
Votes et statistiques

ANALYSE DES VOTES DES CANDIDATS “MEMBRES ACTIFS” À L’AG D’ATTAC DES 17 ET 18 JUIN 2006

Auteur : Christophe OSSWALD
Osswald.Christophe@wanadoo.fr

Table des matières

| | |
|---|---|
| Quelques rappels de statistiques | 1 |
| Pourquoi se limiter à S1 et S2 ? | 1 |
| Hypothèse 1 : les candidats sont décorrélés | 2 |
| Hypothèse 2 : les listes affichées ont bien fonctionné | 2 |
| Hypothèse 3 : il y a bien deux listes, mais pas exactement celles annoncées | 2 |
| Hypothèse 4 : en fait, il y aurait trois listes informelles ? | 3 |
| De l’ordre alphabétique | 3 |
| Choisir deux urnes parmi cinq | 3 |
| Revoter ou ne pas revoter ? Là est la question | 4 |
| Deux ou trois “clans” dans Attac ? | 5 |
| Pourquoi tant de femmes au CA ? Un biais de la méthode ? | 5 |
| Références | 5 |

Deux jours après l'AG, il est certain que cette analyse n'est pas dépassionnée. L'auteur est humain. Les calculs y sont honnêtes¹, et j'espère que ce document est accessible à la plupart des non-statisticiens.

Il éclaire quelques questions que nous sommes nombreux à nous poser sur les "accidents statistiques" depuis ce maudit samedi 17 juin 2006, même sur les accidents qui n'ont pas été perçus comme graves, et tâche de tordre le cou à certaines remarques et analyses faites encore un peu plus dans l'émotion.

Les considérations plus politiques que mathématiques sont concentrées dans la section "Revoter ou ne pas revoter".

Quelques rappels de statistiques

En statistiques, pour déterminer s'il est crédible que deux échantillons proviennent d'une même population, on utilise le test du χ^2 (lire "kideu").

Le principe est le suivant : on détermine un état théorique de la population, et on mesure les écarts des échantillons disponibles avec cet état théorique, selon la distance dite du χ^2 . Notre échantillon est constitué des valeurs x_1, x_2, \dots, x_{62} . Nos valeurs théoriques sont t_1, t_2, \dots, t_{62} . La distance est :

$$d_{\chi^2}(x, t) = \sum_{i=1}^{62} \frac{(t_i - x_i)^2}{t_i}$$

Dans notre cas, la seule estimation que nous ayons de la population (les votants d'Attac) est notre paire d'échantillons, S1 et S2. Nos valeurs théoriques seront donc construites en faisant la moyenne pondérée des échantillons : comment auraient été répartis les votes si les proportions avaient été les mêmes dans les deux échantillons, en gardant le même nombre de voix au total pour chaque candidat.

On additionne les distances d_{χ^2} de chaque échantillon à ses valeurs théoriques, pour obtenir une valeur χ^2 . Plus elle est grande, et moins il est vraisemblable que les échantillons proviennent d'une même population.

Comme nous avons deux échantillons, et 62 mesures, notre système a 61 *degrés de liberté*. Les travaux sur le χ^2 nous annoncent que :

- Si les deux échantillons ont un χ^2 supérieur à 80.23, on a moins de 5% de chances de se tromper en affirmant que les deux échantillons proviennent de populations différentes.
- Si les deux échantillons ont un χ^2 supérieur à 89.6, on a moins de 1% de chances de se tromper en affirmant que les deux échantillons proviennent de populations différentes.
- Si les deux échantillons ont un χ^2 supérieur à 100.9, on a moins d'un chance sur mille

de chances de se tromper en affirmant que les deux échantillons proviennent de populations différentes.

- Si les deux échantillons ont un χ^2 supérieur à 128.5, on a moins d'un chance sur un million de chances de se tromper en affirmant que les deux échantillons proviennent de populations différentes.

Selon la valeur χ^2 obtenue, et les risques d'erreur que l'on est prêt à assumer, on décidera si oui ou non on considère que les échantillon proviennent de populations différentes. Les populations peuvent être différentes parce que :

- L'un des échantillons est truqué et pas l'autre.
- Les deux échantillons sont truqués, mais de manière différente.
- L'hypothèse selon laquelle l'ordre alphabétique n'influe pas ... est fausse (cf. Jacques Testard).

Dans tous les cas, le test ne peut pas nous dire quel échantillon est truqué, et quel échantillon est honnête. Il nous donnera juste une mesure du risque que l'on prend en affirmant que les échantillons sont *significativement* différents. Il ne nous donnera pas la nature de la différence.

