

Limites et risques des analyses statistiques.

Contribution pour la commission d'enquête sur les élections du CA d'Attac.

Michel Lasserre, le 20 septembre 2006

Les différentes méthodes utilisées par les statisticiens pour analyser les résultats des élections du CA reposent sur la même hypothèse : les lots de bulletins sont considérés comme aléatoires et représentatifs du résultat final.

C'est à partir de cette hypothèse que Michel Husson et Thomas Coutrot, suivis par Gérard Duménil et Dominique Lévy, ont montré que certains lots de bulletins situés vers la fin du dépouillement présentaient des divergences plus ou moins importantes avec les résultats globaux, et en on déduit qu'il y avait eu fraude lors du dépouillement.

Les rapports des experts indépendants, dont René Passet a fait la synthèse, s'appuient pareillement sur cette hypothèse, tout en restant toutefois beaucoup plus prudents dans leurs conclusions : *"Le point central est que les résultats globaux semblent statistiquement peu probables si les lots constitués à partir des lettres sont bien aléatoires"* (S. Jugnot).

Cette contribution montre que les lots, et dans une moindre mesure les lettres ou cumuls de plusieurs lots en fonction des jours de dépouillement, ne sont qu'une approximation pas forcément représentative du résultat final.

Expliquer par une fraude le basculement constaté durant le dépouillement de notre scrutin n'est donc qu'une hypothèse qui a l'avantage d'être simple, mais qui n'a rien de probant face à l'insuffisante représentativité des ensembles de bulletins analysés.

1) Echantillons aléatoires et échantillons représentatifs

Le raisonnement que je développerai s'appuie en partie sur une étude de Michel Guillemet, *"Limites et puissance des sondages"*, parue dans une publication de la Direction Régionale de Rennes de l'INSEE. Elle est téléchargeable à partir de cette adresse internet¹ : <http://m-lasserre.com/attac/elections.htm>

L'élection de juin est terminée, nous en connaissons les résultats, alors pourquoi s'intéresser aux méthodes de sondages qui sont normalement utilisées avant la publication des résultats d'une élection ? La raison de cet intérêt est simple, certains travaux des experts font comme si les lots de bulletins qu'ils analysent étaient des échantillons représentatifs du résultat global. Cette hypothèse peut alors être elle-même analysée à l'éclairage des méthodes employées par les instituts de sondages pour définir leurs échantillons et évaluer leur taux de représentativité.

Contrairement à une étude sur des lancers de dés, où le résultat du lancer relève du seul hasard et ne dépend pas de l'identité du lanceur, une élection dépend de choix non hasardeux. Chaque électeur choisit ses candidats en fonction de critères personnels. Pour pouvoir être considéré

¹ Il est possible que le lien ne fonctionne pas en cliquant dessus, on peut toujours alors le copier dans son navigateur.

comme un échantillon représentatif du résultat final, un lot issu d'un tirage aléatoire devrait alors réunir un très grand nombre d'individus. Or, pour des raisons évidentes de coût, les instituts de sondage limitent volontairement la taille de leurs échantillons, le plus souvent autour de 1000 personnes.

Un échantillon de 1000 personnes choisies au hasard, c'est à dire par tirage aléatoire, offre un gros risque de ne pas être représentatif. Pour compenser ce défaut, les instituts de sondage déterminent selon divers quotas la constitution de cet échantillon. C'est ainsi qu'on choisira les différents sondés suivant leur sexe, leur lieu d'habitation, leur âge, leur profession, ... Ce faisant, les résultats de cet échantillon se rapprocheront de ceux obtenus par un tirage aléatoire comprenant beaucoup plus de sondés.

L'échantillon par quotas sera alors considéré comme représentatif, bien que n'étant plus aléatoire.

D'autre part, on précisera qu'un échantillon dit représentatif n'est pas la copie conforme du résultat final. Ce n'est qu'une estimation, et des marges d'erreurs existent.

