|  |
| --- |
| AGENIS-NEVERS Marc |
| Publication : 02/11/2017 |
| Version : 1.1  Repository Gitlab et scripts : gitlab.com/users/agenis/projects |

Influence de la taille des circonscriptions sur le résultat d’une élection législative : l’exemple de juin 2017

Contexte

Cette étude se focalise sur la façon dont les caractéristiques des circonscriptions, spécifiquement, et plus généralement le mode de scrutin, influencent la composition de l'Assemblée Nationale dans l’élection législative française. L’étude se base sur les données électorales complètes du Ministère de l’Intérieur publiées en *open data*, correspondant aux résultats du 1er et 2nd tour des élections de 2017 en France.

On montrera notamment comment l’élection des 577 députés et les scores obtenus par leurs partis reposent sur des mécanismes électoraux non triviaux, dont les justifications philosophiques sous-jacentes sont rarement discutées ou énoncées. Malgré quelques débats médiatisés sur la pertinence de la proportionnelle, les premiers retours que j’ai eus sur l’application *Shiny* m’ont montré que ce sujet restait technique et mal maîtrisé.

**Le présent travail expose quelques statistiques descriptives d’intérêt, puis le développement d’un processus statistique de *clustering* des électeurs afin d’étudier l’effet « taille de circonscription », enfin une application web de simulation de l’élection.**

Un peu de philosophie politique

Fonctionnement de l’élection législative de nos jours

Un rappel sur le mécanisme de l’élection législative n’est pas inutile : elle se fait au **scrutin majoritaire uninominal à deux tours** (SMU2T) depuis le début de la 5ème République (sauf une seule fois en mars 1986 où la proportionnelle intégrale à 1 tour avec un scrutin de liste avait été utilisée). Le deuxième tour départage les candidats ayant obtenu plus de 12.5% des inscrits, et le candidat élu au 2nd tour devient député. De nombreuses particularités existent dans le code électoral, en voici quelques-unes :

* Un parti peut présenter plusieurs candidats dans la même circonscription
* Un candidat peut être élu au premier tour s’il a la majorité absolue ET 25% des inscrits
* Si plusieurs candidats dépassent 12.5%, on peut avoir des « triangulaires »… voire même des « heptangulaires », 7 candidats au 2ème tour est une situation théoriquement possible (même si historiquement on a pu voir des quadrangulaires uniquement, voire des pentangulaires dans le cas des élections municipales). Ce phénomène reste très dépendant du taux de participation et de la multipolarisation du paysage politique.
* En cas d’égalité parfaite des scores au 2nd tour, l’élu est le candidat le plus *vieux*.

Le mot **majoritaire** est ici bien à comprendre comme l’opposé de « proportionnel », et sont aussi différents que le ne sont les notions de « moyenne » et de « maximum ». Un scrutin **proportionnel** affecte un nombre de députés strictement proportionnel au score de chaque parti (ou liste). Précisons enfin que le mode d’élection de nos députés (SMU2T) est plutôt rare **dans le monde** (12% de tous les systèmes électoraux en 1995, la famille des scrutins majoritaires représentant environ 50% des cas), le système proportionnel est lui utilisé dans 42% des cas (source : <http://aceproject.org/main/francais/es/es60.htm>)

Les circonscriptions

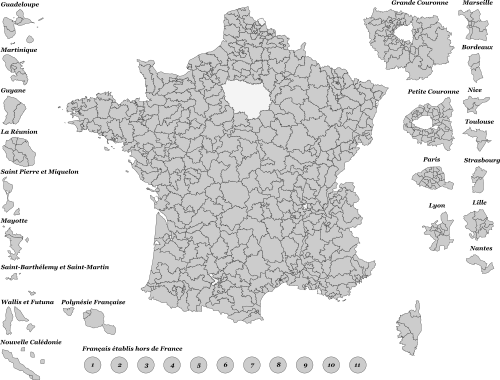
Le nombre de circonscriptions n’a pas toujours été constant, toutefois sous la 5ème République la bijection reste valide. Ainsi le nombre de circonscriptions et de députés augmente avec 5 modifications au total, principalement pour coller à l’accroissement non isotrope de la population. Toutefois des redécoupages sont régulièrement opérés, sans altérer le nombre total. La règle veut qu’une circonscription ne puisse pas être à cheval sur 2 départements et que leur taille (en nombre d’habitants) soit homogène avec une marge de +/-20% autour de la moyenne (81 500 électeurs aujourd’hui).

