dénomination « première » et « seconde » cesse donc d'être conventionnelle) en proposant d'introduire un « petit » nombre α , strictement positif, dit *niveau de signification*, dont le choix est laissé au praticien qui effectue l'étude ; la priorité donnée à l'erreur de première espèce se traduit alors par le fait que l'on impose aux probabilités des erreurs de première espèce d'être inférieures ou égales à α ; et c'est sous cette contrainte que l'on s'efforce de choisir la partition (Ω_0, Ω_1) en minimisant les erreurs de seconde espèce⁽⁸⁾. Autrement dit, introduisons la *fonction puissance du test* qui à tout $\theta \in \Theta$ associe $\pi(\theta) = P_\theta(\Omega_1)$; on s'efforce de maximiser, sous la contrainte

$$\forall \theta \in \Theta_0 \quad \pi(\theta) \leq \alpha,$$

les valeurs de la fonction π restreinte à la contre-hypothèse Θ_1 (sur laquelle la puissance s'interprète en termes de probabilités de donner la bonne réponse, qui est

(8) Prendre α nul conduirait à refuser de rejeter l'hypothèse nulle quelles que soient les données ; c'est sur ce « principe de précaution », souvent invoqué de nos jours mais ici poussé jusqu'à la paralysie, que l'énoncé *meunier* invite les candidats à réfléchir par sa question 3.c ; sans doute serait-ce un peu troublant pour eux.

alors le rejet de l'hypothèse nulle)⁽⁹⁾.

On voit alors l'énorme simplification apportée par les situations dites d'*hypothèse nulle simple* où Θ_0 est un singleton $\{\theta_0\}$ et on comprend que le programme des classes de terminale S ou ES se soit limité aux situations de ce type. En effet alors la contrainte imposée par le *niveau de signification* nécessite uniquement de s'assurer que $P_{\theta_0}(\Omega_1) \leq \alpha$ ⁽¹⁰⁾, autrement dit $P_{\theta_0}(\Omega_0) \geq 1 - \alpha$. Ce sont souvent alors des considérations de « bon sens » qui conduisent à déterminer la « forme » des partitions de Ω « candidates » pour constituer le test (Ω_0, Ω_1) .

Considérons l'exemple de la répartition des sexes à la naissance, observée sur un échantillon de taille n , et prenons par convention (comme dans les numéros de Sécurité Sociale) les codes 1 pour « garçon » et 2 pour « fille », d'où $\Omega = \{1, 2\}^n$. Le couple (p_1, p_2) sera noté $(\theta, 1 - \theta)$; θ , compris entre 0 et 1, est donc le paramètre de ce modèle statistique ; tester l'équiprobabilité revient alors à tester l'hypothèse nulle :

$$\Theta_0 = \{\theta_0\} = \left\{ \frac{1}{2} \right\}.$$

Le « bon sens » nous conduit à rejeter cette hypothèse nulle si f , fréquence observée de naissances de garçons, apparaît suffisamment éloignée de 1/2. En d'autres termes, on va devoir déterminer un nombre $c > 0$ de sorte que le rejet de l'hypothèse nulle s'effectue

quand $\left| f - \frac{1}{2} \right| > c$. Ceci revient à remplacer l'ensemble Ω initial, dans lequel se trouvait

l'observation « brute », par l'intervalle $[0, 1]$ auquel appartient la fréquence f ⁽¹¹⁾ et, dans

cet intervalle, on va prendre pour Ω_0 un intervalle de la forme $\left[\frac{1}{2} - c, \frac{1}{2} + c \right]$. Comment

choisir cette frontière c entre rejet et non-rejet ? C'est là qu'intervient le niveau de signification qu'on s'est fixé au préalable, α : on utilise le fait que le nombre de naissances de garçons, soit $x (= nf)$, suit la loi binomiale de taille n et paramètre θ , notée

(9) Dans certains problèmes, par exemple en contrôle de la qualité industrielle, il arrive qu'on exige aussi une majoration des risques de seconde espèce (autrement dit une minoration de la puissance) sur une partie de la contre-hypothèse, soit $\Theta'_1 \subset \Theta_1$; on introduit alors un « petit »

nombre β , strictement positif, et on exige $\forall \theta \in \Theta'_1, 1 - \pi(\theta) \leq \beta$; souvent cette exigence n'est pas réalisable pour n'importe quelle valeur de la taille de l'échantillon, n , et un problème classique est de déterminer une valeur minima de cette taille pour qu'elle le soit.

(10) En fait, dans de nombreuses situations où Θ_0 n'est pas réduit à un seul élément (les statisticiens parlent alors d'hypothèse nulle multiple), des considérations de calcul des probabilités permettent de déterminer $\sup_{\theta \in \Theta_0} P_{\theta}(\Omega_1)$; il en est ainsi dans la situation évoquée ci-dessus en 2.1 où, s'agissant de la proportion p de votes « oui » dans un référendum, on voulait tester l'hypothèse selon laquelle $p \leq 50\%$; alors le choix classique de la forme de Ω_1 (consistant à rejeter l'hypothèse nulle si le nombre x de votes « oui » dépasse une certaine frontière) vérifie élémentairement $\sup_{p \leq 0,5} P_{0,5}(\Omega_1)$.

(11) En théorie de la Statistique Inférentielle, on justifie mathématiquement le fait d'abandonner la donnée initiale (liste des sexes de tous les enfants de l'échantillon) au profit de cette seule fréquence par le fait que celle-ci constitue une *statistique exhaustive* (on dit aussi un *résumé exhaustif*) dans ce modèle i.i.d..

$B_{n,\theta}$; en particulier, si l'hypothèse nulle est satisfaite, elle suit la loi $B_{n, \frac{1}{2}}$. Donc les seules valeurs de c admissibles sont celles qui vérifient

$$B_{n, \frac{1}{2}} \left(\left\{ x : \frac{n}{2} - nc \leq x \leq \frac{n}{2} + nc \right\} \right) \geq 1 - \alpha,$$

autrement dit

$$B_{n, \frac{1}{2}} \left(\left\{ x : \frac{n}{2} + nc < x \leq n \right\} \right) \leq \frac{\alpha}{2},$$

(car, par symétrie, on a alors aussi $B_{n, \frac{1}{2}} \left(\left\{ x : 0 \leq x < \frac{n}{2} - nc \right\} \right) \leq \frac{\alpha}{2}$).

Nous poursuivons la considération de cet exemple, avec ces mêmes notations, tout au long de ce passage.

Ensuite, parmi de telles partitions, c'est toujours le bon sens qui nous conduit à en choisir une qui assure une valeur $P_{\theta_0}(\Omega_1)$ aussi proche que possible, par défaut, de α , de manière à, en contrepartie, maximiser (au moins approximativement) la puissance sur Θ_1 .

Dans l'exemple de la répartition des sexes à la naissance, grossir Ω_1 se fait en diminuant la frontière c , et donc on doit choisir c aussi petit que possible, compatible avec

la contrainte $B_{n, \frac{1}{2}} \left(\left\{ x : \frac{n}{2} + nc < x \leq n \right\} \right) \leq \frac{\alpha}{2}$; si on note $G_{n,\theta}$ la fonction de répartition

(f.r.) de $B_{n,\theta}$, cette contrainte s'écrit : $1 - G_{n, \frac{1}{2}} \left(\frac{n}{2} + nc \right) \leq \frac{\alpha}{2}$.

En particulier, si cela est mathématiquement possible, on choisira cette partition (Ω_0, Ω_1) de sorte que $P_{\theta_0}(\Omega_1) = \alpha$.

La f.r. $G_{n,\theta}$ n'est pas continue, puisque les lois binomiales sont discrètes ; mais il résulte du théorème de la limite centrale que, pour n grand, on l'approche par l'application

continue $t \mapsto \Phi \left(\frac{t - n\theta}{\sqrt{n\theta(1-\theta)}} \right)$, où Φ désigne la f.r. de la loi normale centrée réduite⁽¹²⁾. On

approximera alors $G_{n, \frac{1}{2}} \left(\frac{n}{2} + nc \right)$ par $\Phi(2\sqrt{nc})$ et donc, à l'équation

$1 - G_{n, \frac{1}{2}} \left(\frac{n}{2} + nc \right) = \frac{\alpha}{2}$, on affectera la solution approchée $c_{n,\alpha} = \frac{t_\alpha}{2\sqrt{n}}$ où t_α désigne le

quantile d'ordre $1 - \frac{\alpha}{2}$ de la loi normale centrée réduite⁽¹³⁾,

(12) Cette approximation est traditionnellement jugée assez satisfaisante si $np(1-p) \geq 5$; pour $p = 1/2$, ceci donne $n \geq 20$ mais le plus souvent on ne l'utilise que pour $n \geq 30$.

(13) Pour $\alpha = 0,95$, on a $t_\alpha = 1,96$, donc $c \approx \frac{1}{\sqrt{n}}$; on constate que le rejet de l'hypothèse

Nous verrons en section 4 comment ces considérations « de bon sens » jouent dans l'exemple de l'adéquation à la loi équirépartie à k modalités (pour laquelle on vient en fait, en s'intéressant au test de l'hypothèse $\theta = 1/2$ où θ est le paramètre d'une loi binomiale, de donner une méthode de présentation dans le cas particulier $k = 2$).

Le choix du niveau de signification α s'effectue en général parmi un petit nombre de valeurs traditionnelles : 0,10 ; 0,05 ; 0,01 ; ... Plus α est petit, plus on dit que le test est « sévère » ; en effet, plus α est petit, plus restreinte est la « région de rejet » (on dit aussi « région critique ») Ω_1 , ensemble des résultats que l'on est autorisé à juger « significatifs » pour rejeter l'hypothèse nulle. En d'autres termes, si l'on cite le *Document d'accompagnement des classes de terminale*, page 146 (dans le cadre, à nouveau, de l'étude de la répartition des sexes à la naissance) : *On prend souvent par défaut $\alpha = 0,05$: il est important que les élèves sachent qu'il s'agit d'un consensus et non d'une constante immuable. L'enjeu est de comprendre sur quoi porte le risque (refuser à tort le modèle) et que, plus le risque est petit, plus on aura tendance à accepter le modèle de l'équiprobabilité.*

Dans l'exemple du test de la valeur 1/2 pour le paramètre de la loi binomiale (introduit ici à partir du problème de l'équirépartition des sexes à la naissance), cet accroissement de sévérité quand α décroît s'apprécie dans le tableau suivant, qui donne, à l'aide de

l'approximation normale, la valeur de $c_{n,\alpha} = \frac{t_\alpha}{2\sqrt{n}}$ pour trois valeurs usuelles de α et trois valeurs de n (on rappelle qu'il y a possibilité de rejet de l'hypothèse nulle si et seulement si la fréquence observée s'écarte de 1/2 de plus de $c_{n,\alpha}$).