Il y a plus gênant. Le test fonctionne bien si tous les éléments de l'échantillon sont indépendants, ce qui n'est pas notre cas. Nos éléments sont "un vote pour X", "un vote pour Z", mais sur un même bulletin, il y a plusieurs votes. Nos éléments indépendants sont les bulletins, pas les votes par candidats.

C'est en essayant de rendre la moins fausse possible cette condition de non-corrélation que peuvent se construire et se déconstruire les quatre hypothèses qui suivent.

Et les rangs dans tout ça ?

Tous ces éléments statistiques portent sur le *nombre de voix*. Ainsi, le fait que 10 voix de plus fasse gagner 0, 1, ou 7 places au classement ne change rien : les rangs n'ont pas de sens dans ce contexte, même s'ils déterminent la composition du CA.

Les chiffres placés dans les colonnes B et C du fichier joint proviennent du tableau remis – au format papier – le samedi 17 juin aux participants à l'AG. Il a été retapé manuellement, j'espère ne pas avoir fait d'erreur².

Pourquoi se limiter à S1 et S2 ?

S1 et S2 sont des bulletins qui sont arrivés en même temps, et qui ne diffèrent que par l'ordre alphabétique.

¹ce qui ne veut pas dire qu'ils soient forcément justes, malheureusement. Toutes les remarques sont les bienvenues

²Mais, justement, le test du χ^2 aide à détecter ce genre d'erreur...

Il est donc crédible qu'ils proviennent de la même population.

S3, et surtout S4, proviennent de votes effectués plus tard, dans un contexte différent, temporellement, mais aussi géographiquement et organisationnellement : voter "à la bourre" traduit une certaine sociologie. Voter encore plus à la bourre en connaissant quelqu'un qui se rend à l'AG, ou voter en étant sollicité par un camarade qui se rend à l'AG encore plus.

Une différence entre S1/S2 et S3 ou une différence entre S1/S2 et S4 n'est nullement un dangereux accident statistique, mais juste l'illustration du fait que les populations ne "semblent" pas être identiques, alors que l'on sait très bien qu'elles n'ont aucune raison de l'être. Si S1/S2 et S3 sont semblables, on pourrait simplement émettre l'hypothèse que le fait de tout remettre à la dernière minute n'influe pas sur le résultat ... dans les cas de votes Attac.

Ainsi, toute cette étude porte sur la différence entre S1 et S2.

Hypothèse 1 : les candidats sont décorrélés

On s'installe dans la situation confortable où l'on peut faire un test du χ^2 sur les données que l'on a sous la main.

Le χ^2 vaut 833.86. Il est extrêmement improbable de se tromper en affirmant que S1 et S2 sont différents (pour $\chi^2 = 193$, cette probabilité est d'une chance sur un million de milliards ... après, les valeurs numériques se calculent mal). Si les différents candidats sont décorrélés, alors, c'est sûr, il faut refaire le vote d'urgence.

Problème : on voit bien / on se doute bien que les candidats se trouvant une des listes qui ont circulé ont plus de chance de se retrouver sur un même bulletin, ou d'être écarté d'un même bulletin. Dans quelle mesure ?

Cette hypothèse correspond aux remarques 4 et 5 de Thomas Coutrot et Michel Husson : il est nettement moins surprenant que les scores de personnes d'une même "liste" varie conjointement plutôt que de façon indépendante.

Hypothèse 2 : les listes affichées ont bien fonctionné

On ajoute deux candidats fictifs :

la liste "SG", constituée des candidats cités sur la lettre de Susan. Les "cas particuliers" (R. Dumas, F. Viale, R. Tassi) et les démentis (R. Silva Jacinto) sont comptés pour un demi.

une liste "JN", constituée des candidats les plus favorisés par S2 par rapport à S1. Les 32 candidats signataires de la lettre associée sont trop nombreux pour former une liste homogène. Notons que R. Tassi se trouve aussi dans cette liste.

Ainsi, ce qui caractérise les urnes de S2 par rapport à S1, c'est qu'il y a plus de bulletins favorables à (la liste de) J. Nikonoff. Ce ne sont pas 19 candidats isolés qui par un hasard improbable auraient remonté lors de la deuxième tranche de décompte.

Ainsi, on va séparer les différentes voix entre voix "isolées" et voix "de listes", de façon à rendre la plus crédible possible l'hypothèse d'une même population.

On arrive à la répartition suivante :

- S1 : 475 bulletins SG, 430 bulletins JN.
- S2 : 220 bulletins SG, 440 bulletins JN.

Et un χ^2 de 396.31. Deux fois moins que pour l'hypothèse 1, mais toujours trop pour accepter l'hypothèse d'homogénéité de la population.