Figure : circonscriptions françaises de 2017

Philosophie de la représentation

La notion de circonscription pose la question de **la fonction symbolique du député** ; la croyance populaire étant que le député représenterait les intérêts de sa circonscription dans le processus législatif. C’est hélas très loin d’être le cas : ainsi dans un décret de 1999 (n° 99-410DC), le Conseil Constitutionnel a rappelé que « si députés et sénateurs sont élus au suffrage universel (…), chacun d'eux **représente au Parlement la Nation tout entière et non la population de sa circonscription** d'élection ». Toutefois cette notion a fluctué à travers les textes et les âges si bien sur les juristes peinent à se mettre d’accord (voir le [mémoire](http://www.droitconstitutionnel.org/congresNancy/comN9/traversacTD9.pdf) de Anne-Sophie Traversac à ce sujet). Cette question est moins présente dans les élections à caractère proportionnel puisqu’il s’agit d’un scrutin de liste, même si le mécanisme proportionnel n’empêche pas la constitution de circonscriptions géographiques, qui sont généralement de taille plus importante comme pour les élections européennes (8 circonscriptions pour toute la France) ou législatives de 1986 (1 circonscription par département).

1 tour ou 2 tours ?

Le choix philosophique derrière la règle du nombre de tours d’une élection est important :

* Une élection à 2 tours implique que l’élu doive nécessairement représenter la majorité absolue des suffrages. Toutefois, si on appliquait strictement ce principe, il serait logique de requérir la majorité des *inscrits* et pas seulement des *exprimés*, donc recommencer l’élection le cas échéant. Pour éviter ce blocage on doit faire l’hypothèse qu’une abstention a pour signification une acceptation du candidat majoritaire quel qu’il soit, ce qui est loin d’être acquis en pratique, d’où les débats récurrents sur la **légitimité** des élus. Ce type de scrutin reste fortement contaminé par des stratégies de « vote utile » du côté de l’électeur ou plus récemment de « désistement de barrage » du côté des candidats, ces tactiques pouvant s’avérer plus ou moins contre-productives ou frustrantes pour l’électeur. Par ailleurs, les candidats sont encouragés à ne conclure des alliances qu’après le 1er tour, une fois leur poids électoral officiellement mesuré ; c’est-à-dire s’allier avec des concurrents qu’ils ont pu vivement critiquer lors de la campagne du premier tour, ce qui est délétère à la fois pour l’image des politiques et pour les programme électoraux qui subissent des remaniements d’urgence quelques jours avant le second tour.
* Une élection à 1 tour (ex : Royaume Uni) implique au contraire que le citoyen doive accepter d’être représenté un candidat souvent minoritaire. Dans ce scrutin, les candidats sont encouragés à s’allier dès la campagne du 1er tour et à préparer un programme commun.

Majoritaire ou proportionnel ?

Le scrutin majoritaire est, fondamentalement, un système dit « winner-takes-all », encore plus accentué en cas de « prime majoritaire » (qui existe pour les régionales) : il n’y a **pas de place pour les vaincus dans la représentation de la nation**, on est dans une philosophie de la compétition et de l’opposition plutôt que de la collaboration. Pourtant, il peut sembler paradoxal de vouloir à la fois un système majoritaire (textuellement qui ne représente pas les courants minoritaires) et à la fois demander au candidat, une fois élu, de « représenter la Nation toute entière » dans toutes ses constituantes, donc même ses anciens adversaires : l’élu devrait-il se départir de ses propres proximités idéologiques, ou de la formation politique qui l’a mis au pouvoir ?