	$\alpha = 0,1$	$\alpha = 0,05$	$\alpha = 0,01$
	$t_{0,1} = 1,645$	$t_{0,05} = 1,960$	$t_{0,01} = 2,576$
$n = 100$	$c_{100;0,1} = 0,082$	$c_{100;0,05} = 0,098$	$c_{100;0,01} = 0,129$
$n = 400$	$c_{400;0,1} = 0,041$	$c_{400;0,05} = 0,049$	$c_{400;0,01} = 0,065$
$n = 1000$	$c_{1000;0,1} = 0,026$	$c_{1000;0,05} = 0,031$	$c_{1000;0,01} = 0,041$

Le fait, observé sur le tableau, que, à α fixé, quand n augmente, $c_{n,\alpha}$ décroisse (à la vitesse $\frac{1}{\sqrt{n}}$) correspond à notre intuition : plus n est grand, mieux on est informé et donc plus facilement une différence donnée entre 1/2 et f peut nous apparaître significative pour procéder à un rejet de l'hypothèse d'équirépartition⁽¹⁴⁾.

nulle s'effectue si 1/2 n'appartient pas à l'intervalle de confiance à 95% déjà évoqué aux notes 3 et 5 ; cette heureuse coïncidence entre intervalle de confiance et test n'est pas une propriété générale ; elle est due aux propriétés de symétrie de la f.r. Φ .

(14) Cette intuition est parfois mal interprétée par des praticiens qui en retirent le sentiment que, à condition d'expérimenter assez longtemps, on arrivera « toujours » à rejeter l'hypothèse nulle ; il n'en est, bien sûr, rien : par exemple, pour $\alpha = 0,05$, la fréquence $f = 0,54$, distante de 0,5 de 0,04, ne conduit pas au rejet pour $n = 100$ et $n = 400$, mais y conduit pour $n = 1000$; cette circonstance est uniquement due au fait que, en cas d'équirépartition, atteindre (et dépasser) un tel éloignement de 0,5 devient beaucoup moins probable pour $n = 1000$ que pour les autres valeurs de n , plus faibles, considérées ici.

Précisons que le vocabulaire relatif au niveau de signification est un peu fluctuant ; on trouve parfois *seuil de signification* ; le document d'accompagnement du programme dit (en p. 146) *risque* ; les énoncés de problèmes donnés en Annexe A parlent de *risque d'erreur inférieur à ...* Ces deux dernières expressions ont pour moi le défaut de masquer qu'il s'agit uniquement du risque de première espèce ; je m'y rallie pour respecter le désir louable du programme de ne céder à aucune inflation de vocabulaire, mais je préciserai (en section 6) de quelles précautions de langage on doit s'entourer pour éviter les confusions possibles.

3.2. Quand l'emploi de la théorie classique des tests se justifie-t-il ?

Il importe de ne recourir à la théorie classique des tests que lorsque la situation étudiée justifie véritablement la hiérarchisation opérée entre erreurs de première et seconde espèces.

Limitons-nous au cas précis retenu par les programmes, pour lequel l'hypothèse nulle est simple et consiste en la répartition équirépartie sur l'ensemble fini dans lequel sont faites les observations indépendantes qui constituent l'échantillon. Il est alors essentiel de pouvoir considérer que, d'une manière ou d'une autre, « il en coûterait » au praticien de ne pas utiliser ce modèle d'équirépartition. Cette idée de coût peut être prise au sens propre ou au sens figuré : ainsi, prendre une décision pratique fondée sur l'équirépartition peut être ce qu'il y a économiquement de moins onéreux ; ou bien, dans une étude scientifique, l'hypothèse d'équirépartition a des avantages d'antériorité ou de simplicité tels qu'il faille vraiment avoir des raisons déterminantes pour s'en écarter ; ou encore rejeter l'hypothèse d'équirépartition peut avoir des conséquences sociales, voire morales ou juridiques, lourdes (c'est ainsi que, dans le troisième exemple développé en Annexe B, rejeter l'hypothèse signifie accuser un commerçant de fraude). L'exemple de la répartition des sexes à la naissance rentre à l'évidence dans ce cadre puisqu'à première vue l'équirépartition y semble évidente ; et l'on sait qu'il faut considérer d'assez gros échantillons pour se convaincre du contraire, lequel semble d'autant plus surprenant qu'aucune justification biologique simple n'apparaît pour cela.

À la lumière de cette exigence, considérons nos problèmes de référence fournis en Annexe.

Dans le problème du *banquier*, on peut considérer qu'affecter les mêmes effectifs d'employés aux guichets tous les jours de la semaine est ce qui est le plus facile et le plus économique comme organisation ; or, à première vue, ces données peuvent inciter à penser qu'il y a moins de clients le mardi ; le chef d'agence ne veut en tirer la conclusion qu'il faut réduire l'effectif ce jour là au profit d'un autre jour (de préférence le samedi) que s'il peut le faire « à bon escient », c'est-à-dire avec une faible probabilité d'erreur ; par ailleurs, à propos du libellé de cet énoncé, on peut remarquer qu'il est peu vraisemblable que le chef d'agence se décide à partir des observations sur une seule semaine !

Dans le problème du *meunier*, s'il est vrai que la confection de la farine nécessite quatre variétés à parts égales, cet industriel dispose sans doute d'un appareillage qui, s'il fonctionne bien, assure cette égalité, sinon exactement, du moins avec une

précision suffisante pour assurer la qualité requise. *A priori*, il considère que ce dispositif fonctionne ; s'il n'en est pas ainsi, ceci implique pour lui des mesures désagréables (arrêt de la production, frais de réparation, voire rejet – ou du moins vente à prix réduit – d'une partie de la production) auxquelles il ne recourt que si cela apparaît vraiment indispensable ; il s'agit là d'une situation de *contrôle de la qualité* tout à fait classique en statistique appliquée.

Dans l'exemple du *pisciculteur*, la distinction entre erreurs de première et seconde espèces me semble beaucoup plus difficile à justifier ; en quoi devoir renoncer à considérer que « le bassin contient autant de truites de chaque variété » serait-il douloureux ? et la pertinence de l'hypothèse d'équirépartition est d'emblée fortement mise en question par le fait que les effectifs exacts varient sans arrêt, au rythme des prises effectuées dans le bassin (mais bien sûr renoncer à considérer cette hypothèse interdirait de poser la dernière question du problème, exercice élémentaire de calcul des probabilités qui m'apparaît ici un peu comme une pièce rapportée !).

Ces considérations sur les justifications des trois problèmes peuvent sembler au lecteur un peu anecdotiques. Je ne le pense pas.

La crédibilité du cours de statistique auprès des élèves suppose que les situations qui leur sont présentées soient à la fois raisonnablement réalistes et bien conformes à la problématique des tests.

S'il en est ainsi, ces situations peuvent donner lieu à une présentation argumentée du contexte, à une discussion sur celui-ci et même éventuellement à la réflexion sur des prolongements possibles. Sinon, il y aurait un risque que s'instaure dans l'esprit de l'élève une impression de « gratuité » de ce qu'on lui fait faire, et c'est là ce que, pour ma part, je redoute le plus en ce qui concerne cet enseignement.

C'est donc **en prenant son temps en classe pour des considérations « autour » de l'énoncé stricto sensu que l'on peut espérer ancrer dans l'esprit des élèves un autre aspect essentiel de la dissymétrie consubstantielle à la théorie classique des tests** : comme nous l'annoncions au point d de la sous-section 2.2, seul le rejet de l'hypothèse nulle est une vraie affirmation (faite en assumant sa probabilité d'erreur, connue, α , associée à cette technique) ; si on ne la rejette pas, ceci ne veut en rien dire qu'on affirme (*même en y adjoignant une probabilité d'erreur*) que l'hypothèse nulle est satisfaite.

En allant ici au delà de la culture des élèves, je vais préciser ce distinguo en considérant le problème le plus général à hypothèse nulle simple $\{\theta_0\}$. L'ensemble Θ des paramètres indexant les probabilités susceptibles de régir l'observation est muni d'une topologie : nous n'envisageons pas d'autres cas que ceux où cet ensemble de paramètres est plongé dans \mathbf{R}^h pour une valeur de h convenable et où donc on dispose d'une topologie naturelle. Admettons que le test utilisé soit tel que sa fonction puissance ($\theta \mapsto P_\theta(\Omega_1)$) soit continue (il en est bien ainsi pour le test dont nous disposons pour l'hypothèse d'équirépartition et ceci se généralise à tous les « bons cas »). La puissance reste donc proche de $P_{\theta_0}(\Omega_1)$, qui est majoré par (et en fait souvent égal à) α , pour les valeurs de θ dans Θ_1 , mais proches de θ_0 ; il s'agit

formellement de valeurs du paramètre en dehors de l'hypothèse nulle, et on pourrait se sentir gêné par le fait que la probabilité d'y rejeter celle-ci reste faible, comme l'est α . Ce n'est pas choquant dans la perspective « décisionnelle » que nous avons adoptée jusqu'ici, en particulier dans mes commentaires aux exemples : dans la pratique, le non-rejet de l'hypothèse conduit à adopter ensuite (en pratique ceci veut souvent dire « continuer à adopter ») un comportement jugé convenable quand θ vaut Θ_0 et il n'y a rien de grave à ce que cette valeur ne soit réalisée qu'approximativement.

Revenant à ce que l'on peut analyser avec les élèves, on peut par exemple leur expliquer que l'exigence du *meunier* n'est pas que son mélange de variétés soit très exactement réalisé à raison d'un quart pour chacune ; sa farine gardera sa saveur si cet équilibre n'est respecté que « à peu près » ; en fait, s'il doit respecter des normes officielles comme c'est souvent le cas en contrôle de qualité, celles-ci comportent toujours des marges d'écarts autorisées par rapport au standard. De même, dans le problème du *banquier*, il est difficile d'imaginer une indépendance absolue de la fréquentation par rapport aux jours de la semaine. Enfin, dans le problème du *pisciculteur*, j'ai déjà fait remarquer que les proportions de truites dans le bassin sont par nature variables, ce que d'ailleurs l'énoncé correspondant traduit en prenant la précaution de parler en introduction de la possibilité de « considérer » que le bassin contient autant de truites de chaque variété (précaution hélas non reprise en compte dans la question 3 de cet énoncé).