D'où vient ce χ^2 ?

- 69 points viennent de la disparité des listes. C'est beaucoup, mais ça ne suffit pas, et de loin, à rejeter l'homogénéité.
- J. Tosti, G. Jouve et F. Denimal sont sur la liste SG, mais sont représentés de la même manière dans S1 et S2 : 110 points.

A partir de ce stade, et pour les hypothèses qui suivent, les trente candidats arrivés en tête, et les listes supposées, ne concentrent plus qu'une part du χ^2 insuffisante seule à rejeter l'homogénéité.

... on construit une hypothèse 3 en réévaluant la liste SG au regard de ces 110 points. *Ceci, sans chercher à résoudre l'origine de cet effet de décorrélation, permet de passer outre la remarque 6 de l'analyse de T. Coutrot et M. Husson.*

Hypothèse 3 : il y a bien deux listes, mais pas exactement celles annoncées

On ajoute deux candidats fictifs :

la liste "SG", constituée des candidats cités sur la lettre de Susan "confirmés" par une présence plus forte dans S1 que dans S2. Les "cas particuliers" (R. Dumas, F. Viale) et les démentis (R. Silva Jacinto) sont comptés pour un demi. J. Tosti, G. Jouve et F. Denimal n'y sont pas.

une liste "JN", constituée des candidats les plus favorisés par S2 par rapport à S1. Les 32 candidats signataires de la lettre associée sont trop nombreux pour former une liste homogène. C'est la même que pour l'hypothèse 2.

On a arrive à la répartition suivante :

- S1 : 535 bulletins SG, 400 bulletins JN.
- S2 : 230 bulletins SG, 430 bulletins JN.

Et un χ^2 de 232.44. Plus de trois fois moins que pour l'hypothèse 1, mais toujours trop pour accepter l'hypothèse d'homogénéité de la population.

D'où vient ce χ^2 ?

- 87 points viennent de la disparité des listes. C'est beaucoup.
- Plusieurs candidats semblent diversement portés par l'effet Nikonoff, au sein de sa liste, alors que la liste SG ne comportait que trois "aberrations", traitées par l'hypothèse 3. Gaziello, Coignard, Barral, Jauffret sont très liés à S2 (on aimerait les mettre de 120% à 140% dans la liste) et gardent donc une contribution assez importante. Ce sont les accidents statistiques les plus forts (pointés pendant l'AG), et il est normal que la méthode ne masque pas cet effet. D'autres sont "à moitié" portés. A moins de les demi-inscrire (comme les cas particuliers de la liste SG), ils gardent un apport important.

Hypothèse 4 : en fait, il y aurait trois listes informelles ?

Si on considère trois listes :

la liste SG , comme en hypothèse 3.

une liste JN1 , pro-JN

une liste JN2 , de gens opposés à la liste SG, mais qui ne soutient pas forcément JN. Notons que cette liste ne contient pas J. Nikonoff lui-même...

La liste JN1 est essentiellement un sous-ensemble de JN2. Les candidats qui ne sont que dans JN2 ne récoltent qu'une part de l'effet de liste identifié dans l'hypothèse 2.

On a arrive à la répartition suivante :

- S1 : 550 bulletins SG, 210 bulletins JN1, 190 bulletins JN2.
- S2 : 230 bulletins SG, 240 bulletins JN1, 190 bulletins JN2.

Et un χ^2 de 171.96. Plus de quatre fois moins que pour l'hypothèse 1, trop pour accepter l'hypothèse d'homogénéité de la population : une chance sur six cents milliards d'avoir un χ^2 aussi important à partir d'une population homogène et des tirages décorrelés.

La contribution des listes au χ^2 est de 91. Beaucoup, mais pas suffisant pour refuser l'hypothèse d'homogénéité à elle seule.

Alors ?

Nous avons bien vu comment diverses hypothèses sur la relation entre listes et votes pouvait influencer sur la possibilité de hurler légitimement à la fraude. Est-ce qu'un modèle en quatre ou cinq classes nous amènerait à accepter le vote ? Est-ce qu'au-delà d'un certain nombre de classes, le χ^2 ne peut plus descendre, et 172 serait le minimum que l'on peut atteindre à partir de S1 et S2 et l'accident statistique est donc grave ?

Est-ce que l'on peut pousser l'étude sur les données disponibles, ou faut-il, pour trancher, en revenir aux bulletins ? Réponse dans la section "Corrélations".