C’est peut-être là le point le plus important dans ce qui va suivre, car si le système proportionnel permet de représenter tous les courants politiques, le système majoritaire favorise, lui, les mouvements politiques plus « consensuels » et « centristes ». Moralement, on peut considérer comme **un défaut ou un bienfait** le fait de favoriser les idées consensuelles au détriment des idées clivantes ou peu répandues, et ce n’est pas l’objet de la discussion ici.

|  |  |  |
| --- | --- | --- |
| Effets | 1 TOUR | 2 TOURS |
| MAJORITAIRE | - Coupure des idées clivantes | - Coupure des idées clivantes  - Sur-représentation des idées consensuelles |
| PROPORTIONNEL |  | Non défini |

La simplicité de la proportionnelle dans son principe se heurte toutefois à la complexité des mécanismes arithmétiques sous-jacents permettant sa mise en œuvre : détermination précise de l’ordre d’attribution des sièges entre les listes ainsi que des règles qui permettent d’éviter de devoir élire des **fractions décimales** de députés (règles de Hondt, règles de Hare, etc.), qui utilisent d’ailleurs souvent un seuillage de représentativité (par exemple 5%). Le seuillage, s’il est trop élevé, tend à dénaturer le bénéfice de la proportionnelle, car finalement un scrutin majoritaire n’est pas autre chose qu’une proportionnelle dotée d’un seuillage à 50%...

Plusieurs programmes politiques des candidats en 2017 proposaient d’instaurer une « dose » de proportionnelle ; les systèmes dits « semi-proportionnels » ne sont pas négligeables dans le monde (12% des cas) et affectent une partie des députés nominativement par circonscription, et une autre partie par liste à la proportionnelle. Reste alors une question importante : quelle dose, et de quel tour utilise-t-on les résultats pour calculer les scores proportionnels au niveau national, le 1er ou le 2nd?

Pourquoi parler de taille de circonscription ?

Distribution spatiale des opinions politiques

Il semble naturellement souhaitable que la diversité des courants politiques qui traversent le pays soit représentée à l’Assemblée Nationale. Or, paradoxalement, **rien dans le système électoral ne le garantit**. En effet, le paradigme de notre mode de scrutin, à la fois fondamental et implicite, est que cette représentativité n'est assurée que par l'hétérogénéité spatiale des opinions politiques dans le pays; on peut montrer que si la population était répartie de façon homogène sur le territoire (disons, aléatoirement), alors il n'y aurait **qu'un seul parti politique à l'Assemblée Nationale**. En effet, on sait que si on échantillonne 1000 personnes au hasard dans la population, les opinions politiques sont à peu près celles de l'ensemble du pays. Il n'y a donc a priori pas de raison pour que deux circonscriptions au hasard, qui comptent 80 000 électeurs chacune, donnent des résultats différents.

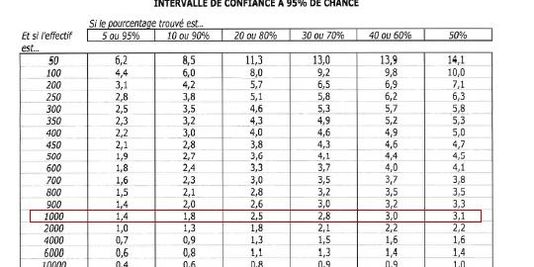


Figure : Intervalles de confiance des sondages IFOP par méthode des quotas. Lecture : dans le cas d’un échantillon de 1 000 personnes, si le pourcentage mesuré est de 20 %, la marge d’erreur est égale à 2.5

Dans le cas des législatives de 2017, au 2nd tour, La République en Marche obtient 43.1% des suffrages exprimés au plan national contre 22.2% pour les Républicains arrivant en deuxième position. Imaginons à présent une France fictive où on aurait réparti aléatoirement les 67 millions d’habitants dans les 577 circonscriptions. Pour que le parti LREM soit battu par LR – par hasard – sur une de ces circonscriptions fictives, il faudrait qu’elle soit aussi petite que 50 votants ! Au-delà, les lois statistiques font systématiquement gagner LREM et l’Assemblée Nationale serait mono-couleur.

