

La formation de l'opinion

Le mirage du sondeur

A.G.HATHOUT

2 décembre 2011

Ce matin du 2 décembre 2011, j'ai discuté au téléphone pendant plus d'une heure avec un ami de Saint-Raphaël sur les problèmes des temps modernes, et plus particulièrement sur le problème des minorités. Nous avons convenu que l'intolérance et la bêtise se développent sur le terrain de l'ignorance et de l'inconscience. A la fin de mon entretien téléphonique, j'ai pris mon clavier et j'ai rédigé ce texte.

La société qui compose un pays est toujours susceptible d'être considérée,- à des fins de recherche,- par référence à des caractéristiques que le langage statistique appelle aussi *critère*. Ainsi, la tranche d'âge, le sexe, le niveau éducatif, le patrimoine financier, la région de résidence, ou à la sensibilité religieuse sont autant de *critères* qui nous aident à cerner la spécificité de l'individu dans la société.

Pour embrasser la société de la manière la plus globale qui soit, au moyen d'un petit nombre d'individus, les sondages procèdent à la construction de qui est appelé « strates », autrement dit, des croisements de critères. Quant à l'acte par lequel les strates sont construites, il est désigné par « stratification ».

Par exemple une stratification qui se réfère à 5 niveaux éducatifs et à 8 régions de résidence nous donnera $5 \times 8 = 40$ strates.

L'opinion que nous nous faisons sur les strates est alors sujette à ce qu'il m'a paru convenable d'appeler le « *mirage du sondeur* ». Par cette locution j'entends une inclinaison de l'entendement humain à se tromper sur les strates avec une erreur d'autant plus élevée que l'effectif de la strate est faible.

Cette erreur se manifeste aussi bien dans le jugement porté sur un seul critère, que dans une strate résultant du croisement de plusieurs critères.

Embrasser le mécanisme mental qui est sous-tendu par le mirage du sondeur, l'illustrer, et le simuler c'est dompter l'inconscient qui risque de nous conduire à de dangereux rivages. C'est pourquoi, raisonner sur des situations qui ne laissent pas la place à l'émotion garantit la compréhension de ce problème de mémoire.

Voici, schématiquement, une situation qui montre comment les choses se passent :

- Appelons « population mère » la composition d'un un sac de riz et d'orge où on a compté une proportion de 1000 grains de riz pour 10 grains d'orge. Etant donné que ce sac contient 100 fois plus de riz que d'orge, on peut dire de l'orge qu'il est ici « minoritaire ».
- Supposons, par construction, que les grains d'orge et de riz ont les mêmes caractéristiques statistiques : même poids moyen et même écart moyen à la moyenne (écart-type).
- Appelons « échantillon » une poignée qui représente 10% de la population mère (celle-ci joue alors le rôle de ce qu'on appelle *urne*). Il est vraisemblable que cette poignée va avoir les mêmes proportions en riz et en orge, soit 100 fois plus de riz que d'orge.

Si maintenant nous calculons le poids moyen des grains de riz et d'orge, à partir de la poignée (échantillon), l'erreur commise dans l'évaluation du poids moyen de chacune de ces deux céréales sera plus importante sur l'orge minoritaire que sur le riz majoritaire. C'est ce que nous pourrions appeler une « a-perception », ou comme on le dit en américain, une « *misperception* ».

Leçon pour le sondage

Un des buts des sondages statistiques est de déjouer l'a-perception que nous venons de voir. Ceci s'effectue comme suit : une fois les strates du sondage établies, la théorie des sondages recommande de faire des relevés (de piocher) dans la population mère avec un pourcentage variable selon les strates : plus l'effectif de la strate est faible ; plus le taux de sondage doit être élevé.

Autrement dit, plus la strate est *minoritaire*, plus nous devons déployer des efforts (de sondage) pour ne pas s'y tromper. Ainsi, dans le cas de l'orge et du blé, l'orge étant *minoritaire* relativement au blé, le « taux de sondage » dans le l'orge aurait dû être supérieur à celui du blé, et nous n'aurions pas dû prendre une « poignée » de manière indifférenciée.

Semblablement, dans un sondage sur l'économie française où les strates de départ ont été celles de la nomenclature de l'INSEE dite APE 40B (en 38 postes), le textile étant *minoritaire* relativement à l'hôtellerie, le taux de sondage doit être plus important dans le textile que dans l'hôtellerie. Et ce, la métaphore est ici permise, de crainte que nous ne soyons « racistes » à l'égard du textile.

Simulation

Afin de visualiser « le mirage du sondeur », j'ai rédigé un programme qui simule une population composée de deux catégories d'éléments qui ont exactement les mêmes caractéristiques statistiques : même poids moyen (50.5) et même écart-type. La seule différence est qu'une catégorie est 6 fois plus nombreuse que l'autre.

Après avoir simulé 30 piochages à la proportionnelle à raison de 10%, j'ai représenté l'erreur commise sur le poids des éléments de chacune des deux catégories, visualisant ainsi le mirage du sondeur, autrement dit, l'erreur de celui qui sonde à la proportionnelle.

Dans le graphique ci-après, le poids moyen ressorti pour la catégorie des petits effectifs est représenté par « p » (comme **p**etit) et celui des grands effectifs est représenté par « g » (comme

grand). Manifestement, le sondage est proche du poids réel (50.5) pour la catégorie « g », alors qu'il s'en écarte grandement pour la catégorie « p » !

Au-delà des simulations, cela signifie, qu'en de nombreuses circonstances où nous parlons de « racisme », il existe un défaut oculaire qui ne porte pas de nom dans la langue et qui n'est pas ce que la langue entend quand elle dit « racisme ». Ce mot est seulement une incapacité mécanique de la mémoire humaine à se faire une juste idée des personnes qui appartiennent à des catégories minoritaires, quel que soit le type de catégorisation effectuée, quel que soit la période étudiée dans l'histoire comme dans les temps présents.

Une politique éducative qui accorde plus de place aux applications informatiques de la statistique devrait améliorer la coexistence pacifique des majorités et des minorités et compléter notre connaissance de nous-mêmes.

Visualisation du mirage du sondeur à la proportionnelle

Légende : « p » comme « petit » désigne les entités minoritaires et « g » comme « grand » désigne les entités majoritaires.

