Exemples de prises de décision

Voici regroupés deux sujets portant sur l'échantillonnage. Dans les deux cas, mais dans des circonstances très différentes, on demande aux élèves de se prononcer sur la véracité d'une affirmation, c'est-à-dire de « prendre une décision ».
Un autre point commun entre les exercices proposés par le Jury dans ces deux sujets est leur habillage surréaliste …

ESD 2013 –02 : Loi binomiale

1. Le sujet

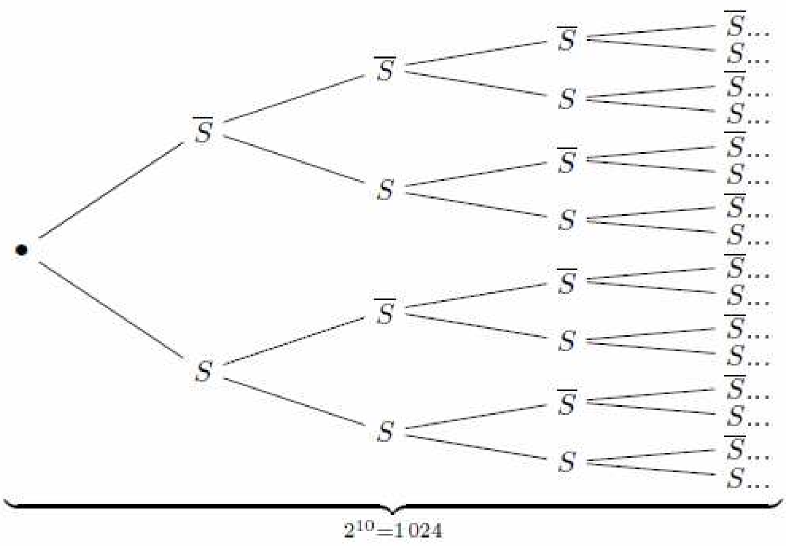
A. L'exercice proposé au candidat

Un homme se présente dans un village gaulois et se déclare devin. Les habitants sceptiques lui proposent un test au cours duquel il devra deviner les résultats de dix lancers d'une pièce équilibrée. Il donne huit fois la bonne réponse.
1. On suppose que les réponses sont données au hasard. Calculer dans ce cas la probabilité de trouver huit bonnes réponses.
2. Les habitants du village devraient-ils croire aux pouvoirs du devin ?
Source : d'après manuel Hyperbole Première S (Nathan 2011)

B. Les réponses proposées par deux élèves de première S

Élève 1

a) $p(X=8)=\dfrac{8}{1\,024}=\dfrac{1}{128}$
b) Oui, ils peuvent croire le devin car il y a une chance sur 128 que ce soit un faux devin.



Élève 2

a) X = nombre de succès donc X suit la loi binomiale.

$$p(X=8)=\binom{10}{8}\times\left(\frac{1}{2}\right)^8\times\left(\frac{1}{2}\right)^2\approx 0{,}04$$

b) Ils peuvent donc le croire puisque la probabilité d'obtenir 8 bonnes réponses sur 10 est faible. Il peut aussi y avoir une marge d'erreur, le devin a fait un seul essai, cela peut fluctuer. Il faudrait savoir si on est dans l'intervalle de fluctuation

C. Le travail à exposer devant le jury

1. Analysez les productions de ces élèves, notamment par rapport à la notion de fluctuation.

2. Exposez une correction de l'exercice comme vous le feriez pour des élèves de première S.

3. Présentez deux ou trois exercices sur le thème « **Loi binomiale** », dont l'un au moins nécessite la mise en œuvre d'une simulation sur tableur.

2. Eléments de correction

L'exercice a pour objectif une « utilisation d'une loi binomiale pour une prise de décision à partir d'une fréquence ».