La représentativité des courants politique est donc une conséquence des corrélations spatiales de l’opinion politique sur notre territoire (L’élection de 2012 montre bien la corrélation spatiale en Figure 3 par exemple), qui elle-même découle de nombreux facteurs comme :

* Les problématiques économiques et politiques locales (ex : les anciens bassins miniers et terres d’immigration favorisent l’extrême droite , voir les travaux de Pascal Perrineau).
* L’histoire des populations (travaux d’Hervé le Bras) et des structures familiales sous-jacentes (travaux d’Emmanuel Todd).
* La mobilité géographique des personnes et les phénomènes d’auto-ségrégation résidentielle (travaux de Thomas Schelling).

Il semble tout de même assez *hasardeux* de se fier uniquement à ce genre de facteurs, qui évoluent au fil du temps de façon non contrôlée, pour garantir la représentativité d’une assemblée législative.

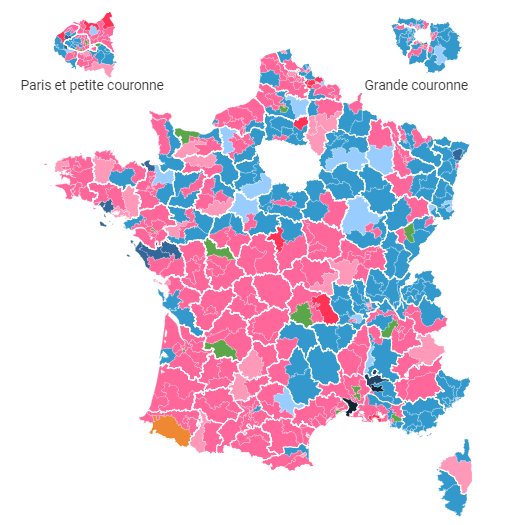
De fait, on voit bien que la taille des circonscriptions est un paramètre majeur, qui découle du nombre de députés (577). On comprend donc aisément que **plus les circonscriptions sont grandes, plus l'homogénéité politique sera forte**. A l'extrême, l'élection présidentielle fonctionne comme une élection législative à une seule circonscription qui regroupe toute la France. Et il n'y a bien qu'un seul parti politique représenté à l'Elysée.

Figure : Résultats du 2ème tour des élections législatives de 2012 en France, source : La Croix

Si on veut étudier ce paramètre de taille, il faut se libérer de la contrainte du nombre de députés fixe à 577. Les travaux de la présente étude simulent des circonscriptions, non pas fictives, mais bien réelles et de taille croissante. On va chercher à agréger les résultats des votes, de proche en proche, en commençant en bas de l'échelle par le bureau de vote et en finissant par la France entière, tout en conservant autant que faire se peut la contrainte de toujours avoir des circonscriptions de cardinaux homogènes. C’est l’objet de la troisième partie.

Définir une mesure de l’inégalité de représentation à l’Assemblée Nationale

En économétrie, une mesure classique d’inégalité de distribution d’une quantité est fournie par l’indice de Gini. Il n’est pas inintéressant de tenter d’utiliser cet indicateur pour caractériser une situation de « sur-représentativité » de certains courants politiques. L’indice de Gini mesure l’écart à une situation d’égalité parfaite de répartition entre tous les acteurs, ce n’est pas ce qui nous intéresse ici car il n’y a aucune raison que tous les partis aient chacun le même nombre de députés (la CNCCFP reconnaissait 451 formations politiques en 2016 !) ; on souhaite en revanche mesurer un écart par rapport à une situation de représentation strictement proportionnelle. Dans sa formule, est le nombre de nuances politiques (certaines comme DIVERS étant des regroupements de nombreuses petites formations) et le nombre de députés obtenu par la nuance . On souhaite que notre nouvel indicateur vale 0 dans la configuration « proportionnelle intégrale » à partir des résultats du 1er tour, et vale 1 lorsqu’il n’y a qu’une seule nuance présente à l’Assemblée (on réalise donc une simple normalisation min-max ):