Prendre ainsi du recul par rapport aux énoncés d'exercices me paraît indispensable pour avoir une chance de donner un sens à la recommandation « optimiste » du programme : *L'élève devra être capable de poser le problème de l'adéquation à une loi équirépartie.*

Mais bien sûr ceci est réalisable en classe et non dans le cadre restrictif d'un sujet d'examen ! Certes on rencontre dans de nombreux enseignements cette contradiction entre d'une part le besoin de développer l'initiative et l'esprit critique des jeunes face à un contexte pratique et d'autre part les contraintes inhérentes à la situation d'examen ; ce dilemme est particulièrement crucial en Mathématiques Appliquées.

Au titre des prolongements possibles d'un énoncé, reprenons par exemple le problème du *banquier* : même si, ce qui est le cas ici, les données ne conduisent pas à rejeter, au niveau de signification proposé (ici 10%) l'hypothèse d'équirépartition, le banquier peut se souvenir que des clients se sont plaints de l'affluence du samedi, alors que le mardi est traditionnellement un jour creux (ce que semblent à première vue confirmer les données présentes). Il peut donc se demander s'il ne serait pas judicieux d'ouvrir un guichet de plus le samedi en réaffectant à ce jour là un employé prévu pour le mardi ; pour cela il va vouloir concentrer son étude sur ces deux seuls jours, en ne polluant pas cette comparaison par les résultats intermédiaires du mercredi au vendredi. Je reviendrai en section 5 sur la manière dont un professeur pourrait, me semble-t-il, assurer avec sa classe l'étude d'un tel prolongement.

4. Comment présenter la statistique de test

4.1. On peut « faire simple »

J'ai dit plus haut que ce sont souvent des arguments de simple bon sens qui peuvent présider au choix de la « forme » de la partition (Ω_0, Ω_1) qui constitue le test.

Quand il s'agit de tester l'équirépartition sur un ensemble fini de résultats, à k modalités, qu'on va noter par convention $\{1, \dots, k\}$, ce bon sens s'appuiera, aux yeux des élèves, sur les acquis du cours de seconde : si on note, pour tout i (où $1 \leq i \leq k$), p_i la probabilité de réalisation du résultat i , dans chacune des n expériences indépendantes qui ont fourni l'échantillon étudié, alors chaque fréquence f_i relevée dans l'échantillon observé fournit une approximation du p_i inconnu correspondant, d'autant meilleure que n est plus grand⁽¹⁵⁾.

Le professeur est donc en mesure d'expliquer aux élèves qu'il est d'autant plus raisonnable de rejeter l'hypothèse d'équirépartition que la suite des fréquences

observées (f_1, \dots, f_k) est plus éloignée de la suite constante $\left(\frac{1}{k}, \dots, \frac{1}{k}\right)$; en effet, cette

suite (f_1, \dots, f_k) étant, elle, en général, proche de (p_1, \dots, p_k) , plus elle est loin de $\left(\frac{1}{k}, \dots, \frac{1}{k}\right)$, moins il est crédible que les p_i soient tous égaux à $\frac{1}{k}$. Le problème est

alors de choisir un indicateur de cet éloignement de la suite des fréquences par

rapport à la suite constante de terme général $\frac{1}{k}$ et il paraît tout à fait naturel (en particulier aux yeux des élèves de terminale qui la connaissent déjà) de considérer la

distance euclidienne ou, ce qui revient au même, son carré $d^2 = \sum_{i=1}^n \left(f_i - \frac{1}{k}\right)^2$, et

de décider que l'on pourra rejeter l'hypothèse d'équirépartition si cet éloignement d^2 est assez flagrant⁽¹⁶⁾ ; au delà de quelle frontière considère-t-on qu'il en est ainsi ?

C'est là qu'intervient le rôle de α ; en effet la probabilité de rejeter à tort l'équirépartition étant celle d'observer d^2 supérieur à cette frontière quand l'équirépartition a lieu, c'est elle dont on doit assurer qu'elle est dominée par α ; il

(15) En termes de théorie de l'estimation, f_i est une estimation sans biais de p_i ; on dispose là d'un estimateur de variance minima (autrement dit ici d'erreur quadratique minima) parmi tous les estimateurs sans biais et de plus il est consistant (autrement dit f_i converge en probabilité vers la vraie valeur p_i quand la taille n de l'échantillon tend vers l'infini).

(16) Dans le cas où $k = 2$, si on pose, comme en sous-section 3.1, $(p_1, p_2) = (\theta, 1 - \theta)$ et par

ailleurs $(f_1, f_2) = (f, 1 - f)$, on a $d^2 = 2\left(f - \frac{1}{2}\right)^2$; on retrouve donc, avec une présentation

différente, le même test que celui présenté en 3.1, qui était fondé sur $\left|f - \frac{1}{2}\right|$; la note 23

précisera la comparaison de ces techniques.

a déjà été expliqué en sous-section 3.1 ci-dessus pourquoi il est alors naturel de choisir la frontière en rendant cette probabilité aussi proche que possible, par défaut de α (voire égale à α si c'est possible) de manière à minimiser autant que faire se peut les probabilités des erreurs de seconde espèce⁽¹⁷⁾.

4.2. Pourquoi, dans certains énoncés, avoir « fait compliqué » ?

Aucun des trois énoncés que je donne en référence ne propose de travailler

directement sur $d^2 = \sum_{i=1}^n \left(f_i - \frac{1}{k} \right)^2$, mais il y est introduit à chaque fois un facteur multiplicatif. Ce facteur est en fait :

- pour le *banquier* : 1 000,
- pour le *pisciculteur* : n (taille de l'échantillon),
- pour le *meunier* : nk (où k est le nombre de modalités des observations, soit ici 4).

Mettons à part le cas du *banquier*, pour lequel l'énoncé justifie son choix par l'explication : *la multiplication par 1 000 permet d'avoir un résultat plus lisible* ; je regrette pour ma part que, en Terminale, on craigne que des résultats presque tous compris entre 0 et 10^{-2} soient « mal lisibles » ; il aurait suffi de porter sur la figure donnée en question 3 (« diagramme en boîtes ») une échelle non en entiers mais en multiples de 10^{-3} ; il me semble un peu dommage de sembler dire que le mathématicien voit une difficulté d'usage dans des outils que simultanément le professeur de physique manie constamment dans son cours.

Les facteurs introduits dans les deux autres problèmes trouvent leur origine (partiellement pour le *pisciculteur*, intégralement pour le *meunier*) dans la théorie des tests du χ^2 dont nous ne considérons ici qu'un cas particulier. Il revient à K. Pearson d'avoir établi dès 1900 que, quand n tend vers l'infini et si la loi de probabilité sur $\{1, \dots, k\}$ est celle donnée par la suite (p_1, \dots, p_k) , alors la loi de la statistique

$$t = n \sum_{i=1}^n \frac{(f_i - p_i)^2}{p_i}$$
 converge, quand n tend vers l'infini, vers la loi dite du χ^2 à

(17) De manière un peu plus savante, pour la culture du professeur, on peut préciser que cette fabrication « intuitive » de la région de rejet Ω_1 assure l'optimisation de la puissance, c'est-à-dire la maximisation des probabilités de rejet à bon escient de l'équirépartition (sans qu'il soit besoin de calculer celles-ci !). Ceci est lié au fait que la famille des distributions de la variable aléatoire D^2 est *stochastiquement croissante* avec l'indicateur de non équirépartition (sur les

paramètres) $\delta(p_1, \dots, p_k) = \sum_{i=1}^k \left(p_i - \frac{1}{k} \right)^2$. Explicitons cette notion de croissance

stochastique : considérons deux lois de probabilités sur $\{1, \dots, k\}$, notées $P = (p_1, \dots, p_k)$ et

$P' = (p'_1, \dots, p'_k)$ telles que $\delta(P) \geq \delta(P')$ et notons, comme il est classique, P^n et P'^n les lois correspondantes des échantillons de taille n : alors il est vrai que, pour tout $t \geq 0$,

$P^n(D^2 > t) \geq P'^n(D^2 > t)$; en d'autres termes, plus $\delta(P)$ est grand, plus la variable aléatoire D^2 a tendance à prendre des valeurs élevées ; c'est bien ce qu'il nous fallait pour valider notre intuition qui fait porter sur d^2 la fabrication du test !

$k - 1$ degrés de liberté (c'est-à-dire la loi de la somme des carrés de $k - 1$ variables aléatoires indépendantes toutes de même loi gaussienne centrée réduite). C'est pourquoi cette statistique est communément appelée dans les manuels de Statistique **distance du χ^2** , quoique ce ne soit pas une distance au sens mathématique propre du terme (en particulier elle n'est pas symétrique et c'est pourquoi j'ai préféré dans ce texte parler d'éloignement). Dans le cas particulier de l'équirépartition (où tous les

p_i sont égaux à $\frac{1}{k}$), et dans ce cas là seulement, on trouve une statistique

proportionnelle au carré de la distance euclidienne : $t = nk d^2$ (ce coefficient multiplicatif nk est celui qui nous est proposé dans l'exercice du *meunier*).

Ce résultat de convergence, qui implique, bien sûr, des possibilités d'approximation pour n assez grand, a eu un impact historique considérable pour permettre la mise en œuvre des tests d'adéquation, à une époque où les moyens techniques ne permettaient pas d'approcher des lois par des simulations intenses, mais où en revanche on était capable d'approcher très convenablement les fonctions de répartition des lois du χ^2 par des méthodes analytiques.

Mais, dans l'esprit de nos programmes actuels, c'est par la simulation d'un grand nombre (soit N) d'échantillons tirés selon la loi équirépartie, tous de même taille n que celui observé, et pour chacun desquels on calcule la valeur prise par la variable aléatoire D^2 , que l'on approxime la loi de celle-ci ; il n'est pas, dans ce programme, envisagé de faire référence à l'approximation fondée sur le comportement asymptotique en n (et donc de recourir aux tables des lois du χ^2 , imprimées ou recalculées en ordinateur).