Un objectif des programmes de la classe de première est d'exploiter l'intervalle de fluctuation à un seuil donné, déterminé à l'aide de la loi binomiale, pour rejeter ou non une hypothèse sur une fréquence et d'amener les élèves à expérimenter la notion de «différence significative » par rapport à une valeur attendue.

On peut émettre deux réserves quant au choix de l'exercice :

- La question 1, qui demande la probabilité d'obtenir *exactement* huit bonnes réponses et non *au moins* huit bonnes réponses, peut induire de fausses représentations (c'est le cas de l'élève 2).
- L'habillage de l'exercice, fantaisiste, amène à la mise en place d'un test d'hypothèse *unilatéral*. L'intervalle de fluctuation à 95 % classique, tel qu'il est défini dans les programmes, ne donne pas la réponse.

1. Analyse des travaux d'élèves
Aucun des deux élèves ne propose de réponse satisfaisante.

Elève 1.
Réussites.
L'élève 1 sait représenter la répétition d'expériences identiques à deux issues et indépendantes par un arbre pondéré. Il sait que le nombre de branches double à chaque étape et calcule correctement le nombre total de branches qu'aurait cet arbre s'il était achevé (1024). La pondération explicite de l'arbre n'est pas nécessaire ici puisque la pièce est supposée « équilibrée ».
Echecs
L'élève ne reconnaît pas une situation de loi binomiale.
L'arbre pondéré étant inachevé, cet élève ne peut pas dénombrer les branches réalisant l'évènement « obtenir huit bonnes réponses ». Peut être utilise-t-il la formule qu'il invente : $P([X=k])_{gjulia2014} = \frac{k}{1024}$ (?).

Il n'a pas la notion de fluctuation et confond l'évènement objectif dont il a cherché à calculer la probabilité (obtenir 8 bonnes réponses sur 10 en répondant au hasard) avec un caractère subjectif (être un faux devin) non probabilisable à propos duquel on peut seulement avoir une présomption.

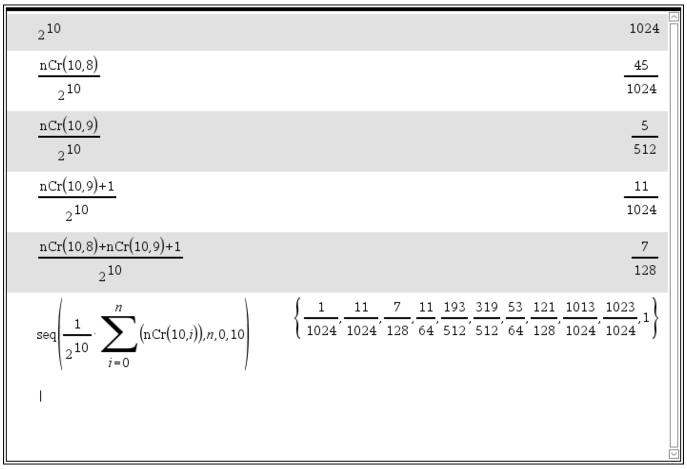
Elève 2.
Réussites
L'élève 2 sait reconnaître une situation relevant de la loi binomiale et sait calculer correctement une probabilité dans le cadre de cette loi.
Il connaît de façon intuitive la notion de fluctuation et d'intervalle de fluctuation (selon lui un devin non infaillible a une certaine « marge d'erreur »)

Echecs
Cet élève en appelle à la notion d'intervalle de fluctuation sans pour autant déterminer un tel intervalle (il ne sait pas très bien ce qui peut « fluctuer »).

La probabilité calculée à la question 1 l'induit à ne raisonner qu'à son propos. Elle est considérée comme « faible » sans qu'il y ait d'exploitation du résultat ni de comparaison explicite à une valeur de référence. Cet échec peut être attribué au moins en partie au peu de pertinence de la question 1 de l'énoncé.

2. On émet l'hypothèse (H_0) que le test proposé s'adresse à une personne prédisant le résultat d'un lancer « au hasard ». Sous cette hypothèse, le nombre X de prédictions exactes est une variable aléatoire qui suit une loi binomiale de paramètres $n=10$; $p=\frac{1}{2}{}_{\text{gjulia2014}}$.