De l'ordre alphabétique

Les bulletins ont été répartis par ordre alphabétique. On peut supposer que cet ordre alphabétique influe peu sur les votes, mais ceux qui ont feuilleté l'annuaire et/ou enseigné dans un environnement se sont sans doute aperçus que la première lettre du patronyme est corrélée à certains attributs :

Bretons : L

Arabes : A

Nobles : D

Alsaciens : M, S, Z

Si ces attributs sont corrélés au vote, alors les premières lettres sont corrélées au vote ... mais ces éventuelles corrélations ne peuvent guère se mesurer. Encore que ... les deux candidats alsaciens était sur la lettre Nikonoff, et les candidats bretons comptaient deux Nikonoff, trois George, et deux non-alignés...

Par ailleurs, si dans une même famille – au même patronyme, donc – il y a plusieurs adhérents à Attac, il y a de fortes chances pour que leurs votes soient semblables. Ainsi, les phénomènes observés sur les votes d'un candidat à l'intérieur de chaque lettre sont amplifiés. Pour prendre cet effet en compte, il faudrait connaître le nombre de patronymes différents parmi les 5600 votants de S1 et S2.

Pour plus de sécurité, il aurait été préférable de ranger les bulletins selon le *dernier chiffre du numéro d'adhérent*³. Pour le prochain vote...

Choisir deux urnes parmi cinq

Afin d'illustrer la thèse de la possibilité de fraude, la séparation en deux classes S1 et S2 des cinq urnes constituées le samedi, les urnes les plus défavorables à JN ont été considérées comme étant dans S1, les autres comme étant dans S2.

³Pour une utilisation efficace, éprouvée et intéressante des derniers chiffres d'un tirage aléatoire comme un nouveau tirage aléatoire, voir la règle de localisation des coups dans l'excellent jeu de rôle *Warhammer*

Il nous faut donc considérer que le χ^2 trouvé est le plus grand parmi toutes les 30 répartitions de cinq urnes en deux ensembles non-vides ($2^5 - 2 = 30$).

Ainsi, sous l'hypothèse 4, il y a maintenant une chance sur 20 milliards d'atteindre ce score de χ^2 , en acceptant l'hypothèse que nos trois listes absorbent toutes les corrélations entre les 62 candidats au sein de la population des votants.

Revoter ou ne pas revoter ? Là est la question

En tenant compte de ce biais de choix des urnes, il suffirait que la modélisation en listes nous fasse atteindre un χ^2 de 105 pour que ce score soit atteint ou dépassé par des échantillons homogènes avec une probabilité de 1%. A ce stade, il ne serait pas raisonnable du tout d'invoquer les statistiques pour rejeter le vote : il est bien plus probable d'aligner trois six d'affilée sur un dé.

Descendre jusqu'aux urnes

Si l'on ne considère pas deux ensembles, mais cinq urnes, on passe à un système à 244 degrés de liberté. Le seuil à une chance sur un million nous autoriserait un χ^2 à 363 au lieu de 128. Il s'agit de voir comment les listes déjà établies peuvent nous aider dans une modélisation à cinq urnes.

Ce résultat "intermédiaire" du décompte doit exister. Maintenant l'AG passée, il pourrait être rendu public, et les calculs être faits rapidement.

Les vraies listes induites par les bulletins : calculer les corrélations

Pour les hypothèses 3 et 4, nous avons extrapolé des listes informelles pour expliquer au mieux les variations. Rien ne dit que ces listes aient un sens au niveau des bulletins. Pourtant, ce sont bien les bulletins qui représentent – seuls – la volonté des adhérents. C'est donc à partir des bulletins, et non des sommes de voix, qu'il faut chercher les listes de candidats qui ont été "utilisées" par les électeurs.

Comment aller jusqu'à ces listes ? Les résultats des décomptes par ensemble (S1 ou S2) ne sont pas utiles, pas plus que les résultats par urne.

Il faut calculer les corrélations entre candidats à partir de leurs présence ou leur absence commune d'un bulletin, et donc informatiser chacun des 4960 bulletins de S1+S2.

A partir de là, on pourra faire une modélisation des voix entre effets de listes (induites par des conseils de vote, ou par des candidatures perçues comme compatibles les unes avec les autres par des votants qui ne se sont pas concertés) et votes qui rentrent dans le modèle des événements décorrélés.

Si nous choisissons d'aller jusque là, il serait raisonnable de mettre à contribution les CL éloignés, en leur envoyant quelques centaines de photocopies de bulletins, et ne pas laisser toute la charge aux CL parisiens.

Une alternative serait de considérer comme classes théoriques toutes les façons de choisir 24 candidats ou moins parmi une liste de 62. Problème : il y a plus de 200 millions de milliards de telles possibilités. Comme une classe théorique, dans le cadre de la règle de χ^2 , ne peut contenir moins de cinq éléments, ce n'est pas réalisable.

