

MATH & MEDIAS

On trouvera ci-dessous une brève extraite de libération, ainsi que les commentaires qu'elle a suscité. Merci d'envoyer tous extraits de presse qui vous paraissent intéressants et exploirables (notamment en classe) à jacquesverdier@free.fr

PLUTOT FACE QUE PILE ?

Cher Wim,

Tout avait si bien commencé, or voici que les scientifiques s'en mêlent. Une équipe de statisticiens polonais affirme que la pièce d'un euro a plus de chance de tomber côté face que côté pile. Tu mesures l'ampleur du problème. Nous faudra-t-il garder quelques francs - ou lires, marks, etc. - pour trancher les grandes questions, genre pile je fais la vaisselle, face tu t'y colles ? Pour l'heure, cette asymétrie n'a toutefois été démontrée que pour les pièces belges. Les mathématiciens Tomasz Gliszczynski et Wacław Zawadowski ont procédé à 250 essais sur un euro frappé outre-Québécois. Résultat: 140 "face" (56 %) et 110 "pile" (44 %). La faute au roi Albert dont l'auguste profil orne le côté face des pièces belges. Le problème serait que la tronche d'Albert ne pèse pas assez lourd.

L'euro "français" sera-t-il impartial, lui, quand sonnera l'heure de la vaisselle ? La rédaction de *Libération* a procédé hier à sa propre campagne de tests. L'échantillon polonais de 250 essais nous semblant un peu limité, nous avons décidé de le porter à 1 000. Il en a résulté ce qui suit. Au cinquième essai, la pièce a roulé sous une armoire de forte carrure, et de ce fait indéplaçable. Nous en étions à quatre "pile" contre un "face". Puis une difficulté expérimentale a surgi : le propriétaire de la pièce a refusé de remettre la main à la poche, prétextant un urgent besoin de café et une bourse hélas ! plate. L'argument a été jugé recevable par les autres membres de l'équipe. C'est donc sur une base de cinq essais qu'il nous a fallu conclure : l'euro "français" est lui aussi biaisé.

Plus aguerrie sans doute, l'équipe de la revue britannique *New Scientist* est parvenue à refaire l'expérience sur l'euro belge. Pour arriver à un résultat exactement inverse à celui des Polonais.

(*Libération*, 7 janvier 2002)

A propos de cet article, voici quelques informations pour le statisticien "débutant". Tout le problème est de savoir à partir de quel écart on peut décider, sans prendre trop de risque, que la pièce n'est pas "symétrique". On comprend tout de suite que s'il y avait eu 215 'FACE' contre 35 'PILE', on aurait tout de suite pensé "Il y a quelque chose qui cloche dans cette pièce, elle n'est pas symétrique". Si on avait eu 128 'FACE' contre 122 'PILE', on aurait pensé que, somme toute, le hasard aurait fort bien pu donner un tel résultat avec une pièce parfaitement symétrique. Mais 140 'FACE' contre 110 'PILE' ... ça se discute peut-être.

(Suite page 12)

Pour la bonne compréhension du lecteur qui n’a jamais abordé les statistiques inférentielles (car c’est de cela qu’il s’agit), je commencerai par la dernière expérience, celle de la rédaction de *Libération* : 4 ‘PILE’ contre 1 ‘FACE’ en 5 coups. Cela fait quand même quatre fois plus de ‘PILE’ que de ‘FACE’ : cela semble très déséquilibré.

Même si on n’a jamais entendu parler de loi binomiale, on peut lister tous les événements qui peuvent se produire quand on lance 5 fois de suite une pièce : PPPPP, PPPPF, PPPFF, PPPFF, PPFPP etc. ; c’est facile, il y en a $2^5 = 32$. Parmi ces 32, il y en a 5 qui donnent exactement 4 ‘PILE’ contre 1 ‘FACE’ (PPPPF, PPPFP, PPFPP, PFPPP et FPPPP), soit cinq chances sur trente-deux que ce qui est arrivé se produise (en pourcentage, 15,6 % environ) : ce n’est pas beaucoup, mais ce n’est pas négligeable.

Mais en réalité, si l’on voulait savoir si la pièce avait plutôt une propension à tomber du côté ‘PILE’, il aurait fallu ajouter à notre liste l’événement PPPPP ; cette fois, cela fait six chances sur 32 (environ 18,7 %).