L'écran ci-contre montre quelques résultats utiles, en particulier le résultat attendu à la question 1 mais aussi quelques autres probabilités que le lecteur reconnaîtra.



La probabilité sous l'hypothèse (H_0) de prédire exactement le résultat de 8 lancers est $\frac{45}{1024}$.

A lui seul, ce résultat ne permet pas de développer une argumentation.

Il s'agit de calculer la probabilité d'observer une valeur *au moins aussi grande* (donc de calculer aussi les probabilités de prédire exactement les résultats de 9 et de 10 lancers).

L'intervalle de fluctuation à 95 % tel qu'il est défini dans les programmes est l'intervalle $[a\ ;\ b]$ tel que a est le plus petit entier vérifiant $P([X \le a]) > 0{,}025$ et b est le plus petit entier vérifiant $P([X \le b]) \ge 0{,}975$. La probabilité que X prenne une valeur telle que $a \le X \le b$ est au moins 95 % et les probabilités que $X \le a-1$ ou bien que $X \ge b+1$ sont chacune plus petites que 2,5 %.

Ici, $a=2$ puisque $P([X \le 1]) = \frac{11}{1024} < 0{,}025$ et $P([X \le 2]) = \frac{7}{128} > 0{,}025$ et $b=8$ puisque $P([X \le 8]) = \frac{1013}{1024} \ge 0{,}975$ et $P([X \le 7])_{gj2014} = \frac{121}{128} < 0{,}975$.

La valeur 8 est dans l'intervalle de fluctuation à 95 %. Il n'y a aucune raison de croire le « devin ».

Le contexte est cependant différent ici. Si « pouvoir divinatoire » il y a, ce pouvoir devrait augmenter la probabilité de prédiction exacte. Il s'agit par conséquent de décider si la valeur observée 8 peut être attribuée à une fluctuation naturelle ou bien s'il s'agit d'une valeur anormalement grande (on ne cherche pas à savoir si la valeur observée s'écarte significativement de la moyenne dans un sens ou dans l'autre, mais à savoir si elle est significativement plus grande).

Pour prendre une décision, on définit dans une telle situation l'intervalle $[0\ ;\ c]$ où c est le plus petit entier vérifiant : $P([X \le c]) \ge 0{,}95$. La probabilité que X prenne une valeur telle que $0 \le X \le c$ est au moins 95 % et la probabilité que $X \ge c+1$ est plus petite que 5 %.

Si l'expérience aboutit à une valeur située dans $[0\ ;\ c]$, aucune anomalie n'est mise en évidence. Dans le cas contraire, on conclura au risque 5 % à une anomalie.

Ici $c=8$, la même valeur que b ci-dessus, puisque $P([X \le 8]) = \frac{1013}{1024} \ge 0{,}95$ et $P([X \le 7]) = \frac{121}{128} < 0{,}95$. Les valeurs que l'on considèrera anormalement grandes sont celles qui sont à l'extérieur de $[0\ ;\ 8]$, il s'agit seulement de 9 et 10 mais non de 8.

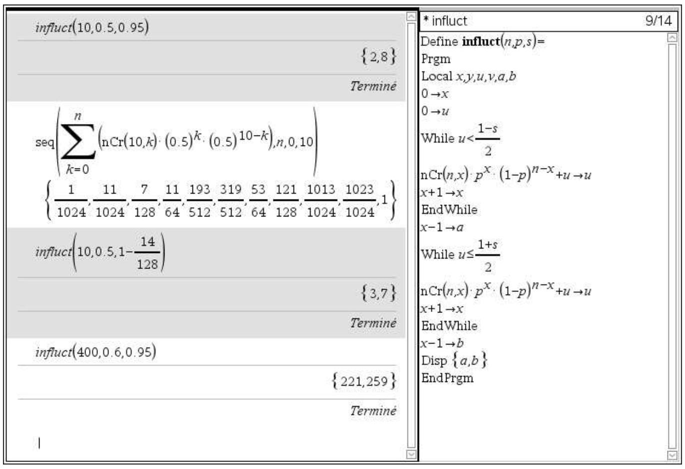
En conclusion, il n'y a aucune raison de croire aux pouvoirs du devin.

