

POLLUTION ET SEX RATIO

Niveau

Seconde.

Situation étudiée

Dans un village du Canada situé dans une réserve indienne, on constate qu'entre 1999 et 2003, sur 132 naissances il n'y a eu que 46 garçons. Ce phénomène peut-il raisonnablement être du au seul hasard ou bien faut-il chercher d'autres causes (la proximité d'usines chimiques par exemple) ?

Type d'activité

Exercice en classe ou évaluation.

Durée

20 minutes.

Objectifs

Contenus mathématiques au programme

- Calcul de fréquences.
- Notion d'intervalle.
- Utilisation d'une formule qui permet de préciser la notion de fluctuation d'échantillonnage.

Enjeux citoyens

Comprendre comment une étude statistique peut permettre de détecter une anomalie touchant une certaine population (ici le sex-ratio chez une population vivant dans un réserve située au cœur d'industries chimiques) et par la suite encourager des recherches pour cerner les causes de cette anomalie.

Aborder la notion de « différence significative ».

Compétences et attitudes

Comprendre et savoir utiliser dans le contexte de l'activité les définitions données : sex ratio et « différence significative à 95% ».

Interpréter les calculs effectués : s'interroger...

Organisation

L'activité (évaluation ou exercice) se situe en fin de séquence. La formule donnant l'intervalle de fluctuation doit déjà avoir été travaillée dans le cadre d'une expérimentation.

Description de l'activité

L'énoncé suivant a été donné en devoir surveillé dans une classe de seconde en fin de chapitre de statistique.

Les données proviennent d'une étude effectuée au Canada et montrant une différence (très) significative sur le sex-ratio à la naissance (déficit de garçons) sur une population exposée à une pollution chimique. Dans ce cas particulier, l'inquiétude provient du fait que bien que ces industries canadiennes respectent les normes, une exposition prolongée à de faibles doses de polluants puisse avoir un impact sanitaire mesurable.

Énoncé élève

Le « sex-ratio » est le rapport du nombre de garçons à celui des filles à la naissance. Il est habituellement de 105 garçons pour 100 filles.

1) La fréquence habituelle des garçons est $p = \frac{105}{105+100} = 0,512$. Si cette proportion est exactement respectée, combien de garçons et de filles a-t-on sur 1000 naissances ?

2) Si l'on prélève des échantillons aléatoires de taille n dans une population où la fréquence étudiée est $p = 0,512$, dans plus de 95 % des cas, la fréquence f observée sera comprise dans l'intervalle $[0,512 - \frac{1}{\sqrt{n}}; 0,512 + \frac{1}{\sqrt{n}}]$.

Si f n'appartient pas à cet intervalle, on dira que f présente une « différence significative » au niveau 0,95 avec $p = 0,512$.

a) Dans la réserve indienne d'Aamjiwnaag, située au Canada, il est né entre 1999 et 2003, $n = 132$ enfants dont 46 garçons.

Que vaut la fréquence f des garçons pour cette période à Aamjiwnaag (arrondir à 10^{-3}) ?

b) La fréquence des garçons observée à Aamjiwnaag pour la période 1999-2003 présente-t-elle une « différence significative » avec $p = 0,512$ (justifier par un calcul) ?

Sources : Science et Vie février 2006 – Environmental Health Perspectives octobre 2005 (article en ligne).

Éléments de réponse

1) Si la proportion $p = 0,512$ est exactement respectée, on doit avoir, sur 1000 naissances, 512 garçons et 488 filles.

2) a) La fréquence de garçons f observée à Aamjiwnaag est $f = 46 / 132 = 0,348$.

b) La différence est significative lorsque la fréquence observée n'appartient pas à l'intervalle $[0,424 ; 0,599]$.

0,348 n'appartient pas à cet intervalle donc la valeur observée présente une différence significative avec $p = 0,512$.

Déroulement et commentaires

Voici quelques exemples de copies d'élèves réalisées en devoir surveillé.

Les deux copies suivantes sont des exemples de réponses « satisfaisantes ».

① SEX-RATIO

① Si cette proportion est respectée, il y aura :

- 512 garçons
- 488 filles

... sur 1000 naissances. *Bien*

② a/ Ici le sex-ratio ρ vaut : $\frac{46}{132} = \boxed{0,348}$

par cette période *Bien*

③ $\left[0,512 - \frac{1}{\sqrt{132}} ; 0,512 + \frac{1}{\sqrt{132}} \right]$ donc

$\hat{\rho} [0,424 ; 0,593]$

- Oui il y a donc une différence significative car le résultat obtenu (ρ nouvelle) est très éloigné du résultat censuré de la ville.