Rien ne justifie donc ici d'affecter d^2 de quelque multiplicateur que ce soit et j'aurais tendance à déconseiller de le faire. En effet une telle « rectification », **non motivable auprès des élèves**, portant sur la distance d^2 qui, elle, peut être introduite « avec naturel », ne peut apparaître qu'arbitraire, aux deux sens de ce terme : sens scientifique (qui signifie « non fixé au préalable ») et sens politique (qui renvoie au « fait du prince », le prince étant ici le professeur ou l'auteur d'un manuel ou d'un énoncé). **Or tout ce qui est arbitraire nuit à la crédibilité et au caractère proprement scientifique de ce qui est enseigné.**

5. Comment étudier la loi de la statistique de test sous l'hypothèse d'équirépartition

5.1. Obtention approchée de cette loi

Rappelons que les instructions du programme stipulent : *L'élève devra être capable ... de se reporter aux résultats de simulation qu'on lui fournira.* Il n'est donc pas question, contrairement à ce qui se passe en classe de seconde, qu'on lui fasse effectuer lui-même ces simulations, et pour cause : la taille de l'échantillon (n) comme le nombre de répétitions nécessaire pour avoir une approximation convenable (N) sont bien trop élevés (ou alors il faudrait y consacrer un temps appréciable de Travaux Pratiques sur ordinateur pour écrire et faire tourner le programme informatique correspondant). Les valeurs de N annoncées dans nos

problèmes de référence sont de 1 000 (*pisciculteur*), 2 000 (*banquier*) et 10 000 (*meunier*) ; mais nous (c'est-à-dire moi, vous mes lecteurs, les professeurs, les élèves) sommes bien obligés de croire sur parole les auteurs des problèmes ; s'ils n'avaient pas effectué, en se plaçant en situation d'équirépartition⁽¹⁸⁾, les N simulations annoncées mais fabriqué leurs résultats en se fondant sur l'approximation asymptotique par la loi du χ^2 , « nous n'y verrions que du feu ».

Cette petite remarque sur la facilité du mensonge en l'occurrence m'amène à une réflexion sur les avantages comparés des deux méthodes d'approximation de la distribution de D^2 , dont aucune n'est plus « noble » que l'autre. Les professeurs de Mathématiques ont souvent une tendresse particulière pour l'approximation asymptotique du χ^2 , car il y a un théorème derrière ; et il est vrai qu'il s'agit de bien belles mathématiques, dans le domaine de la théorie des probabilités consacré aux lois normales multidimensionnelles. Mais ce comportement asymptotique est relatif à n , taille de l'échantillon, dont le statisticien, souvent, n'a pas la maîtrise et, selon les exigences de précision usuelles, l'usage de cette approximation est à proscrire si les k valeurs $np_i(1-p_i)$ ne sont pas toutes plus grandes que 5, ce qui, quand on teste

l'équirépartition, signifie que l'on doit avoir $n > 5 \frac{k^2}{k-1}$. En revanche, l'approximation par simulations⁽¹⁹⁾ a l'avantage qu'elle prend en compte la vraie valeur de n et que nous y sommes maîtres du paramètre qui détermine la précision, à savoir le nombre N de répétitions⁽²⁰⁾.

Mais la méthode asymptotique garde pour l'enseignant l'avantage que les calculs s'y effectuent avec seulement d'une part une calculatrice rudimentaire et d'autre part une table des lois du χ^2 , ou bien une calculatrice (ou un ordinateur) sur laquelle on peut appeler la fonction de répartition et le calcul des quantiles de ces lois⁽²¹⁾.

Le professeur peut ainsi se fabriquer à faible coût des données utilisables autres que celles des énoncés préfabriqués ou bien utiliser ces données pour se poser, avec ses élèves, d'autres questions que celles de l'énoncé. Rien ne lui interdit, à mon avis, de dire aux élèves que les informations, sur la loi de D^2 , qu'il leur fournit auraient pu

(18) Insistons sur le fait que la simulation est un outil de calcul, mis en œuvre indépendamment des données à analyser ; le fait que, pour ce faire, on se place en situation i.i.d. avec équirépartition est donc une certitude (sauf à mettre en doute le générateur RANDOM utilisé !) ; je trouve donc inapproprié que, dans deux de nos trois énoncés de référence (*banquier* et *meunier*), les résultats de ces simulations soient introduits en parlant d'une « hypothèse » (ou « supposition ») d'équiprobabilité.

(19) Cette démarche s'inscrit en fait dans une évolution de la Statistique moderne (depuis 1980 environ) qui tend à insérer de plus en plus de simulations au sein même des procédures de calcul statistique, avec en particulier les techniques dites de *bootstrap*, bien plus élaborées que ce que nous présentons ici, bien sûr ; il y a là aussi de belles mathématiques derrière !

(20) Une réflexion sur le choix de N serait utile ici ; nous envisageons de la présenter dans un article ultérieur.

(21) On peut aussi aller les chercher sur Internet ; voir par exemple le site <http://perso.wanadoo.fr/alain.pichereau/chi2test.html> ; ce site comporte également une bonne introduction en français au test du χ^2 .

être obtenues par simulations, mais qu'il existe une autre méthode et que c'est celle qu'il a utilisée chez lui !

À titre d'exemple, c'est ainsi que j'ai procédé pour prolonger le problème du *banquier* ci-dessous en sous-section 6.2.

5.2. Présentation de cette loi

Si nous désirons fabriquer un test d'adéquation de données expérimentales à une loi équirépartie, au niveau de signification α , ce qui nous importe, c'est q_α , *quantile* d'ordre $1 - \alpha$ de la loi de la variable aléatoire D^2 quand l'équirépartition est satisfaite : si on note H la fonction de répartition de la loi (discrète) de D^2 en cas d'équirépartition, ce quantile sera défini ici⁽²²⁾ par :

$$q_\alpha = \inf \{x : H(x) \geq 1 - \alpha\};$$

si n est assez grand pour qu'on ait choisi de procéder à l'approximation asymptotique, q_α est approché par $\frac{1}{nk} G_{k-1}^{-1}(1 - \alpha)$, où G_{k-1} est la f.r. (strictement croissante et continue sur $[0, +\infty)$) de la loi du χ^2 à $k - 1$ degrés de liberté.

Alors, si les informations qui nous sont fournies sur cette loi permettent le calcul (exact ou du moins avec une très bonne approximation) de ce quantile q_α , c'est quand la valeur observée d^2 sera strictement supérieure à q_α que l'on osera procéder au rejet de l'équirépartition⁽²³⁾⁽²⁴⁾. Mais si l'on ne dispose que de la connaissance d'un intervalle auquel appartient q_α , autrement dit de valeurs approchées de q_α par défaut (soit q_α^-) et par excès (soit q_α^+)⁽²⁵⁾, c'est évidemment quand d^2 sera supérieur

(22) Pour une réflexion sur la définition des quantiles de lois discrètes, voir par exemple [14].

(23) Pour $k = 2$, complétons la comparaison, amorcée en note 16, avec le test introduit en 3.1, dans sa version utilisant l'approximation normale de la loi binomiale, pour laquelle le rejet

s'effectue si $\left|f - \frac{1}{2}\right| > c_{n;\alpha}$; si on prend pour q_α la valeur approchée issue de l'approximation

de la loi de D^2 par la loi du χ^2 à un degré de liberté (loi du carré d'une v.a. de loi normale centrée réduite), les frontières de ces deux présentations du même test sont liées par la relation : $nkq_\alpha = 2c_{n;\alpha}^2$.

(24) Que l'on utilise la loi discrète de D^2 ou son approximation à l'aide d'une loi du χ^2 , l'inégalité $d^2 > q_\alpha$ équivaut à $1 - H(d^2) < \alpha$ (pour la loi discrète, il faut, pour que cette équivalence soit correcte, se limiter aux valeurs effectivement atteintes par D^2 , autrement dit les points de discontinuité de la f.r., en escalier, H). $1 - H(d^2)$ est appelé la *p-valeur* associée aux données et de nombreux logiciels de statistique la fournissent. Sa connaissance nous renseigne donc sur ce que serait la conclusion du test pour chaque choix possible du niveau de signification (il s'agit bien sûr d'une notion générale de théorie des tests, transposable bien au delà du test d'équirépartition qui nous intéresse ici).

(25) Le lecteur peut vérifier, par exemple, que les 10 000 simulations faites dans le problème du *meunier* conduisent aux valeurs approchées suivantes de certains q_α^- et q_α^+ :

– encadrement du 9^e décile (quantile d'ordre 0,9) : $400 q_{0,1}^- = 6$ et $400 q_{0,1}^+ = 7$,

à q_α^+ que nous saurons procéder au rejet de l'équirépartition, puisque c'est ce que nous savons alors faire de « plus gros » comme région de rejet du type $[D^2 > q]$, tout en garantissant que sa probabilité (qui est celle de l'erreur de première espèce) est inférieure à α ⁽²⁶⁾.

Les énoncés de problèmes doivent donc présenter cette loi, en se référant aux simulations qui ont permis de l'approcher et en mettant en évidence des informations sur ce quantile (ou ces quantiles si on envisage plusieurs valeurs de α).

Examinons comment cette exigence est satisfaite dans nos différents énoncés de référence.

Les auteurs de *banquier* ont choisi de se reposer sur les **diagrammes en boîte** enseignés en cours de statistique descriptive dans les classes antérieures. J'exprimerai deux regrets quant à ce choix : d'une part le recours à la simulation effective est gommé, l'énoncé fournissant en fait directement les quantiles d'ordres 0,1 (premier décile), 0,25 (premier quartile), 0,5 (médiane), 0,75 (troisième décile) et 0,9 (neuvième décile) ; d'autre part ce matériel ne permet pas d'envisager d'autre valeur de α que 0,1, qui est vraiment bien peu sévère. En d'autres termes cet énoncé fournit une information « prémâchée » (où la référence au nombre de simulations, ici 2 000, n'est donc que rituelle) et très pauvre quant à son usage pour les tests.