3. Pour aller plus loin

Voici un exemple de programme donnant les bornes a et b de l'intervalle de fluctuation pour la loi binomiale $B(n, p)$ au seuil s.

Le programme **influct** est muni de trois arguments, l'entier n, la probabilité p et le seuil s. Il renvoie les entiers a et b dont il est question.

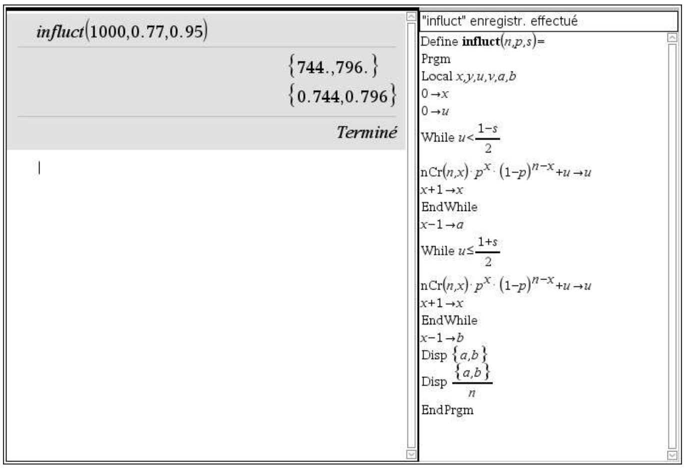
Ci-contre, on a exécuté le programme pour $n = 10$; $p = 0,5$ au seuil 0,95 puis, à titre de vérification, pour $s = 1 - \frac{14}{128}$



Pour cette valeur, le plus petit entier a tel que $P([X \le a]) >_{gjulia\,2014} \frac{7}{128}$ est l'entier 3, et le plus petit entier b tel que $P([X \le b]) \ge \frac{121}{128}$ est l'entier 7.

On a ensuite exécuté le programme dans un cas plus significatif. ($n = 400$; $p = 0,6$)

On a rajouté une ligne au programme, donnant en plus les fréquences



Exécuté avec les valeurs $n = 1000$; $p = 0,77$, ce programme donne un intervalle de fluctuation des fréquences pour le sujet suivant (celui trouvé par l'élève 3 de ce sujet)

ESD 2013 –17 : Prise de décision

1. Le sujet

A. L'exercice proposé au candidat

Une société commercialise des pochettes surprises et affirme que 77 % d'entre elles sont des pochettes bonus, donnant droit à un cadeau.
Afin de vérifier cette affirmation, un organisme de contrôle effectue un tirage aléatoire dans les stocks de l'entreprise (compte tenu de la taille des stocks, ce tirage peut être assimilé à un tirage avec remise). Cet organisme constate que sur les 1000 pochettes choisies aléatoirement, 740 sont des pochettes bonus.
D'après vous, l'organisme doit-il, au vu de ce constat, conduire d'autres investigations pour savoir si l'entreprise n'a pas produit moins de pochettes bonus qu'annoncé ?

B. Les réponses proposées par trois élèves de première

Élève 1
J'ai trouvé $I=[0{,}738\ ;\ 0{,}802]$ en appliquant la formule de l'an dernier. Puisque 0,740 appartient à *I*, je peux d'après le cours affirmer que les différences ne sont dues qu'au hasard. Il n'y a pas d'inquiétude à avoir.