L'échantillon s'inscrit dans une fourchette de probabilités, qui dépend du score obtenu par le candidat dans l'échantillon ainsi que de la taille de l'échantillon. C'est ainsi que pour un échantillon "représentatif" de 1000 personnes, un candidat obtenant un score de 40 % aura 95 % de chance que son résultat final soit compris dans une fourchette allant de 36,9 % à 43,1 %.

Soit 5 % de chance que son score final se situe hors de cette fourchette.

Une première série de questions se pose alors à propos de nos lots. Sont-ils aléatoires ? Sont-ils représentatifs ? Dans quelles mesures ?

2) Les lots ne sont ni aléatoires, ni représentatifs

Un tirage aléatoire implique un hasard absolu. Si l'on mélangeait bien tous les bulletins, comme on peut le faire avec les boules du loto, et qu'on en tirait un lot au hasard, ce lot serait aléatoire. Sa représentativité ne dépendrait alors que de sa taille. Or, différentes particularités de notre scrutin font que nos lots ne sont ni forcément aléatoires, ni forcément représentatifs.

a) Le déroulement du dépouillement

Les témoignages des adhérents ayant dépouillé les bulletins font état de nombreuses manipulations intempestives (cf. rapport sur les témoignages). Il y a eu des transferts de bulletins entre lots, des possibilités de transferts de lots entre lettres. Il apparaît que plusieurs lots de bulletins résultent de classements divers², de décomptes de voix par demi bulletins pouvant être à l'origine d'assez grosses erreurs³, ou de jours d'arrivée différents et faisant alors

² "Une technique faisant l'objet de plusieurs témoignages a consisté à faire, à partir du même « lot », deux paquets de bulletins, l'un contenant les 31 premiers candidats de la liste ; l'autre contenant les 31 derniers. La justification de cette dérogation à la procédure est le gain de temps réalisé pour dépouiller. Les lettres concernées, d'après les témoignages, sont au moins B, C et T. "

³ « Nous avons dépouillé les bulletins des candidats actifs en nous partageant le travail : du début jusqu'au milieu de la liste des 62 candidats à pointer par 2 personnes (l'une lisait et surveillait ce que l'autre notait) et du milieu jusqu'à la fin de la liste à pointer d'autres bulletins par 2 autres personnes simultanément : chaque binôme

apparaître un effet chronologique. Il est même possible que de telles conditions puissent avoir engendré des erreurs de décompte de voix. En tout cas, il en résulte que la constitution des lots n'a absolument pas été régie par les seules lois du hasard.

Tous ces dysfonctionnements ont eu pour effet de rendre des lots encore moins aléatoires, et parfois absolument non représentatifs du résultat global.

Les travaux de Michel Fenayon montrent d'ailleurs qu'en regroupant les lots dans la lettre dont ils ont été fractionnés, des atypicités s'annulent par compensation et la lettre apparaît comme beaucoup plus "normale".

b) La taille des lots

Quelle que soit la manière dont il est constitué, par tirage aléatoire ou échantillonnage par quotas, plus un lot est important, et plus il a de chance d'être représentatif. Ainsi que l'indique M. Guillemet dans son étude citée au début de cette contribution : *"L'élément important est la taille de l'échantillon et non le taux de sondage. Une enquête auprès de 1000 personnes donne des renseignements aussi précis sur la population d'un pays de 55 millions d'habitants que sur celle d'une ville de 5000 personnes"*.

M. Guillemet parle ici d'échantillon sélectionné selon la méthode des quotas, contrairement à nos lots qui sont issus d'un tirage à peine aléatoire. Pour obtenir la représentativité d'un tel échantillon, la taille de nos lots devrait donc très largement dépasser les 1000 bulletins, probablement d'une dizaine de fois. Or, aucun de nos lots, ni même de nos lettres, n'atteignent un tel nombre. Cette contrainte est d'ailleurs impossible vu que l'ensemble des votes n'atteint pas ce nombre, on peut alors tout au plus en déduire que la taille d'un échantillon réellement représentatif du résultat final devrait être très proche du total de tous nos bulletins.

