

# PÉDAGOGIE STATISTIQUE DES SONDAGES ÉLECTORAUX

Benoît Riandey <sup>1</sup>

<sup>1</sup> *Ined, 133 Bd Davout, F75980 Paris Cedex 20, riandey@ined.fr*

Les citoyens français ont des raisons d'être troublés lorsqu'en 24 heures deux instituts privés de sondage publient deux estimations d'intentions de vote divergentes tant en niveau qu'en évolution. Les enquêtes sont-elles en cause? La taille limitée des échantillons? Des méthodologies divergentes? Notre profession des statisticiens a-t-elle apporté les éléments utiles pour lever ces doutes? Non et archi-Non.

Le *Que sais-je?* d'Anne-Marie Dussaix et Jean-Marie Grosbras et l'annexe pédagogique de Pascal Ardilly au rapport du Sénat ont clairement présenté la notion d'intervalle de confiance mais les questions posées ne se ramènent pas à l'incertitude d'échantillonnage des estimations d'intentions de vote. Nous devons amplifier notre contenu pédagogique.

Ce qu'en disent les instituts de sondage se limite à l'affichage sommaire d'une précision de 2 à 3 % pour leurs estimations. Or la signification statistique de cette même appréciation mériterait d'être explicitée et pourrait être approfondie pour illustrer la meilleure précision des estimations présidentielles de second tour par post-stratification sur la répartition des votes au premier tour. Par ailleurs, en corrigeant des biais de non réponse, le redressement de l'échantillon différencie les pondérations à la manière de probabilités inégales non optimisées. La variance des estimations en est accrue en toute ignorance. Au premier tour, l'effet de réduction des biais de réponse l'emporte certainement sur le bénéfice de la post-stratification de l'échantillon, contrairement au second tour. Mais nous n'avons pas connaissance qu'un statisticien universitaire français se soit intéressé à cette question en accédant aux données individuelles des sondages d'intentions de vote. Mais peut-être le fut-ce sous couvert du fameux "secret de fabrication".

Mais l'intervalle de confiance attaché à ces estimations d'intentions de vote ne répond pas aux questions que se posent les citoyens. Il faut distinguer les questions suivantes et apporter à chacune sa justification statistique propre :

- 1- Quelle est aujourd'hui l'intention de vote pour le candidat X?
- 2- Le candidat X devance-t-il aujourd'hui le candidat Y?
- 3- Le candidat X a-t-il progressé depuis la précédente enquête?
- 4- Les enquêtes A et B fournissent-elles des résultats statistiquement compatibles?
- 4bis- L'enquête finale A fournit-elle des résultats statistiquement compatibles avec le résultat du vote?

## Bibliographie

Riandey, B. (2012). Sondages électoraux : estimations ou inférences ?, Journal de l'APMEP.