

Écarts de voix et classement au soir du 11 Juin : un modèle « anti-hasard » Michel Husson, 11 juillet 2006

Le modèle cherche à tester la dépendance entre :

- l'écart, noté **E**, entre le nombre de voix de chaque candidat et son nombre de voix hors lots atypiques.
- le rang qu'il obtenait dans le dépouillement intermédiaire du 11 Juin.

Le modèle fait deux hypothèses :

- a) que l'écart est négatif pour les candidats de la liste de Susan George, positif pour les autres ;
- b) que l'écart est fort pour les candidats arrivés dans la première moitié du classement du 11 Juin, et faible pour les autres.

Le modèle introduit donc deux variables indicatrices.

indilist est un indicateur de liste qui vaut :

1 si le candidat appartient à la liste de Susan George ;
0 sinon.

indirang est un indicateur de rang qui vaut :

1 si le candidat est arrivé parmi les 31 premiers du classement établi dans le cumul du 11 Juin ;
0 s'il est arrivé parmi les 31 derniers.

Le modèle s'écrit donc :

$$E = A [(1-2.\text{indilist}).\text{indirang}] + B (1-2.\text{indilist})$$

Cela veut dire que l'écart de voix vaudra $-(A+B)$ pour les candidats de la liste de Susan George arrivés en tête et $(A+B)$ pour les autres candidats. Ces écarts seront respectivement de $-B$ et $+B$ pour les candidats arrivés dans la seconde moitié du peloton. Le test de ce modèle a été réalisé avec *Eviews4.1* (voir <http://hussonet.free.fr/rangdep.xls> pour les données et <http://hussonet.free.fr/rangdep.wfl> pour la feuille de calcul Eviews). Il donne les résultats suivants :

Dependent Variable: E
Method: Least Squares
Date: 07/11/06 Time: 16:10
Sample: 1 62
Included observations: 62

Variable	Coefficient	Std. Error	t-Statistic	Prob.
A	88.79554	10.02287	8.859292	0.0000
B	27.18669	6.854808	3.966076	0.0002
R-squared	0.816593	Mean dependent var		2.049387
Adjusted R-squared	0.813537	S.D. dependent var		91.19174
S.E. of regression	39.37788	Akaike info criterion		10.21601
Sum squared resid	93037.02	Schwarz criterion		10.28463
Log likelihood	-314.6964	Durbin-Watson stat		2.041774

Les *t* de Student (*t-Statistic*) sont très significatifs (largement supérieurs à 2), ainsi que la coefficient de corrélation (*R-squared*). Bref, le modèle est validé, comme le montre aussi le graphique ci-dessous qui permet même de repérer quelques hésitations sur la frontière entre candidats *in* et candidats *out*.

On a donc vérifié ce fait majeur dont toute explication des anomalies doit rendre compte : **la distribution des écarts de voix engendrés par les lots atypiques dépend du classement au soir du 11 Juin.**