Les auteurs de *pisciculteur* et *meunier* se conforment mieux à l'esprit du programme en nous renseignant sur la répartition des valeurs de D^2 lors des N simulations de l'échantillon de taille n tiré selon la loi équirépartie ; évidemment cette répartition ne peut être donnée de manière utilisable qu'après regroupement des observations en un nombre raisonnable de classes qui, dans les deux énoncés, sont des intervalles de même longueur (sauf bien sûr le plus élevé qui est une demi-droite supérieure). Cette information est donnée dans un cas (*pisciculteur*) sous forme d'histogramme et dans l'autre cas (*meunier*) sous forme de fonction de répartition inverse (nombre d'observations supérieures à chacune des frontières de classe). Il reste donc ici, pour chaque α proposé, un travail à faire au candidat, travail qui permet d'apprécier sa compréhension de l'approche par simulation du quantile

– encadrement du 95^e centile (quantile d'ordre 0,95) : $400 q_{0,05}^- = 7$ et $400 q_{0,05}^+ = 8$.

Notons que ces valeurs sont compatibles avec les quantiles correspondants de la loi du χ^2 à 3 degrés de liberté, qui valent respectivement 6,25 et 7,81.

(26) Dans [2], il est proposé, à titre de « compromis pédagogique », de prendre $q_{0,05}^+ = \frac{2}{n}$, en

se basant sur une circonstance heureuse qui veut que le quantile d'ordre 0,95 de la loi du χ^2 à $k - 1$ degrés de liberté soit toujours inférieur à $2k$. Cette approximation est assez bonne pour k faible ; en particulier, pour $k = 2$ (donc $k - 1 = 1$, et on a alors affaire au carré d'une variable aléatoire normale centrée réduite). Ceci revient à l'approximation classique du quantile d'ordre 0,975 de la loi normale centrée réduite, qui vaut 1,96, par 2 ; on trouve dans [2] les valeurs de ce quantile pour k variant entre 2 et 10 (par exemple pour $k = 5$, on obtient 9,49, encore assez proche de 10). Mais cette approximation devient de plus en plus mauvaise quand k augmente (par exemple pour $k = 30$, on obtient 42,6, bien inférieur à 60) ; donc on fabrique ainsi un test qui est certes de niveau de signification 0,05, mais dont la puissance se détériore quand k augmente.

d'ordre $1 - \alpha$: il doit « localiser », à l'aide des extrémités d'intervalles, la valeur qui est dépassée par une proportion α des N simulations effectuées. Mais les choix faits, dans ces énoncés, pour guider cette localisation, ne sont pas tout à fait conformes aux exigences de la théorie des tests : dans *pisciculteur* on demande une valeur approchée par défaut et dans *meunier* l'extrémité d'intervalle la plus proche. Or, on l'a bien vu dans le préambule de cette sous-section, c'est sa **valeur approchée par excès** qu'il aurait été utile de demander aux candidats.

6. Comment présenter la conclusion de l'étude

6.1. Rédaction de la « question finale » de l'énoncé

J'espère que, si le lecteur m'a suivi jusqu'à ce point de mon propos, il est bien convaincu qu'il n'y a, à des nuances de formulation près, qu'un seul type de conclusion possible, à deux modalités, affirmative ou négative, pour l'interrogation sur l'adéquation de données expérimentales à une loi équirépartie :

– soit **on peut**, avec un risque d'erreur inférieur à α , **rejeter l'adéquation des données fournies à une loi équirépartie**,

– soit **on ne peut pas**, avec un risque d'erreur inférieur à α , **rejeter l'adéquation des données fournies à une loi équirépartie**.

Quoique ceci aille un peu au delà de la lettre du programme, qui ne pousse pas à l'emploi du mot *hypothèse* (et en tout cas exclut *hypothèse nulle*) je préférerais dans ces phrases types, dire *rejeter* (ou *ne pas rejeter*), avec un risque ..., **l'hypothèse d'adéquation ...**

Cette exigence en matière de formulation conforme à la problématique classique des tests statistiques doit bien sûr être respectée dans le libellé des énoncés ; c'est pourquoi je crois nécessaire d'éviter toute rédaction de question qui, même sous forme interrogative, ferait porter le membre de phrase sacramentel (et indispensable) avec un risque d'erreur inférieur à ... sur l'acceptation de l'adéquation, et non sur son rejet. Que l'on m'excuse d'insister : si on « accepte », on ne sait jamais quelle est la probabilité de s'être trompé, puisque celle-ci serait à calculer relativement à chacune des lois non-équiréparties et qu'il y en a une infinité, certaines « très loin » mais d'autres aussi « très proches » de l'équirépartition⁽²⁷⁾. Tout ce que l'on sait alors, c'est que, au vu des données dont on dispose, rejeter l'équirépartition nous aurait fait encourir un risque d'erreur qui, vu notre connaissance sur la valeur de q_α (le quantile d'ordre $1 - \alpha$ de la loi de la statistique de test en cas d'équirépartition), aurait pu être de probabilité supérieure à α ; et j'insiste dans cette dernière phrase sur le conditionnel *aurait pu* qui est lié au fait que les manipulations faites sur le résultat (par exemple regroupement en classes des valeurs de d^2) ne nous donnent de la valeur de q_α qu'une connaissance partielle ; je renvoie ici à mes considérations, en sous-section 5.2, sur l'obligation d'approximer q_α par excès et non par défaut.

(27) Bien sûr il est possible de choisir quelques-unes de ces lois non équiréparties et de pousser à bout le calcul de la puissance du test en chacune d'entre elles (faisable à l'aide de simulations ou, asymptotiquement en n , à l'aide de lois limites dites « du χ^2 décentré ») ; s'agissant de notre problème du *pisciculteur*, on trouve de tels calculs dans l'article [18].

Je regrette donc la formulation à cet égard de chacun des trois énoncés que j'ai étudiés en référence, car aucun ne respecte la consigne que je viens de recommander. En particulier on lit dans *meunier* :

Si l'hypothèse d'équiprobabilité est vraie, peut-on affirmer, avec un risque d'erreur de 10%, que le mélange étudié à la question 1 est homogène ?

La précaution de langage *si l'hypothèse d'équiprobabilité est vraie* est ici inopportune. Certes, c'est *en se plaçant dans le cadre de l'hypothèse d'équiprobabilité* qu'ont été effectués les calculs, plus ou moins précis, de la valeur de q_α qui va nous servir de référence⁽²⁸⁾. Mais en ce qui concerne l'étude menée, sur le vu du d^2 observé, il est illogique de dire « *si l'hypothèse d'équiprobabilité est vraie* », alors que justement la question posée est de savoir si elle est vraie ou non ; et finalement on ne le saura jamais « vraiment » (le terme *affirmer* est donc trop fort) mais on se contentera d'annoncer si on rejette ou non cette hypothèse, et ce avec les précautions que l'on sait relatives au risque d'erreur accompagnant impérativement cette annonce.

Le lecteur m'aura peut-être trouvé bien pointilleux ; j'assume⁽²⁹⁾ et j'irai même un peu plus loin en l'invitant à réfléchir sur la nuance de sens entre les constructions suivantes :

1. *on peut* (resp. *on ne peut pas*), avec un risque d'erreur inférieur à α , *rejeter l'hypothèse ...* ;
2. *on peut* (resp. *on ne peut pas*) *rejeter*, avec un risque d'erreur inférieur à α , *l'hypothèse ...*

Je préfère la première qui marque mieux que le risque d'erreur est fonction de la méthode de test qui nous apprend quand il est licite ou non de rejeter l'hypothèse.

6.2. Comment aller, en classe, au delà des énoncés de problèmes

Tout au long de ce texte, je me suis plu à varier les formulations visant à bien ancrer le rôle central de l'alternative rejet ou non-rejet de l'hypothèse testée (j'ai parlé ainsi de *oser rejeter*, de *être autorisé à rejeter*, de *licite de rejeter*, ...). J'invite les enseignants à procéder ainsi lors de leurs commentaires oraux devant les élèves, car ces variations un peu littéraires peuvent renforcer leur pouvoir de conviction. Comme je le disais en sous-section 2.2, j'aimerais bien en particulier que l'on ne bannisse pas l'emploi des mots *significatif* et *significativement*, en raison de l'usage « populaire » qui en est fait ; on pourrait dire ainsi par exemple, en classe, que *la suite*

(28) Dans l'esprit du programme, chaque énoncé présente comme faites « maintenant » les simulations fournissant cette valeur. L'énoncé *pisciculteur* est même très explicite dans l'affirmation de cette contemporanéité : on y lit : *le pisciculteur simule le prélèvement au hasard...* Ce producteur a vraiment une multiple compétence (pas tellement, à la réflexion, car son nombre de répétitions, 1 000, est trop faible). En fait, dans la pratique, ces simulations peuvent très bien avoir été effectuées antérieurement (ou bien on peut avoir recours plutôt à l'approximation asymptotique par les lois du χ^2).

(29) Je ne veux quand même pas jeter l'anathème sur tous ceux, statisticiens, usagers de la Statistique ou enseignants, qui disent « acceptation » là où j'ai veillé à dire « non-rejet » (cela m'arrive !), mais on n'insistera jamais assez sur le fait qu'ici « accepter » ne peut pas vouloir dire « affirmer vrai » mais seulement « ne pas être en état de réfuter ».

des fréquences observées s'écarte (ou ne s'écarte pas) significativement, au niveau de signification α , de la suite de fréquences constante traduisant l'équirépartition.

D'autre part, puisque l'intérêt principal de ces études en terminale est, pour moi, de développer l'esprit critique du futur citoyen, ou bien de l'initier à une démarche scientifique fondamentale, il me semble très souhaitable de ne pas s'en tenir à des énoncés « secs » mais de réfléchir avec les élèves sur la signification concrète des contextes proposés, d'observer les données et d'en tirer la motivation de nouvelles interrogations.

À titre d'exemple, c'est ainsi que je vais procéder pour prolonger le problème du *banquier* dans deux directions :

- considérer $\alpha = 0,05$ (alors que l'énoncé prend $\alpha = 0,1$),
- s'intéresser à la seule considération du mardi et du samedi, comme on l'a justifié plus haut en sous-section 3.2.

Je vais avoir besoin de valeurs de quantiles qui ne me sont pas données par l'énoncé. Je vais les chercher ici à l'aide d'approximations par la loi du χ^2 ; ces approximations vont être excellentes car :

- avec les cinq jours ouvrables, $k = 5$ et donc $5 \frac{k^2}{k-1} = 5 \frac{16}{3} \approx 27$, auquel est bien supérieur $n = 250$,

- avec le mardi et le samedi, $k = 2$ et donc $5 \frac{k^2}{k-1} = 20$, auquel est bien supérieur $n = 97$ (somme de 37 pour le mardi et 60 pour le samedi) ; le fait que la valeur de n ne soit plus préfixée⁽³⁰⁾, mais dépende de l'expérimentation ne modifie en rien la technique de test telle que nous la pratiquons ici.

