

Quelques questions de méthode

rappel : les références sont en ligne sur <http://hussonet.free.fr/attac0.htm>

1) Alain Lecourieux fait référence, dans un de ses derniers courriels, au théorème d'incomplétude de Gödel. Alain nous en résume ainsi la teneur : « Toutes les formulations axiomatiques consistantes de la théorie des nombres incluent des propositions indécidables ». Puis il passe immédiatement à des considérations sur les hypothèses de la démonstration de la fraude et énonce un nouveau théorème (de Lecourieux ?) : « ou bien donc leurs hypothèses sont vraies, ou bien elles sont fausses, ou bien elles sont incomplètes ».

On voit bien ce qu'entend suggérer Lecourieux : que la question de savoir s'il y a eu ou non fraude est une « proposition indécidable ». La référence au théorème de Gödel, dont la formulation exacte est donnée en post-scriptum, sert ici d'argument d'autorité.

Avant de répondre plus en détail, je ne résiste pas à la tentation de rappeler un précédent. Il y a quelques années, Régis Debray avait lui aussi invoqué le théorème de Gödel pour « démontrer » l'impossibilité de l'autogestion, parce que « le gouvernement d'un collectif par lui-même serait une opération logiquement contradictoire ». Il s'était fait sérieusement étriller, et d'autres avec lui, pour cette transposition impressionniste d'un théorème de logique dans celui des sciences sociales et politiques. Je me permets de renvoyer au livre d'Alain Sokal et Jean Bricmont, *Impostures intellectuelles* (Ed. Odile Jacob, 1997).

2) Il faut partir ici d'une distinction fondamentale entre axiome, hypothèse (et peut-être même conclusion). En épistémologie, un **axiome** est (je prends la définition de Wikipedia) est « une vérité évidente en soi sur laquelle une autre connaissance peut se reposer, autrement dit peut être construite dessus ». Un axiome est donc quelque chose que l'on se donne comme matériau initial d'une élaboration théorique, mais qui ne peut être démontré. Le théorème de Gödel concerne les formulations axiomatiques et n'a donc rien à voir ici où nous n'en posons aucun.

3) Les études statistiques du vote qui ont été présentées par Thomas et moi-même partent de **postulats** et d'**hypothèses**. Toujours selon Wikipedia, un postulat est « un principe utilisé dans la construction d'un système déductif, mais qu'on ne démontre pas lui-même, sans pour autant s'interdire la possibilité de s'y essayer plus tard (en ce sens, le postulat se distingue de l'axiome, toujours posé au départ comme fondamental qu'on ne cherchera donc pas à démontrer) »

Nous n'avons utilisé qu'un seul postulat : que les trois récapitulatifs successifs portant sur 5862 bulletins cumulés sont représentatifs du dépouillement. C'est une approximation : il y a eu sans doute des erreurs de pointage, une lettre B qui a été lue comme une lettre D après photocopie de la feuille d'émargement, et des votes de dernière minute qui n'ont pas été pris en compte puisque le résultat définitif porte sur 6051 bulletins exprimés. Mais j'ai pris soin (voir tableau 3 de *Radiographie d'une blessure*) de vérifier que ce récapitulatif était suffisamment proche du décompte final dont je ne dispose pas. Le recomptage par huissier permettra d'ajuster ces écarts du second ordre, sur lesquels on ne peut évidemment fonder une contre-argumentation.

Venons-en aux hypothèses. J'emprunte à nouveau à Wikipedia : « Une *hypothèse* est une proposition ou une explication que l'on se contente d'énoncer sans prendre position sur sa véracité c'est-à-dire sans l'affirmer ou la nier. Une fois énoncée, on se propose de l'étudier, de travailler avec, d'adopter une attitude par rapport à elle ». Cette définition correspond bien à la démarche suivie je rappelle les étapes.