On ne peut donc dire que les « échantillons » que constituent les lettres, et surtout les lots qui en sont issus, soient strictement représentatifs :

- **pour les lots, parce qu'ils n'ont pas été constitués selon les lois du hasard.**
- **pour les lettres, parce que leurs résultats se situent en fait au milieu d'une fourchette dont la probabilité de sortir n'est pas négligeable.**

Quant aux ensembles de bulletins regroupés sur quelques jours, ils subissent divers aléas qui les rendent non aléatoires. D'une part, par leur relative petite taille, et d'autre part par l'effet des tris constatés lors du dépouillement qui ont pu affecter l'équilibre entre les premiers jours et les derniers. En effet, des dépouilleurs favorables à la liste George ont pu avoir tendance à décompter en premier lieu les lots de bulletins pré triés qui allaient dans leur sens (case Harrybey cochée), ou tout aussi bien parce que ces lots étaient plus gros. Ce qui aboutit à avantager les résultats de cette liste durant les premiers jours du dépouillement. On peut alors tout à fait comprendre qu'il puisse exister d'importants décalages entre les résultats issus de

traitait son propre paquet d'enveloppes. Cette méthode nous avait été recommandée par une personne mais je ne me souviens plus qui. N'ayant pas pu terminer le pointage, nous avons inscrit tous les 4 les sous-totaux afin que le travail puisse être poursuivi par la suite. Il restait des bulletins [le soir], on a mis un maximum de notes, sur chaque paquet, finis, à moitié finis par un bout, à moitié finis par l'autre bout. »

chaque journée de dépouillement, de même qu'entre ceux regroupant des séquences de deux ou trois jours.

Expliquer par une fraude le basculement constaté durant le dépouillement de notre scrutin n'est donc qu'une hypothèse qui a l'avantage d'être simple, mais qui n'est pas pleinement probante.

Les chiffres qui suivent contribuent à illustrer cette affirmation.

3) Analyse des tableaux des lettres

a) Les "fourchettes" d'écart

Ainsi que l'explique M. Guillemet, "le schéma d'estimation d'un échantillon aléatoire peut être formalisé de manière simple". "On peut alors associer à chaque pourcentage obtenu une "fourchette" dans laquelle le véritable pourcentage a des chances de s'inscrire".

En fonction des indications techniques fournies dans l'étude, j'ai donc calculé les "fourchettes" pour chaque lettre et lots. Les différents tableaux Excel (lettres ou lots ⁴) sont consultables à partir de ce lien ⁵: <http://m-lasserre.com/attac/elections.htm>

La "fourchette" utilisée correspond à une estimation de 95 % . C'est à dire que le score final du candidat a 95 % de chances d'entrer dans la "fourchette". Il y a donc 2,5% de chance pour qu'il soit au dessus de la fourchette, et autant pour qu'il soit au dessous.

Le tableau suivant récapitule les constatations faites à partir du tableur.

Lettre	Score final hors fourchette	Score final au dessous de la borne inférieure de la fourchette		Score final au dessus de la borne supérieure de la fourchette		Présence de "lots atypiques"
		Total	dont candidats George	Total	dont candidats George	
A	1	0	0	1	0	
B	25	18	3	7	7	B3, B4
C	10	4	2	6	1	C1
D	1	0	0	1	0	D4
E	5	0	0	5	1	
F	17	5	3	12	0	
G	1	0	0	1	0	G3, G4
H	20	1	0	19	2	
IJK	8	0	0	8	0	
J	9	5	1	4	2	
L	7	6	0	1	0	L4

⁴ Si je n'ai pas utilisé dans cette contribution le tableau sur les lots, j'ai travaillé dessus à la manière dont le fait Michel Fenayon à partir d'autres méthodes de test. Les résultats obtenus confirmaient globalement ceux obtenus par M. Fenayon.

⁵ Il est possible que le lien ne fonctionne pas en cliquant dessus, on peut toujours alors le copier dans son navigateur.