Si on prend en compte les cinq jours ouvrables, la valeur observée d^2 est $5,248 \times 10^{-3}$, d'où

$$nkd^2 = 1\,250d^2 = 6,56.$$

Les quantiles d'ordre 0,95 et 0,90 de la loi du χ^2 à 4 degrés de liberté valent respectivement 9,49 et 7,78⁽³¹⁾ ; 6,56 est inférieur à chacun d'eux et donc on ne peut ni avec un risque d'erreur inférieur à 0,05, ni même, moins sévèrement, en tolérant un risque d'erreur inférieur à 0,10, rejeter l'hypothèse d'équirépartition.

Avec uniquement le mardi et le samedi, la valeur observée d^2 est

$$2 \left(\frac{37}{97} - \frac{1}{2} \right)^2 = 2,81 \times 10^{-2}$$

(30) En fait, dans ce cas précis, on peut douter que la taille de l'échantillon complet ait été préfixée puisque, malgré sa valeur « bien ronde » (250), elle est présentée comme le nombre total de clients ayant effectué des retraits dans une semaine donnée.

(31) Pour ce dernier la valeur obtenue par simulations donnée dans l'énoncé est 7,5 (= $1\,250 \times 0,006$) ; on est assez loin de la valeur donnée par l'approximation asymptotique du χ^2 (7,78), qui, elle, est en fait excellente ; mais il nous est dit que la simulation a été faite avec 2 000 répétitions, ce qui est assez faible.

d'où

$$nkd^2 = 194d^2 = 5,45.$$

Les quantiles d'ordre 0,95 et 0,90 de la loi du χ^2 à un degré de liberté d'une variable (loi du carré d'une variable aléatoire normale centrée réduite) valent respectivement 3,84 et 2,71 ; 5,45 est supérieur à chacun d'eux et donc cette fois on peut, avec une probabilité d'erreur de 0,10, rejeter l'hypothèse d'équirépartition ; on le peut même, plus sévèrement, en admettant une probabilité d'erreur de 0,05. En revanche on ne pourrait plus la rejeter au risque, encore plus sévère, de 0,01 car le quantile d'ordre 0,99 de la loi du χ^2 à un degré de liberté vaut 6,63.

Si on traite cet exemple avec des élèves, il peut être intéressant de tâcher de leur faire comprendre pourquoi ces résultats n'ont rien de paradoxal : la différence entre les effectifs de clients du mardi et du samedi est significative, aux niveaux usuels 0,05 ou 0,10 ; en revanche si on considère toute la semaine, elle est en quelque sorte « noyée » parmi les valeurs intermédiaires des autres jours.

7. Conclusion

À propos d'un thème que je considère comme important et dont je souhaite faire connaître la logique interne, je me suis trouvé confronté à un dilemme didactique classique : comment bénéficier de la « liberté » du travail en classe pour faire s'approprier une notion et les techniques y afférentes, mais comment se plier aussi aux fortes contraintes de l'exercice d'examen sur cette même situation ?

Ce dilemme est d'autant plus marqué ici que :

- la vocation même de ce thème dans les programmes est de faire réfléchir les élèves sur différents aspects de situations concrètes et sur la signification, précise et limitée, de ce qu'on peut inférer à leur sujet ;
- le programme se limite à l'apprentissage d'une technique unique et encadre très strictement ce que l'on peut faire à son propos.

À la limite, on peut dire : **mathématiquement, il n'y a, aux valeurs numériques près, qu'un seul exercice d'examen faisable sur ce point du programme.** Or ce n'est pas dans le temps limité de l'épreuve qu'on peut inciter le candidat à faire preuve de réflexion créatrice sur les contextes où se justifie le recours au test d'adéquation à l'équirépartition.

Il faut donc, à mon sens, admettre que la présence possible d'exercices sur ce point au baccalauréat n'a d'autre but que de manifester son existence dans le programme et de s'assurer très modestement que le protocole associé est maîtrisé par le candidat (et c'est pourquoi le mode d'expression associé à ce protocole doit être si scrupuleusement respecté).

Je ne serais donc pas trop gêné, pour ma part, que l'exercice d'examen soit présenté comme « décontextualisé ». C'est ainsi que j'ai choisi de rédiger un modèle en Annexe C, tout en reconnaissant que cette attitude est un peu provocatrice, car bien sûr des exercices d'examen faisant intervenir des contextes soigneusement choisis peuvent aussi guider les enseignants dans leur choix d'exemples.

Mais le vrai travail utile et intéressant se fait en classe, en réfléchissant avec

les élèves, avant de se lancer dans les calculs, sur des contextes adaptés à notre problématique du test d'équirépartition.

C'est surtout là, à mon sens, qu'il y a matière à réflexion pédagogique neuve et qu'il est souhaitable de faire appel à des praticiens de la Statistique pour suggérer des cadres d'emploi pédagogiquement transposables (voir [11], ou ici en Annexe B).

8. Annexe A : nos trois énoncés de référence

8.1. Le problème du meunier

Cet exercice est extrait du fascicule *Accompagnement des programmes, Mathématiques, Baccalauréat séries S et ES, Exemples d'exercices pour des sujets, publié en vue de la session 2004 par le Ministère de l'Éducation Nationale et diffusé par le Centre National de Documentation Pédagogique (page 35)⁽³²⁾.*

Un meunier a besoin, pour sa farine, d'un mélange de quatre variétés différentes de grains de blé, d'égales quantités chacune, et notées 1, 2, 3, 4.

1. Il veut savoir si, dans son silo, les différentes variétés sont bien mélangées. Pour cela, il prélève, à la sortie du silo, un échantillon de 100 grains de blé rendus radioactifs par des marqueurs différents selon les variétés. Il obtient les résultats suivants :

Variété	1	2	3	4
Nombre de grains radioactifs	18	27	35	20

Le meunier veut savoir si ces données sont vraisemblables lorsqu'on fait l'hypothèse d'un mélange homogène des quatre variétés, ce qui correspond à un quart des marqueurs pour chaque variété.

On appelle f_i la fréquence dans l'échantillon de la variété i et on pose

$$d^2 = 400 \sum_{i=1}^n \left(f_i - \frac{1}{4} \right)^2.$$

Calculer la valeur de d^2 .

2. On suppose l'équiprobabilité de la présence d'un grain de blé de chaque variété et on simule 10 000 séries de 100 tirages de grains de blé.

Pour chaque série de 100 tirages, on calcule d^2 . La tableau suivant donne le nombre de séries pour lesquelles la valeur de d^2 est strictement supérieure à l'entier j :

j	3	4	5	6	7	8	9	10
Nombre des séries pour lesquelles $d^2 > j$	3 915	2 618	1 715	1114	728	467	306	180

(32) Disponible sur Internet, par exemple sur le site : www.ac-poitiers.fr (information valable en juillet 2005) ; suivre : Espace pédagogique → Disciplines et filières → Mathématiques → LGT → Pour les enseignants → Baccalauréat 2004, séries S et ES. Ces batteries d'exemples d'exercices sont également disponibles sur www.eduscol.education.fr, regroupées séparément pour S et pour ES, mais, bizarrement, le problème qui nous intéresse ici, et qui est le seul commun aux deux filières, en est absent !

Lire la valeur du 9^e décile, arrondie à l'entier le plus proche, puis celle du 95^e centile.

3. Si l'hypothèse d'équiprobabilité est vraie :

- a) Peut-on affirmer avec un risque d'erreur de 10% que le mélange étudié à la question 1 est homogène ?
- b) Même question avec un risque d'erreur de 5%.
- c) Que peut-on dire si quelqu'un demande un risque d'erreur de 0% ?

8.2. Le problème du pisciculteur

Cet exercice a été posé au centre d'examen de Pondichéry, en 2003, au Baccalauréat ES⁽³³⁾.

Un pisciculteur possède un bassin qui contient 3 variétés de truites : communes, saumonées et arc-en-ciel. Il voudrait savoir s'il peut considérer que son bassin contient autant de truites de chaque variété. Pour cela il effectue, au hasard, 400 prélèvements d'une truite avec remise et obtient les résultats suivants :

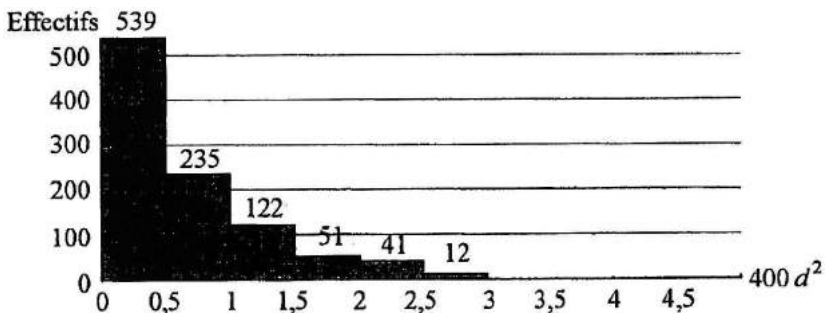
Variétés	Commune	Saumonée	Arc-en-ciel
Effectifs	146	118	136

1.a. Calculer les fréquences de prélèvement f_c d'une truite commune, f_s d'une truite saumonée et f_a d'une truite arc-en-ciel. On donnera les valeurs décimales exactes.

b. On pose $d^2 = \left(f_c - \frac{1}{3}\right)^2 + \left(f_s - \frac{1}{3}\right)^2 + \left(f_a - \frac{1}{3}\right)^2$.

Calculer $400d^2$ arrondi à 10^{-2} ; on note $400d_{obs}^2$ cette valeur.

2. À l'aide d'un ordinateur, le pisciculteur simule le prélèvement au hasard de 400 truites suivant la loi équirépartie. Il répète 1 000 fois cette opération et calcule à chaque fois la valeur de $400d^2$. Le diagramme à bandes ci-dessous représente la série des 1 000 valeurs de $400d^2$, obtenues par simulation.



Déterminer une valeur approchée à 0,5 près par défaut, du neuvième décile D9 de cette série.

(33) Disponible sur Internet, par exemple sur le site : www.ac-aix-marseille.fr (information valable en juillet 2005) ; suivre : La pédagogie → Les ressources disciplinaires et transversales → Sciences → Mathématiques en Collège et Lycée → Examens et concours → Bac 2003.