En y réfléchissant bien, ce qui nous intéresse n’est pas de savoir si la pièce tombe plutôt sur ‘PILE’ que sur ‘FACE’, mais si elle tombe plutôt d’un côté que de l’autre.

De la même façon que ci-dessus, on peut calculer les probabilités des quatre événements suivants :

$P(4 \text{ ‘PILE’ et } 1 \text{ ‘FACE’}) = 5/32 \approx 15,6 \%$

$P(\text{au moins } 4 \text{ ‘PILE’}) = 6/32 \approx 18,7 \%$

$P(4 \text{ ‘FACE’ et } 1 \text{ ‘PILE’}) = 5/32 \approx 15,6 \%$

$P(\text{au moins } 4 \text{ ‘FACE’}) = 6/32 \approx 18,7 \%$

Finalement, on en déduit que $P(\text{au moins quatre fois du même côté}) = 12/32 \approx 37,5\%$.

Tout cela avec une pièce supposée parfaitement symétrique (notons au passage que si la pièce n’avait pas été un ‘modèle’ de pièce symétrique, on n’aurait rien pu calculer du tout).

Finalement, l’événement “ au moins quatre fois du même côté ” n’est pas rare du tout quand on lance cinq fois de suite une pièce, puisqu’il a plus d’une chance sur trois de se produire. Autrement dit, s’il se produit, on ne pourra pas incriminer la pièce (en langage de statisticien, on ne pourra pas **rejeter l’hypothèse** que la pièce est ‘parfaite’).

Recommençons la même chose avec l’expérience de Gliszczynski et Zawadowski : quelle est la probabilité, en lançant 250 fois une pièce ‘idéalement parfaite’, d’obtenir au moins 140 fois le même côté ? Si cette probabilité est négligeable, on pourra sans prendre trop de risque, dire “ *Le hasard n’a pas pu donner ce résultat, donc la pièce est asymétrique* ”. Si cette probabilité n’est pas assez petite, on ne dira rien ... ou on recommencera l’expérience 1 000 fois, comme voulaient le faire les journalistes de *Libération*.

Si jusque là, lecteur débutant, tu as tout compris, parfait...tu peux poursuivre.

Quelle est donc la probabilité qu’une pièce ‘parfaite’, lancée 250 fois, tombe au moins 140 fois du même côté ? Les choses se compliquent bigrement, car il y a 2^{250} éléments dans la liste des issues possibles, ce qui est de l’ordre de 10^{75} : le calcul précédent est infaisable...

Nous allons prendre les choses autrement, et calculer “ l’écart ” entre ce qui a été observé (ici, 140 ‘FACE’ et 110 ‘PILE’) et ce qui était attendu (en théorie, 125 ‘FACE’ et ‘125 ‘PILE’). Première façon de compter l’écart : 15 (= 140 – 125). Deuxième façon de compter l’écart : 30 (15 pour les ‘PILE’ et 15 pour les ‘FACE’). J’en vois qui ricanent : s’il y a 15 d’écart pour ‘PILE’, bien sûr qu’il y en a 15 pour ‘FACE’.

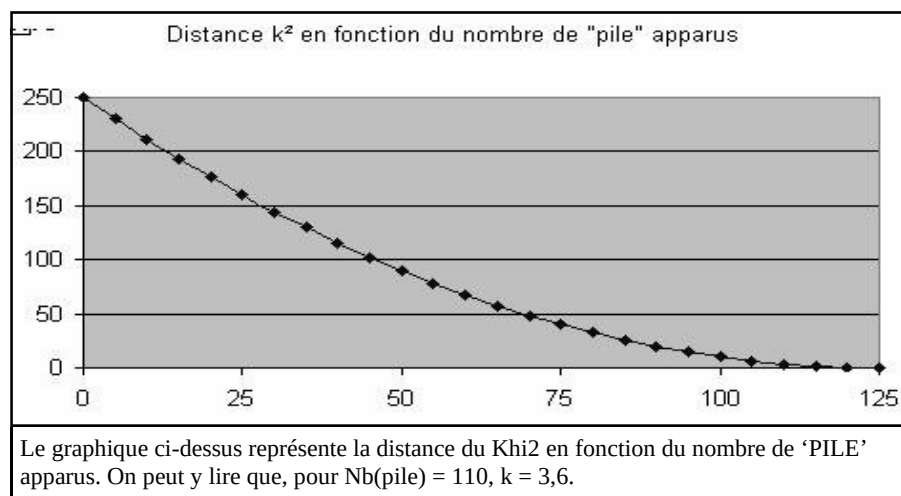
Pas si simple... J'ouvre une parenthèse. Au lieu de lancer une pièce, j'ai lancé un dé. 300 fois. En théorie, je devrais avoir 50 'AS', 50 'DEUX', 50 'TROIS', 50 'QUATRE', 50 'CINQ' et 50 'SIX'. En réalité, j'ai eu 42 'AS', 61 'DEUX', 58 'TROIS', 44 'QUATRE', 46 'CINQ' et 49 'SIX'. Pour compter l'écart "total", j'ai envie de faire la somme des écarts (en valeur absolue, bien sûr) : $8 + 11 + 8 + 6 + 6 + 1 = 40$. C'est ici parfaitement logique de faire la somme de **tous** les écarts. Fin de la parenthèse.