On pourra alors calculer pour chaque configuration de scrutin qu’on souhaite, et en particulier pour chaque taille de circonscription.

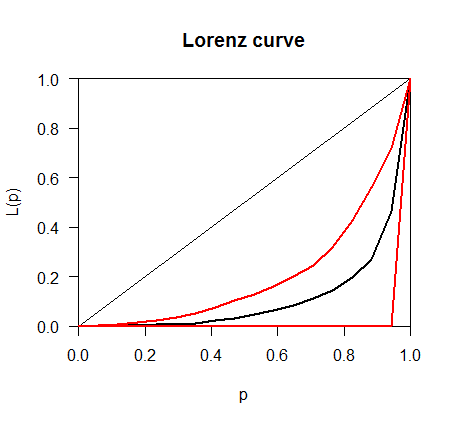


Figure : courbe de Lorenz de l'inégalité de répartition des députés : pourcentage des sièges cumulés obtenus par un pourcentage cumulés des nuances. La courbe noire (composition réelle de l'AN) est encadrée par deux situations fictives en rouge : équité maximale (proportionnelle) au-dessus et en dessous d'iniquité maximale (1 seule nuance).

Méthode de clustering et résultats

Critères pour un clustering des bureaux de vote

Cette partie traite des méthodes permettant d’obtenir des clusters de points dans l’espace (deux dimensions (X, Y)), et respectant un certain nombre de contraintes opérationnelles. L’unité géographique minimale est le bureau de vote, il n’y a pas de comptage à une échelle inférieure. Un bureau de vote représente environ un millier d’électeurs.

* Les clusters doivent avoir une taille égale à 20% près en nombre d’électeurs.
* Les clusters doivent regrouper des points selon un critère de proximité « relâché » et avoir des formes convexes dans la mesure du possible.
* A chaque augmentation d’échelle, les clusters du niveau inférieur sont conservés tels quels et servent d’individus à clusteriser pour le niveau . Ce critère correspond au principe d’imbrication des échelles successives qui vaut pour tout découpage administratif.

On ne cherchera pas ici à développer un nouvel algorithme de clustering mais à identifier un algorithme existant qui réponde globalement à ces critères, sachant que les principaux algorithmes classiques ne garantissent absolument pas des tailles de cluster constantes.

|  |  |  |  |
| --- | --- | --- | --- |
|  | Taille de cluster ? | Imbrication des échelles ? | Complexité calculatoire |
| k-means |  | Pas géré | Faible |
| k-medoids | Non | Pas géré | Faible |
| DbScan | Taille minimale possible | Pas géré | Faible |
| Clustering hiérarchique | Non | Géré via le seuil de coupure de l’arbre | Faible |
|  |  |  |  |
|  |  |  |  |

La technique par

Cette agrégation est alors dite « hiérarchique », elle se fait par un algorithme de clustering basé sur la distance, mais fonctionnant à taille de cluster constante (ou quasi-constante). C'est là la grande difficulté de ce projet, car un tel algorithme est par défaut sous-spécifié. Une des contraintes est d'agréger des bureaux de vote proches, pour conserver l’information de corrélation spatiale du vote.

Plusieurs approches ont été testées, l'approche retenue finalement étant celle du voyageur de commerce. Quelques exemples de résultats intermédiaires:

<https://www.la-croix.com/France/Politique/Elections-legislatives-2017-carte-resultats-2017-06-11-1200854147>

http://www.cnccfp.fr/index.php?art=220