Par contre, il est possible de regrouper les bulletins par une règle de proximité (classification ascendante hiérarchique munie de la règle du lien maximum, par exemple, pour constituer des ensembles de bulletins autour de noyaux interprétables).

La poursuite de l'analyse, en tout état de cause, ne pourra pas se faire sans informatiser chaque bulletin.

Vers un χ^2 limite ?

Personnellement, j'ai le sentiment que pousser l'étude jusqu'aux corrélations nous permettra d'accepter l'hypothèse d'homogénéité, avec une probabilité certes faible (un pour cent à un pour mille) mais trop élevée pour rejeter la possibilité du manque de bol dans une élection honnête.

Mon opinion, après les premiers calculs de χ^2 menés sur une calculatrice à quatre fonctions en amphi de l'AG, n'allait pas dans ce sens. Mais la technique et le choc m'ont cantonné, à ce moment-là, dans l'hypothèse 1 ; quantité d'effets surprenants peuvent en fait s'expliquer assez bien au sein des données, en sortant de cette hypothèse intuitive mais classique.

Coûts politiques

Ils sont déjà élevés : destruction dans les médias (et cette fois, les journalistes ont été honnêtes : nous leur avons tendu le bâton et le dos pour nous faire battre), CA largement malmené, bureau mal élu, direction globalement illégitime, adhérents en desherance.

Les promesses de parité à base de désistement sont oubliées, alors que le déséquilibre est flagrant.

Des élus se sont assis trois fois sur le vote des adhérents en moins de 24 heures : refus siéger alors qu'il sont élus pour cela ; refus de siéger alors qu'une motion majoritaire le demande ; refus de sortir d'un amphithéâtre alors qu'une motion très majoritaire le demande.

Le prix payé est déjà important, en interne comme vis-à-vis de l'extérieur.

Je crois que ces coûts politiques seront plus lourds si nous conservons un demi-CA que si nous relançons des élections, et qu'il seront plus lourds si nous relançons une élection plutôt que pratiquer une réconciliation de

façade, le temps d'organiser "sereinement" des élections anticipées.

Je pense que cette analyse préliminaire devrait nous encourager à faire fonctionner l'association dans l'apparence de la sérénité ces prochains mois.

Deux ou trois "clans" dans Attac ?

P. Khalifa a signalé – longuement – que 66% des votants avaient approuvé la liste des fondateurs. Il faut aussi noter que 62% des votants ont choisi J. Nikonoff : entre 28% et 62% des votants ont donc choisi *à la fois* la liste des fondateurs et J. Nikonoff.

Un gros tiers de l'association, au moins, ne souhaite donc qu'une chose : voir nos "élites" se réconcilier. Il se retrouvent parmi 4300 à 4600 bulletins que les hypothèses 2 à 4 ne parviennent pas à ranger dans un clan : ils ont dû panacher les listes. Ceux-là manquent singulièrement de relais au sein du CA... et cette analyse à partir d'effets de liste maximisant la probabilité de scrutin honnête montre même qu'il seraient plus des deux tiers des adhérents.

Pourquoi tant de femmes au CA ? Un biais de la méthode ?

Parmi les 62 candidats, il y a 21 femmes, et 41 hommes. Au final, parmi les 24 élus, il y a 17 femmes

et 7 hommes. Félicitation aux élues ? Bien sûr ... mais voyons quand même quelques calculs.

Si chaque votant, ayant fait sien l'esprit de la parité, remplit son bulletin avec 12 femmes et 12 hommes : chaque homme reçoit, en moyenne, $12/41 = 29\%$ des voix ; chaque femme reçoit, en moyenne, $12/21 = 57\%$ des voix. Pour peu que les candidatures soit "équilibrées", le CA se retrouve avec 21 femmes et 3 hommes.

Il apparaît que les votants n'ont pas tous respecté la parité ... mais ce biais semble malgré tout avoir joué un rôle important.

Pour la prochaine édition d'un système à désistements, il faudra songer à assurer la parité *parmi les candidats* et pas seulement parmi les élus. Pour cette élection-ci, il faudra sans doute faire un rappel de promesses de démission ... promesses faites à une époque où l'on croyait que la différence de sexe était politiquement plus importante que la différence de "clan".

Références

Analyse proposée par la liste "Susan George" lors de la proclamation des résultats.

Analyse proposée par Michel Husson et Thomas Coutrot (mail du 20/06, <http://www.attac-toulouse.ouvaton.org/forum/viewtopic.php?p=1941#1941>)

à titre personnel,
Christophe Osswald, Attac Brest.