Malheureusement, les mathématiciens n'aiment pas les sommes de valeurs absolues, parce que ça ne donne jamais rien de bien calculable. Allez savoir pourquoi, ils préfèrent les sommes de carrés... Dans le cas de notre dé, cela donnerait : $8^2 + 11^2 + 8^2 + 6^2 + 6^2 + 1^2 = 323$. Dans le cas de nos pièces, $15^2 + 15^2 = 450$.

On se rend tout de suite compte qu'une telle façon de compter n'est pas satisfaisante : un même écart obtenu sur 40 lancers (5 'PILE' et 35 'FACE'), sur 250 lancers (110 'PILE' et 140 'FACE') ou sur 100 000 lancers (49 985 'PILE' 50 015 'FACE') n'a évidemment pas la même signification ! Le bon sens nous dit que ce dernier écart est proportionnellement bien plus faible. Nous allons donc "**normer**", en rapportant chaque écart (au carré) à l'effectif théoriquement attendu. Dans le cas de la pièce, on attendait 125 'PILE' et 125 'FACE', nous allons donc diviser 15^2 par 125 et 15^2 par 125. Ce qui nous fait un écart de 3,6 au total. Cette mesure de l'écart porte un nom : c'est la "**distance du Khi2**"

2 / 1 0

Il est évident que si la pièce est 'parfaite', on a peu de chances d'avoir un grand écart (on a d'autant moins de chances que l'écart est plus grand). Le problème se reformule donc autrement : quelle est la probabilité, pour notre expérience (lancer 250 fois une pièce), que la distance du Khi2 observée soit supérieure à 3,6 ? Si cette probabilité est faible, on rejettera l'hypothèse de la pièce symétrique (en disant, en quelque sorte, "*il est fort peu probable qu'avec une telle pièce le hasard puisse nous donner un si grand écart*"). Le problème est transformé, mais nous n'avons toujours pas la réponse...



Heureusement, Monsieur PEARSON (prononcer 'personne') est arrivé ... Il a étudié de façon théorique la loi de probabilité de cette distance, et a dressé des tables (N.B. c'est justement parce qu'il a additionné les carrés et non les valeurs absolues qu'il a pu faire ces calculs). Lesquelles tables nous disent que, dans le cas qui nous intéresse (deux pièces), la probabilité que cette distance soit supérieur à 3,6 est d'à peine 5 %. Nous sommes obligés de lui faire confiance...

Le mathématicien ayant calculé, il nous reste à prendre une décision : il s'est produit un événement qui n'avait qu'une chance sur vingt de se produire. Si nous décidons que le hasard n'a pas pu intervenir seul (autrement dit, dire que la pièce d'un euro belge ne peut pas être équilibrée), nous prenons un risque : 5 % de chances de nous fourvoyer. Si nous ne voulons pas prendre un tel risque, nous ne concluons pas (ou nous recommencerons l'expérience à une plus grande échelle). Et là, le mathématicien ne peut plus rien pour nous... C'est à l'utilisateur de prendre la décision, donc de prendre le risque..

J'espère que ces quelques lignes auront permis au lecteur non initié des arcanes de la statistique inférentielle d'en comprendre l'enjeu. C'est, pour les élèves, ce qu'il y a de plus difficile.

Jacques VERDIER