J'ai commencé par comparer les résultats en % de chacun des lots aux résultats obtenus sur le Segment 1 (les 3000 premiers bulletins dépouillés). L'hypothèse est ici que le Segment 1 constitue une référence « propre » par rapport à laquelle sont évaluées d'éventuelles « anomalies ». Les sceptiques vont dire : vous voyez bien ! Qui vous dit que ces 3000 bulletins constituent une

référence « conforme » ? Voilà bien une hypothèse « indécidable » ! Je conseille à ceux qui voudraient se saisir de cette ligne de critique de ne pas y consacrer trop de temps, car il s'agit là d'une hypothèse transitoire, qui devient superflue à la fin de l'étude.

Les 10 lots repérés comme « atypiques » le sont aussi par rapport aussi à l'ensemble des résultats. Nous avons donc clairement isolé 10 lots sur 59, qui regroupent 837 bulletins sur 5862, soit 14 % seulement de l'ensemble.

Je souligne aussi un résultat très fort de l'analyse statistique du vote. Il y a dans les résultats 23 lots qui comptent moins de 50 bulletins. Leur taille moyenne est de 28 bulletins, contre 140 bulletins en moyenne pour les 36 autres lots. En raison de leur taille réduite, ces petits lots présentaient évidemment une fiabilité moins grande et je pensais au départ les laisser de côté. Mais mes réflexes de statisticien m'ont conduit à regarder ce qui se passait quand on les regroupait en un seul, et je dois dire que j'ai été surpris du résultat : le « segment » ainsi obtenu par addition des petits lots se rapprochait à peu près parfaitement de la distribution moyenne hors lots atypiques. On peut vérifier ici que le hasard ne signifie pas l'absence de loi, et c'en est une qui fonctionnait sous mes yeux émerveillés.

Mais il en va tout autrement avec les lots « atypiques » qui restent « atypiques », même après regroupement. **Nous parvenons donc à un fait massif, qui n'est pas une « hypothèse » : il existe deux catégories de lots nettement différenciés.**

Corrélation et Khi-deux

Un petit détour méthodologique, pour répondre à cette remarque d'un autre correspondant (J.L. Linosse) qui m'écrit : « Je ne suis pas persuadé que le coefficient de corrélation soit pertinent car ce paramètre n'est habituellement pas utilisé pour comparer des lots issus d'une population mais pour établir une possibilité de lien entre plusieurs variables. Ce qui n'est pas le cas de cette analyse où il n'y a pas de variables différenciées. Et des coefficients trop faibles sont réputés souvent sans signification. Il serait sans aucun doute plus significatif de comparer les lots entre eux puisque c'est la base de "l'accusation". »

Ces assertions sont curieuses. Le coefficient de corrélation est parfaitement adéquat lorsqu'il s'agit de répondre à la question suivante : les votes sont-ils différents d'un « lot » à l'autre ? Quant à comparer les « lots » entre eux, c'est exactement ce que je fais.

Il existe un autre test, celui du Khi-deux, et les calculs effectués par Christophe Osswald montrent qu'il y a moins d'une chance sur un million pour que les lots « atypiques » et « conformes » proviennent d'une même population.

4) Il faut maintenant, après l'avoir repéré, construire une explication, un « scénario », qui permette de rendre compte de l'existence de deux grandes catégories de lots. La démarche suivie s'est déroulée en deux temps.

Premier temps : envisager et discuter tous les biais possibles permettant de rendre compte de cette différenciation. Je renvoie aux documents précédents pour montrer sur quelle base argumentaire on peut les rejeter tour à tour.

Second temps : répondre à cette question qu'Alain Lecourieux pose à nouveau : qui sont les « atypiques », qui sont les « conformes » ? Voici les trois arguments qui permettent de trancher :

a) la proportion : d'un côté, 4845 bulletins, de l'autre 837, soit six fois moins. Il aurait fallu une fraude extraordinairement massive (pour ne pas arriver au résultat souhaité).