Élève 2
J'ai effectué une simulation sur tableur. Pour simuler le test du tirage d'une pochette avec une probabilité de $p=0{,}77$ que ce soit une pochette bonus j'ai entré la formule qu'on avait vu en séance info et j'ai tiré vers le bas et vers la droite pour créer 100 simulations de 1000 tirages.
Sur 100 simulations de 1000 tirages, 97 m'ont donné une fréquence de pochettes bonus supérieure à 0,75. Or, on a trouvé une fréquence de seulement 0,74. Ce n'est pas normal, il faut rapidement faire une enquête.

Élève 3
J'ai appelé *X* une loi binomiale de paramètres (1000 ; 0,77). Avec le tableur, j'ai déterminé comme dans le cours le plus petit entier *a* tel que $P(X\le a)>0{,}025$ et le plus petit entier *b* tel que $P(X\le b)\ge 0{,}975$. J'ai ainsi abouti à l'intervalle $I=[0{,}744\ ;\ 0{,}796]$. Puisque 0,740 n'est pas dans *I*, je rejette l'hypothèse « après le test, *p* vaut toujours 0,77 » et je conseille de faire une enquête.

C. Le travail à exposer devant le jury

1.- Analysez la production de chaque élève en mettant en évidence les compétences[1] acquises.

2. Proposez une correction de l'exercice telle que vous l'exposeriez dans une classe de première S, en prenant en considération les différentes stratégies mises en œuvre par les trois élèves.

3. Proposez deux ou trois autres exercices sur le thème **prise de décision**, dont l'un au moins pourra donner lieu à une simulation sur tableur.

[1] On peut s'interroger à propos ce que l'on entend ici par « compétence ». Il s'agit plutôt de savoirs et savoir faire. par

2. Eléments de correction

Tout comme dans le sujet précédent, on peut se demander s'il convient à propos de la situation proposée de mettre en place un test unilatéral ou un test bilatéral.

Tel que l'exercice est posé, l' « Organisme d'Inspection des Pochettes Surprises » semble craindre que la société carotte sur le nombre de cadeaux. Dans ce cas, l'OIPS mettra en place un test unilatéral (tant mieux s'il y a plus de cadeaux que prévu, personne ne s'en plaindra, l'important, c'est qu'il n'y en ait pas moins).

Une modification de l'énoncé s'impose : « D'après vous, au vu de ce constat, l'organisme doit-il considérer que l'information donnée par la société est exacte, ou bien doit-elle conduire d'autres investigations ? ». On est plus logiquement amené à tester l'écart dans un sens ou dans l'autre entre la fréquence observée sur l'échantillon et le pourcentage annoncé par la société. Conformément au programme, c'est cette interprétation que nous donnons à l'énoncé.

Le choix d'une fréquence égale à 0,74 semble être délibéré, à la lisière des intervalles de fluctuation que les élèves vont trouver. Selon la démarche adoptée, certains vont accepter l'information de la société, d'autres la rejeter. Cela permettra à l'enseignant d'organiser un petit débat sur les possibilités et les limites des tests d'hypothèse.

1. Analyse des travaux d'élèves

Les trois élèves sont représentatifs de trois méthodes différentes. Deux d'entre eux déterminent un intervalle de fluctuation autour de *p*, le troisième adopte une attitude différente, induite par l'énoncé.

Elève 1.

Cet élève utilise l'intervalle de fluctuation au seuil de 95 % centré en p tel qu'il a été défini en classe de seconde, c'est-à-dire l'intervalle $\left[p-\frac{1}{\sqrt{n}}\ ;\ p+\frac{1}{\sqrt{n}}\right]$. La borne inférieure a été correctement arrondie au millième par défaut, la borne supérieure au millième par excès.

Cet élève sait :

- Mener à bien un test d'hypothèse conformément au programme de seconde. Pour cela, il montre qu'il connaît la formule de seconde donnant un intervalle de fluctuation au seuil 0,95 (savoir) et qu'il sait appliquer le critère décisionnel usuel (savoir faire).