Intervalle $[42,4\% \text{ et } 59,3\%]$ alors que le résultat obtenu est de $\boxed{34,8\%}$ à 95% de chance *Bien*

Autre exemple correct :

I. Sex-ratio et pollution.

1) On aura 512 garçons et 488 filles. *oui*

2) Dans la réserve $p = 0,348$. *exact*

3) $\left[0,512 - \frac{1}{\sqrt{132}} ; 0,512 + \frac{1}{\sqrt{132}} \right] = [0,426 ; 0,99]$

→ p n'appartient pas à cette intervalle, il présente donc une différence significative. *oui*

Dans la copie suivante, l'élève calcule à la fois un intervalle de confiance autour de la fréquence f observée (ce qui n'est pas la démarche demandée même si elle est ici équivalente) et un intervalle de fluctuation autour de la valeur p théorique. Les deux intervalles ont une partie commune et la conclusion de l'élève est confuse.

I. 1) Sur 1000 naissances p y aura 512 et 488.
si cette répartition est exactement respectée. *oui*

2. a) $f \left[p - \frac{1}{\sqrt{nt}} ; p + \frac{1}{\sqrt{nt}} \right]$
 $f \left[0,35 - \frac{1}{\sqrt{1321}} ; 0,35 + \frac{1}{\sqrt{1321}} \right]$
 $f \left[0,26 ; 0,44 \right]$
 $f \left[26\% ; 44\% \right]$. ← *Ce n'est pas la question posée*

b) $f \left[0,512 - \frac{1}{\sqrt{1321}} ; 0,512 + \frac{1}{\sqrt{1321}} \right]$
 $? \left(f \left[0,42 ; 0,60 \right] \right)$ *oui*

La sex-ratio à Ramjivnagar présente une différence significative avec $p = 0,512$. *exact, mais*
Comment utilises-tu le calcul précédent ?

Dans la copie suivante, la « définition » donnée dans l'énoncé de « différence significative » n'est pas respectée. L'élève n'utilise pas, pour conclure, les fluctuations des échantillons de taille 132.

1) pour $p = 0,512$ il y aura normalement 512 garçons et 488 filles sur 1000 naissances. oui

2)

a) Le sex-ratio f pour cette période est $\frac{46}{132} = 0,348$ oui

b) $A_{\text{amjiwnaag}} = 0,348$ et $p = 0,512$ pour cette période le sex-ratio de Amjiwnaag présente une différence significative plus de 15% de différence (16,4%)
 \rightarrow Ce n'est pas la définition donnée dans le texte

2

Dans cette copie, c'est la définition de « fréquence » qui n'est pas maîtrisée.

1) Si cette proportion est exactement respectée on aura sur 1000 naissances 512 garçons et 488 filles. oui

2) $[46 - \frac{1}{\sqrt{132}}; 46 + \frac{1}{\sqrt{132}}]$
 $[45,91; 66,08]$
 soit $[45,91\%; 66,08]$
 46 n'est pas une fréquence

3) $[0,512 - \frac{1}{\sqrt{132}}; 0,512 + \frac{1}{\sqrt{132}}]$

Les erreurs les plus souvent rencontrées proviennent d'erreurs de calcul sur les fréquences ou de confusions entre estimation (intervalle centré sur la valeur observée) et test (intervalle d'acceptation de l'hypothèse centré sur la valeur correspondante de la

fréquence). Dans l'ensemble, le caractère « anormal » des statistiques d'Aamjiwnaag a été vu et plusieurs élèves ont posé, à la fin du contrôle, la question de savoir s'il s'agissait d'une histoire « vraie ».

Bilan et perspectives

Tel que l'énoncé est présenté ici, il soulève inmanquablement la question « Que peut-on tirer comme conclusion de ces statistiques ? ».



Une première version de l'énoncé ajoutait comme précision que la réserve d'Aamjiwnaag est située « à proximité de nombreuses industries chimiques ». Cela laissait entendre qu'il y avait une relation de causalité entre la pollution chimique et le « défaut » observé du sex-ratio, causalité qui n'est pas établie par l'étude statistique.

A la question « Que peut-on tirer comme conclusion ? », on peut seulement ici répondre que « cette étude pose question ». La statistique donne l'alerte, ce qui est déjà beaucoup. Le fait que la réserve soit située au cœur d'industries chimiques devient un élément troublant sur lequel on doit enquêter.

De façon générale, c'est la notion de « preuve statistique » qui est ici en jeu. Il ne s'agit pas d'une « preuve » au sens habituel mais d'un élément probant. Plutôt que de parler de « preuve statistique », les anglo-saxons disent plus justement *piece of evidence*.

D'autres explications possibles du déséquilibre du sex-ratio pourraient être liées au modes de vie de ces indiens ou à leur patrimoine génétique. Une étude statistique comparative a été menée sur des indiens de la même tribu vivant dans un autre environnement et a (dé)montré que ce n'était (sans doute) pas le cas. En revanche l'influence de certains produits chimiques sur le sex-ratio a été établie « statistiquement » par d'autres études.

Une recherche sur Internet permettra d'avoir d'autres éléments sur ce dossier (qui a fait polémique au Canada).

<http://www.ehponline.org/docs/2005/8479/abstract.html>