3. En argumentant soigneusement la réponse, dire si on peut affirmer avec un risque d'erreur inférieur à 10% que « le bassin contient autant de truites de chaque variété ».

4. On considère désormais que le bassin contient autant de truites de chaque variété. Quand un client se présente, il prélève au hasard une truite du bassin. Trois clients prélèvent chacun une truite. Le grand nombre de truites du bassin permet d'assimiler ces prélèvements à des tirages successifs avec remise. Calculer la probabilité qu'un seul des trois clients prélève une truite commune.

8.3. Le problème du banquier

Cet exercice a été posé en France métropolitaine, en 2003, au Baccalauréat ES⁽³⁴⁾.

Les guichets d'une agence bancaire d'une petite ville sont ouverts au public cinq jours par semaine : les mardi, mercredi, jeudi, vendredi et samedi.

Le tableau ci-dessous donne la répartition des 250 retraits d'argent liquide effectués aux guichets une certaine semaine.

jour de la semaine	mardi	mercredi	jeudi	vendredi	samedi
rang i du jour	1	2	3	4	5
nombre de retraits	37	55	45	53	60

On veut tester l'hypothèse « le nombre de retraits est indépendant du jour de la semaine ».

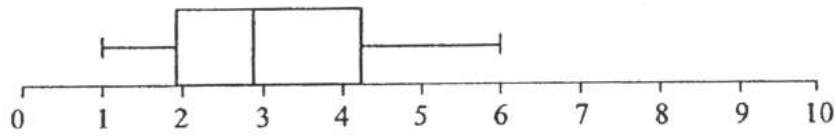
On suppose donc que le nombre de retraits journaliers est égal à 1/5 du nombre de retraits de la semaine.

On pose $d_{obs}^2 = \sum_{i=1}^5 \left(f_i - \frac{1}{5} \right)^2$ où f_i est la fréquence des retraits du i -ème jour.

1. Calculer les fréquences de retraits pour chacun des cinq jours de la semaine.
2. Calculer alors la valeur de 1 000 d_{obs}^2 (la multiplication par 1 000 permet d'obtenir un résultat plus lisible).
3. En supposant qu'il y a équiprobabilité des retraits journaliers, on a simulé 2 000 séries de 250 retraits hebdomadaires.

Pour chaque série, on a calculé la valeur du 1 000 d_{obs}^2 correspondant. On a obtenu ainsi 2 000 valeurs de 1 000 d_{obs}^2 .

Ces valeurs ont permis de construire le diagramme en boîtes ci-dessous, où les extrémités des « pattes » correspondent respectivement au premier décile et au neuvième décile.



(34) Disponible sur Internet : voir note précédente.

Lire sur le diagramme une valeur approchée du neuvième décile.

4. En argumentant soigneusement la réponse, dire si pour la série observée au début, on peut affirmer, avec un risque d'erreur inférieur à 10%, que « le nombre de retraits est indépendant du jour de la semaine » ?

9. Annexe B : quelques propositions de contextes utilisables en classe

9.1. À propos de la fabrication de « nombres au hasard »

Il serait évidemment maladroit, dans l'esprit de nos programmes, de vouloir tester la qualité du générateur de nombres au hasard (l'appel au RANDOM) fourni par un moyen de calcul à la disposition des élèves, calculatrice ou ordinateur, et ce pour deux raisons, l'une conjoncturelle et l'autre fondamentale.

La raison conjoncturelle est que, puisque le programme de Terminale se repose, dans la fabrication même du test, sur l'usage de tels moyens, c'est qu'il considère qu'on peut se fier à eux ; si cette confiance risquait d'être non fondée, on ne pourrait évidemment pas s'appuyer dessus pour les tester eux-mêmes (ce serait comme demander à un menteur s'il ment) !

La raison fondamentale est que toute notre théorie repose sur des modèles dans lesquels la qualité d'échantillon (c'est-à-dire l'indépendance et l'identique distribution) des données n'est pas en cause, et pour lesquels c'est uniquement la nature de cette distribution commune qui pose problème ; or on sait bien que, dans les générateurs de nombres au hasard, c'est l'indépendance bien plus que l'uniformité qui est difficile à réaliser et peut donc être mise en question (pour une réflexion critique sur ces générateurs, voir par exemple [20]).

En revanche, il peut être amusant d'imaginer avec les élèves des simulateurs analogiques (par opposition à numériques) et de les tester. Ainsi on peut leur proposer de se poster dans la rue et de noter le chiffre des unités dans la première partie du numéro minéralogique de voitures immatriculées en France qu'ils voient passer devant eux, dans un sens donné, pendant 15 minutes (par exemple, pour 574 CTR 75, on notera 4), puis de rassembler des résultats d'élèves postés en des lieux fort éloignés pour constituer un gros échantillon. L'indépendance de ces observations peut être considérée comme assurée (c'est pourquoi nous avons précisé « dans un sens donné » pour éviter autant que possible les aller-retours d'une même voiture et avons posté nos observateurs en des lieux distants, pour éviter les passages systématiques de mêmes véhicules). Et on pensera que « ce serait bien le diable » qu'il n'y ait pas équirépartition et qu'il faudrait donc, pour rejeter cette conviction, que l'analyse de l'échantillon nous surprenne nettement ; nous voici donc typiquement dans le cadre d'emploi de la théorie classique des tests. Et pourtant, il peut arriver qu'on détecte une majorité nette de chiffres pairs (par exemple) à Paris un jour de très forte pollution, où, parmi les voitures sans pastille verte, seules celles à numéro pair sont autorisées à circuler !

9.2. Un problème de contrôle industriel : comparaison de machines

Dans une usine fonctionnent k machines produisant un même type d'objet, qui peut être soit bon, soit défectueux ; on s'intéresse au fonctionnement de l'usine un jour donné, au cours duquel k machines ont eu la même production, à savoir m objets chacune. On considère que, pour une machine donnée i (où $1 \leq i \leq k$), les qualités des m objets produits sont indépendantes et de même loi, caractérisée par la probabilité q_i pour chacun de ces objets d'être bon. Sitôt produit, chaque objet est vérifié et les objets défectueux sont immédiatement mis au rebut. Tous les objets conservés sont donc bons et ils sont tous mélangés avant d'être expédiés en fin de journée vers un entrepôt ; mais l'observation d'un objet permet toujours d'identifier quelle machine l'a produit.

Un contrôle statistique est effectué sur le stock d'objets produits ce même jour, par prélèvement d'un échantillon de taille n à l'entrepôt. n étant petit devant la taille du stock, on admettra que ces n observations sont indépendantes. Il n'est évidemment plus possible à ce stade d'estimer les probabilités q_i (puisque les objets défectueux ont disparu). En revanche on va pouvoir, en examinant l'origine des objets (tous bons) dans cet échantillon, s'intéresser à tester l'hypothèse d'égalité des machines, aux termes de laquelle ces probabilités q_i sont toutes égales. Rejeter cette hypothèse a des conséquences lourdes car cela induit la suspicion que l'une au moins des machines se serait détériorée, d'où le besoin éventuel d'une nouvelle étude statistique le lendemain (dans l'usine et non dans l'entrepôt), des vérifications techniques onéreuses, l'arrêt possible de la production, ... ; on se trouve donc bien face à une problématique justifiant le recours à la théorie classique des tests.

Notons x_i le nombre d'objets en provenance de la machine i dans l'échantillon prélevé à l'entrepôt. Il est intuitif que des qualités analogues des machines se traduiront par le fait que les rapports $\frac{x_i}{n}$ seront proches les uns des autres, c'est-à-

dire tous proches de $\frac{1}{k}$; inversement, une machine plus mauvaise (resp. meilleure) que les autres devrait être sous-représentée (resp. sur-représentée) dans cet échantillon tiré parmi de bonnes pièces.

Cette intuition peut être justifiée par un calcul de probabilités conditionnelles élémentaire (à la portée des élèves de terminale S ou ES), relatif au tirage au hasard (c'est-à-dire avec équiprobabilité) d'une pièce dans l'ensemble de toutes les km pièces produites à l'usine dans la journée, de sorte que la probabilité que cette pièce

soit issue de la machine i est $\frac{1}{k}$ quel que soit i . Notons alors p_i la probabilité qu'une pièce soit produite par la machine i conditionnellement au fait qu'elle est bonne ; il

vient : $p_i = \frac{q_i}{q_1 + \dots + q_k}$. Il en résulte l'équivalence entre les deux propositions suivantes :

A. $q_1 = q_2 = \dots = q_k$ (peu nous importe ici cette valeur commune) ;

B. $\forall i \ p_i = \frac{1}{k}$.

L'équivalence entre les propositions A et B exprime bien que l'hypothèse, que l'on désire tester, de même qualité des machines, se traduit par l'hypothèse d'équirépartition relative à l'appartenance aux différentes machines dans le stock disponible à l'entrepôt. On peut donc, utilisant la statistique de test

$$d^2 = \sum_{i=1}^n \left(\frac{x_i}{n} - \frac{1}{k} \right)^2, \text{ lui appliquer la technique connue}^{(35)(36)}.$$

9.3. Un test sur la médiane

La manière dont je vais présenter cet exemple est inspirée d'une anecdote célèbre sur Gauss. Mais il peut être adapté à de nombreux autres contextes.

Le poids officiel d'un certain type de baguette de pain est de 250 grammes. Mon boulanger m'affirme qu'il ne peut bien sûr pas garantir ce poids exact sur chaque baguette vendue, mais qu'il y a dans sa production « autant de baguettes plus lourdes que de baguettes moins lourdes ». Ayant l'impression qu'en fait il triche un peu et que ses baguettes sont plus souvent trop légères que trop lourdes, je décide de peser systématiquement (au gramme près) ma baguette chaque jour et de pratiquer un test statistique. Mais je ne peux pas me permettre d'accuser mon boulanger de tricherie sans en être bien convaincu et je me trouve donc typiquement dans la situation de dissymétrie de la gravité des deux types d'erreurs qui caractérise la théorie classique des tests.