b) la chronologie : tous les lots atypiques ont été dépouillés **après** la première moitié des bulletins et ont été suivis de nouveaux lots redonnant les mêmes résultats que les 3000 premiers, dont ceux reçus par la poste le dernier jour (15P) et qui ont été dépouillés immédiatement. On voit mal comment et pourquoi la fraude aurait porté d'une manière extraordinairement bien répartie sur les 3000 premiers bulletins, puis aurait cessé (au moment où c'était le plus facile puisque tous les bulletins étaient entassés et sortis de leurs enveloppes) pour reprendre sur les derniers bulletins triés immédiatement après leur arrivée. Pour que cette fraude ne soit pas décelable statistiquement, il aurait fallu en outre qu'elle porte sur l'intégralité des 49 lots, avec une très grande régularité. Pour être encore plus concret, une nuit sépare le bilan à mi-parcours et l'apparition des premiers bulletins atypiques.

c) le résultat final : il faut imaginer des scénarios particulièrement tordus pour penser que les fraudeurs, capables de manipuler (avec un art consommé) 83 % des lots représentant 85 % des bulletins, n'auraient pas été fichus de faire basculer le score final en leur faveur.

d) il reste un quatrième élément qu'il sera possible, je l'espère, d'analyser : c'est la structure interne des lots atypiques afin de tester la fréquence de bulletins très voisins, voire identiques. Mais cela suppose que le classement des bulletins par lots ait été conservé avant la pose des scellés. Un recomptage global, même par huissier, n'apportera de ce point de vue aucun élément utile.

5) Les coups super-tordus

Alain Lecourieux répercute l'hypothèse imaginée par François Cambillard, (Attac Var Toulon) qu'il trouve « astucieuse » et « plus facile à réaliser que la substitution de 500 à 700 bulletins de vote dans le lot 2 » : il y aurait pu avoir « manipulation des échantillons réputés "solides" par transfert de voix d'un échantillon à l'autre pour créer la suspicion de fraude sans changer le résultat ».

On peut féliciter François Cambillard pour son imagination fertile (j'avoue ne pas y avoir pensé !) mais cette hypothèse ne tient pas la route non plus. On voit mal comment les tenants de la liste Susan George auraient pris le risque immense de la fraude pour ne rester qu'à mi-chemin et ne pas aller carrément jusqu'au bout. Il faut se rappeler que le vote était très serré et qu'il suffisait de peu de choses pour le faire basculer dans un sens ou l'autre. Il ne fallait pas remplacer 500 à 700 bulletins : cette estimation initiale s'est resserrée. Les lots atypiques contiennent 837 bulletins mais la substitution n'avait pas besoin de porter sur l'ensemble de ces bulletins. Les estimations données dans mon tableau 5 montrent que les transferts de voix ne dépassent pas 200.

Mais cette hypothèse ne tient pas, pour une raison encore plus simple, c'est que la substitution sur les 3000 premiers bulletins auraient dû faire apparaître des distorsions entre les lots de ce Segment 1. Ce dernier aurait alors contenu des lots significativement différents du résultat d'ensemble supposé inchangé. Idem sur le Segment 3. Or, l'analyse par lots ne permet pas de repérer de telles différenciations. Donc, *exit* l'ultra-machiavélisme.

6) Les contre-hypothèses pauvres

Faisant décidément feu de tout bois, Lecourieux évoque « de nombreuses autres hypothèses » comme, par exemple, « un comptage et/ou un report profondément défectueux » qui réduirait la "solidité" des échantillons à néant. « Dans ces cas la démonstration de Michel Husson s'écroule » avertit Lecourieux.

Je pense qu'il n'a pas bien compris les ressorts de cette démonstration. Il y aura bien sûr des erreurs minimales de comptage, mais ces erreurs, si ce sont des erreurs, seront distribués de manière aléatoire. Il n'y a par conséquent aucune raison qu'elles modifient la structure très robuste de la répartition des lots entre « atypiques » et « conformes ». Il faudrait non seulement que le comptage ait été « profondément défectueux » mais qu'il soit systématiquement défectueux au détriment d'une catégorie de candidats. Cela ne ferait que modifier les modalités de la fraude, mais pas la réalité de la fraude.