On ne dispose pas d'assez d'éléments pour mettre en évidence une quelconque autre « compétence » au sens propre de ce terme.

Sa réponse peut être considérée comme « acceptable », bien qu'ambiguë. Selon lui, on pourrait « affirmer » que si f appartient à I, alors l'écart est forcément dû au hasard à l'exclusion de toute autre cause. Dans ce cas, cet élève aurait une fausse représentation du test décisionnel. Il ne semble pas avoir conscience qu'accepter l'hypothèse du hasard comporte un risque (on ne peut pas « affirmer » que le pourcentage 77 % est exact, on peut seulement dire qu'il n'y a pas de raison de le mettre en doute).

Elève 2.

Cet élève utilise une simulation. Il utilise implicitement un test unilatéral, influencé par l'énoncé (« … moins de pochettes … ». Pour cela, il considère une liste de « 100 simulations de 1000 tirages » qui sert de référence. Une fois les 100 fréquences rangées par ordre croissant, il considère l'intervalle $[f_4\ ;\ f_{100}]$, qui contient 97 des 100 fréquences observées. Il compare la fréquence observée aux termes de sa liste de référence : Si $f < f_4$, ce qui est le cas pour sa simulation, l'hypothèse nulle est rejetée.

Cet élève sait :

- Elaborer et utiliser une simulation numérique prenant appui sur la modélisation et utilisant un logiciel.

En revanche, cet élève applique un critère de décision empirique et pour cette raison, son argumentation n'est pas recevable.

Elève 3.
Cet élève utilise son tableur pour déterminer exactement les probabilités qu'une variable aléatoire X suivant la loi binomiale $B(1000\ ;\ 0{,}77)$ prenne les diverses valeurs possibles de 0 à 1000. Il en déduit les bornes a et b de l'intervalle de fluctuation au seuil de 95 % d'une fréquence. Les valeurs calculées par cet élève sont exactes (voir fin du sujet précédent).

Cet élève :

- Sait élaborer un programme complexe utilisant un logiciel (compétence). Pour cela, cet élève montre qu'à l'aide de son logiciel, il sait calculer la distribution de probabilité associée à une loi binomiale X et des probabilités liées à la fonction de répartition de X (savoir faire)
- Mener à bien un test d'hypothèse conformément au programme de première.
- Développer une argumentation mathématique correcte en expliquant clairement sa démarche et la raison de sa décision.

Sa réponse peut être considérée comme « acceptable » mais devrait être précisée : Il n'y a pas une probabilité avant le test et une autre après le test, il y a une probabilité supposée, et le test nous permet de dire si cette probabilité est plausible ou bien s'il y a une raison de la mettre en doute.

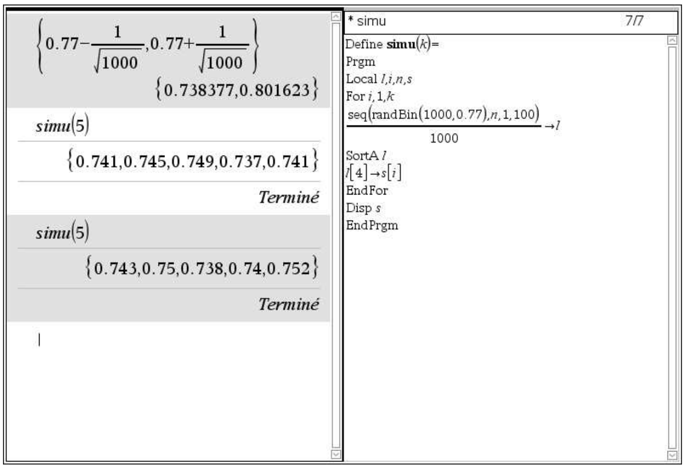
2. Correction de l'exercice dans une classe de première S.