Formellement, l'affirmation du boulanger peut se traduire par le fait que la médiane de la distribution des poids des pains produits se situe à 250 grammes. Comme cette distribution est discrète (on ne donne la mesure qu'en nombre entier de grammes), nous devons préciser cette affirmation par le fait que, si on note a la probabilité que le poids d'une baguette soit mesuré exactement à la valeur de 250 grammes, les probabilités qu'on enregistre des poids strictement plus élevés ou

(35) Dans le cas particulier où $k = 2$, d^2 est évidemment proportionnel à $\left(\frac{x_1}{n} - \frac{x_2}{n} \right)^2$ et on

retrouve le test classique de *comparaison de fréquences* fondé sur $\left| \frac{x_1}{n} - \frac{x_2}{n} \right|$. Une version « à

signe », portant sur la statistique, à valeurs réelles de signe quelconque, $\frac{x_1}{n} - \frac{x_2}{n}$, peut

permettre de tester une hypothèse dite « unilatérale » du type $q_1 \leq q_2$, c'est-à-dire mettre à l'épreuve une assertion selon laquelle la machine 2 serait meilleure que la machine 1 (l'hypothèse $q_1 = q_2$ considérée plus haut étant dite « bilatérale »).

(36) L'égalité des productions quotidiennes des k machines a joué un rôle essentiel ; si elles étaient différentes, l'hypothèse d'égalité des machines se traduirait par le fait que, pour tout i , p_i serait égal à la fraction de la production totale assurée par cette machine i . Cela ne gênerait pas le statisticien, mais est en dehors du programme des terminales.

strictement moins élevés sont toutes deux égales à $\frac{1-a}{2}$; autrement dit, conditionnellement au fait de ne pas peser le poids légal, les probabilités de peser strictement plus ou strictement moins sont toutes deux égales à $\frac{1}{2}$. Voilà où gît l'équirépartition que nous devons tester et, comme dans l'exemple précédent, il va s'agir d'un test conditionnel (ici conditionnellement au nombre d'observations différentes de 250).

Si on note x_+ (resp. x_-) le nombre de baguettes qui, dans l'échantillon que j'ai constitué, sont plus lourdes (resp. plus légères) que 250 grammes, la statistique de test est ici

$$d^2 = \left(\frac{x_+}{x_+ + x_-} - \frac{1}{2} \right)^2 + \left(\frac{x_-}{x_+ + x_-} - \frac{1}{2} \right)^2 = 2 \left(\frac{x_+}{x_+ + x_-} - \frac{1}{2} \right)^2.$$

Prolongement possible. Un bon artisan doit non seulement assurer la valeur centrale de sa production, mais également en limiter la dispersion. Imaginons que notre boulanger déclare, en sus du respect de la médiane à 250 déjà évoqué : « Je garantis que :

- parmi mes pains de poids strictement inférieur à 250 grammes, la moitié ont un poids supérieur à 247,5 grammes ;
- parmi mes pains de poids strictement supérieur à 250 grammes, la moitié ont un poids inférieur à 252,5 grammes. »

Si on veut étudier s'il y a lieu de mettre en doute cette déclaration, on est conduit à un test d'équirépartition avec observations dans un ensemble à 4 éléments.

10. Annexe C : une proposition d'énoncé « décontextualisé »

Bien sûr, pour un « vrai » énoncé, les notations n, k, N, α et $1 - \alpha$ de l'exemple ci-dessous seront remplacées par des valeurs numériques.

On a recueilli un échantillon de taille n d'observations à valeurs dans un espace fini à k éléments ; ces k modalités sont codées par les entiers de 1 à k .

Les résultats sont donnés par le tableau suivant :

TABLEAU

(je propose de faire un tableau à $k + 1$ colonnes, la dernière colonne étant consacrée au total, pour retrouver la taille de l'échantillon).

Notre but est de tester l'hypothèse selon laquelle la loi des modalités est équirépartie.

1. Calculez les fréquences des différentes modalités observées dans l'échantillon ; on les notera f_1, \dots, f_k .

2. Calculez, pour cet échantillon, la valeur de d^2 , carré de la distance euclidienne entre la suite des fréquences observées f_i (où $1 \leq i \leq k$) et la suite constante de

longueur k dont tous les termes valent $\frac{1}{k}$.

3. Il a été effectué, en situation d'équirépartition, la simulation de N (je conseille de prendre au moins 5 000) échantillons de taille n à k modalités et il a été calculé à chaque fois la valeur de d^2 .

Le tableau suivant donne, pour certaines valeurs y , le nombre d'échantillons simulés pour lesquels d^2 est strictement supérieur à y .

TABLEAU

Diverses variantes sont possibles pour la fourniture de l'information relative aux résultats des simulations ; on peut les donner sous forme graphique ou sous forme de tableaux ; l'important me semble être que l'on y parle bien en termes de nombre de simulations et que l'aspect « regroupement en classes » y soit explicite.

Sur la base de cette information, proposez une valeur approchée par excès du quantile d'ordre $1 - \alpha$ de la loi de d^2 en situation d'équirépartition.

Diverses terminologies sont possibles ici ; on peut ainsi selon les besoins parler du 9^e décile, du 95^e centile, du 99^e centile, ... de cette loi⁽³⁷⁾.

Une variante qui laisse un peu plus d'autonomie à l'élève peut être : sur la base de ces informations, que pouvez-vous dire de la localisation du quantile d'ordre $1 - \alpha$ de la loi de d^2 en situation d'équirépartition ?

4. Reprenant la valeur de d^2 calculée en 2, dites, en justifiant votre réponse, si, au vu de l'échantillon recueilli, on peut, avec une probabilité d'erreur inférieure à α , rejeter l'hypothèse d'équirépartition des k modalités possibles⁽³⁸⁾.

Dans le libellé de cette question, il faut éviter l'emploi des vocables « accepter » et « affirmer ».

Éventuellement on peut faire reprendre les questions 3 et 4 pour d'autres valeurs de α .

Références

- [1] BARTHÉLÉMY M.-J. ET AL. (2005) *Lois continues, test d'adéquation : une approche pour non spécialiste*, Publications de l'IREM de Besançon, Presses Universitaires de Franche-Comté.
- [2] BONNEVAL L.-M. (2002) Test d'équirépartition : qui a dit khi-deux ? *Bulletin de l'APMEP*, **441**, 512-521.
- [3] BONNEVAL L.-M. ET HENRY M. (2005) Test d'équirépartition à une loi de probabilité, pratique des tests du Khi-deux, *Statistique au Lycée, 1. Les outils de la description statistique*, ouvrage réalisé par la Commission Inter-IREM Statistique et Probabilités, coordonné par CHAPUT B. ET HENRY M., Brochure APMEP n° 156, 261-273.

(37) Il n'y pas lieu, me semble-t-il, d'avoir dans la rédaction de cet énoncé l'attitude puriste consistant à distinguer entre la variable aléatoire D^2 et son observation présente d^2 , attitude que j'ai par endroits adoptée à l'intention des enseignants dans le corps de cet article.

(38) Il ne s'agit pas ici de « troubler » le candidat : je suggérerais de faire en sorte que d^2 soit nettement en dehors (d'un côté ou de l'autre) de l'intervalle dans lequel on sait situer le quantile d'ordre $1 - \alpha$, afin que la décision en faveur du rejet ou du non-rejet soit bien tranchée.

- [4] BONNEVAL L.-M. (2005) Test d'équirépartition, quel risque d'erreur ? *Bulletin de l'APMEP*, **456**, 31-45.
- [5] BONNEVAL L.-M. (2005) Test d'équirépartition, quel risque d'erreur ? <http://www.aromath.net/moodle/>
- [6] DACUNHA-CASTELLE D. ET DUFLO M. (1982) *Probabilités et statistique. 1. Problèmes à temps fixe*, Masson, Paris.
- [7] DACUNHA-CASTELLE D. ET DUFLO M. (1982) *Exercices de probabilités et statistique, 1. Problèmes à temps fixe*, Masson, Paris.
- [8] DROESBEKE J.-J. ET TASSI P. (1990) *Histoire de la statistique*, P.U.F., Que sais-je ?, Paris.
- [9] DUTARTE P. ET PIEDNOIR J.-L. (2001) *Enseigner la statistique au lycée*, Brochures de la Commission Inter-IREM Lycées Technologiques, **112**.
- [10] DUTARTE P. (2005) Pour une éducation à l'inférence statistique au lycée, *Repères-IREM*, **60**, 5-19.
- [11] DUTARTE P. (2005) Adéquation statistique à un modèle ; exemples dans le domaine de l'environnement, <http://www.aromath.net/moodle/>
- [12] FOURDRINIER D. (2002) *Statistique inférentielle*, Dunod, Paris.
- [13] GERBAL J.-P. ET AL. (2005) Adéquation à une loi équirépartie (à partir des résultats d'une simulation - terminale) *Réflexions sur l'enseignement des statistiques au lycée*, IREM, Orléans, 24-29.
- [14] GIRARD J.-C. (2005) Quartiles, déciles et tutti quantiles, *Statistique au Lycée, 1. Les outils de la description statistique*, ouvrage réalisé par la Commission Inter-IREM Statistique et Probabilités, coordonné par CHAPUT B. et HENRY M., Brochure APMEP n° 156, 39-52.
- [15] HENRY M. (2003) Simulation d'un sondage, fourchettes d'échantillonnage et intervalles de confiance, *Bulletin de l'APMEP*, **444**, 88-96.
- [16] HENRY M. (2005) Introduction aux tests d'hypothèses, exemples, *Statistique au Lycée, 1. Les outils de la description statistique*, ouvrage réalisé par la Commission Inter-IREM Statistique et Probabilités, coordonné par CHAPUT B. ET HENRY M., Brochure APMEP n° 156, 247-260.
- [17] KENTZEL J.-F. (2005) Quelques précisions sur le test de l'adéquation de données à une loi équirépartie, <http://www.aromath.net/moodle/>
- [18] KOCH B. ET AL. (2005) Les truites de Pondichéry, un problème d'adéquation au baccalauréat, *Repères-IREM*, **59**, 55-60.
- [19] LEJEUNE M. (2004) *Statistique : la théorie et ses applications*, Springer, Paris.
- [20] PARZYSZ B. (2005) Quelques questions à propos des tables et des générateurs aléatoires, *Statistique au Lycée, 1. Les outils de la description statistique*, ouvrage réalisé par la Commission Inter-IREM Statistique et Probabilités, coordonné par CHAPUT B. ET HENRY M., Brochure APMEP n° 156, 181-199.
- [21] ROBERT C. (2003) *Contes et décomptes de la statistique, une initiation par l'exemple*, Vuibert, Paris.
- [22] SAPORTA G. (1990) *Probabilités, Analyse des données et Statistique*, Technip, Paris.