M	15	7	2	8	0	
NO	3	3	1	0	0	
PQ	6	2	1	4	1	P2
R	8	3	1	5	5	R2
S	22	7	7	15	0	
TU	33	17	0	16	14	T3
V	3	0	0	3	2	V1
WYZ	4	3	2	1	0	
Moy	10,4	4,3	1,2	6,2	1,8	

La colonne "Score final hors fourchette" indique le nombre de candidats dont le score final sort de la fourchette de 95 %. Les colonnes suivantes indiquent les nombres de candidats dont le score final est soit en dessous, soit en dessus de la fourchette, avec des sous colonnes concernant les candidats de la liste George. Un score final au dessous de la borne inférieure de la fourchette correspond à un très bon score du candidat dans cette lettre, et dont la probabilité serait inférieure à 2,5 % si toutefois les ensembles de bulletins correspondants aux lettres étaient considérés comme représentatifs.

On constate que pas une lettre ne sort totalement indemne de ce test. Le nombre de scores final sortant de la fourchette varie de 1 à 33 suivant les lettres, avec une moyenne de 10,4.

Cet important décalage n'a rien d'étonnant puisque les lettres ne sont pas totalement représentatives du fait de leur petite taille.

b) Les lettres les plus divergentes

Toutefois, ce tableau montre qu'il existe différents degrés d'écarts, et que certaines lettres sont plus ou moins favorables à la liste Georges que d'autres. Les lettres qui divergent le plus sont dans l'ordre les lettres TU, B, S, H, F et M. Les lots présentés comme atypiques par les experts partisans de l'hypothèse de la fraude, et supérieurs à 50 bulletins, sont B3, B4, C1, D4, G3, G4, L4, P2, R2, T3 et V1.

On constate alors que :

- seules les lettres TU et B sont signalées comme contenant des lots atypiques
- les lettres S, H, F et M, pourtant fortement "hors fourchette", ne contiennent aucun lots atypiques
- les lots "atypiques" C1, D4, G3, G4, L4, P2, R2 et V1 appartiennent à des lettres dont les nombres de scores hors fourchette sont parmi les plus faibles.

En dehors de B3, B4 et T3, dont on a vu précédemment qu'ils pouvaient peut-être résulter d'erreurs de décompte lors du dépouillement, la répartition des lots présentés comme atypiques ne correspond donc pas aux résultats obtenus par le "test des fourchettes" pour les lettres. On est alors face à un important paradoxe, qui infirme la thèse selon laquelle les lots atypiques résulteraient d'une fraude.

Comment se fait-il que des lots atypiques apparaissent dans des lettres "normales" ?

Comment se fait-il que des lettres fortement divergentes ne contiennent aucun lot atypique ?

La principale réponse à ces questions se trouve dans l'insuffisance de représentativité inhérente à nos lettres. Nous avons vu précédemment que la principale cause de cette faible représentativité est la petite taille des ensembles de bulletins.

3) Pour conclure

L'analyse des tableaux de lettres confirme donc la première partie de cette contribution. Leur petite taille, ainsi que leur origine plus ou moins aléatoire, font qu'aucune d'entre elles ne peut être considérée comme véritablement représentative du résultat final. Il ne peut a fortiori qu'en être de même pour les lots de bulletins issus d'un fractionnement de ces lettres. **Cette faible représentativité montre que la plus grande prudence aurait été nécessaire avant de conclure à une fraude.**

Cette contribution démontre seulement l'absence de fiabilité des études statistiques construites sur des hypothèses douteuses. Elle apporte une autre explication que la fraude pour justifier le basculement observé lors du dépouillement, sans pour autant prouver l'absence d'une fraude. Seul un examen physique des bulletins pourra trancher entre les deux explications.

Méfions nous tous de tomber dans le travers décrit par Andrew Lang⁶ en parlant d'un de ses collègues : "Il utilise les statistiques comme l'ivrogne les lampadaires : pour s'appuyer plutôt que pour s'éclairer".

⁶ Andrew Lang, homme de lettre britannique du XIX^e siècle. Cité par Michel Guillemet dans son étude : "*Limites et puissance des sondages*".