Mettre au point avant toute chose la règle de décision qui va être appliquée. Si la fréquence observée est dans l'intervalle de fluctuation au seuil de 95 %, l'information donnée par la société ne sera pas remise en question. Sinon, elle le sera. Cette mise au point est importante, car les réponses des élèves 2 et 3 montrent que la signification d'un test d'hypothèse n'est pas complètement assimilée.

Commencer par la production de l'élève 1. Cet élève ne remet pas en question la valeur 0,77, il représente de façon correcte la démarche de la classe de seconde, qu'il est opportun de rappeler. Cette démarche est légitime, on se trouve dans les conditions usuelles de son application ($n > 25\ ;\ 0{,}2 < p < 0{,}8$), la grande valeur de n compensant abondamment la proximité de p avec la borne supérieure 0,8.

On considère ensuite en parallèle les productions des élèves 2 et 3. Tous deux ont une conclusion différente, ils remettent en question l'information. Qui a raison ?

Commençons par analyser la démarche de l'élève 2.



Le programme **simu** effectue le même travail que celui produit par cet élève en retenant la quatrième fréquence de la liste des 100 fréquences obtenues, et cela k fois.

On dispose ci-contre des résultats de deux fois cinq simulations identiques à celle de l'élève 2. On constate que huit fois sur dix, 0,74 est en plus petit que la quatrième fréquence et deux fois sur dix plus grand.

On peut émettre une critique majeure et rédhibitoire : le seuil de décision dépend de la simulation, il n'est pas le même pour tout le monde.

En l'occurrence, vu que 0,74 est une valeur lisière, suivant le résultat de la simulation l'information de la société sera ou acceptée ou rejetée alors que la méthode suivie est la même.

La démarche de l'élève 2 n'est donc pas convaincante.

Quant à l'élève 3, il calcule les valeurs exactes de la distribution de probabilités. C'est pourquoi l'intervalle $I_3 = [0{,}744\ ;\ 0{,}796]$ est un peu plus performant que l'intervalle donné en seconde, et ce gain de performance fait basculer la décision dans le cas qui nous intéresse.

On conclut qu'avec la méthode de première, on dispose d'un critère décisionnel de meilleure qualité … mais après bien des efforts.

Il reste en effet à relativiser les choses. Le choix de 0,74 en lisière des intervalles de fluctuation est volontairement « provocateur », l'enseignant voulait sans doute insister sur ce point. Le travail de l'élève 3 est coûteux en temps et en énergie, tout cela pour un gain de six millièmes. Le jeu en vaut-il la chandelle ?

Plus tard, en terminale, les élèves vont disposer de l'intervalle $\left[p - \frac{1{,}96\sqrt{p(1-p)}}{\sqrt{n}}\ ;\ p + \frac{1{,}96\sqrt{p(1-p)}}{\sqrt{n}}\right]$, c'est-à-dire ici $[0{,}743\ ;\ 0{,}797]$, forcément inclus dans l'intervalle de seconde et plus performant que lui, qui va discriminer la valeur 0,74 sans pour autant nécessiter des calculs importants.

Pour terminer, j'attribuerais le mot de la fin à l'élève 1. Selon lui, « il n'y a pas à s'inquiéter ». Je suis bien de son avis. N'est-ce pas lui le plus philosophe des trois ? Est-ce bien important que pas tout à fait 77 % des pochettes surprises contiennent effectivement un cadeau ? Je soupçonne l'auteur de l'exercice d'avoir délibérément « poussé le bouchon un peu loin » en choisissant cet habillage d'exercice caricatural …

3. Exercices complémentaires

Dans l'exercice du Jury, on est amené à comparer la méthode seconde et la méthode première S dans un cas où toutes les deux sont légitimes et où elles donnent des résultats voisins (seule la valeur 0,74 met un pied dans un petit plat). Il serait intéressant de proposer une situation où la méthode de seconde ne peut pas être appliquée et où uniquement la méthode de première donnera des résultats fiables (par exemple en choisissant une assez grande valeur de n comme 500 ou 1000, mais $p < 0{,}05$)