L'enseignement des sondages au plus grand nombre : quelques réflexions tirées de l'expérience.

Benoît Riandey, Blöss-Widmer Isabelle

L'ENSEIGNEMENT DES SONDAGES À L'USAGE DU PLUS GRAND NOMBRE : QUELQUES RÉFLEXIONS TIRÉES DE L'EXPÉRIENCE

Benoît RIANDEY[1] et Isabelle WIDMER[2]

TITLE

Lecture on sampling to all users drawn from experiences

RÉSUMÉ

Le discours commun sur les sondages en France témoigne d'une grande méconnaissance et incompréhension des théories sous-jacentes. Heureusement, les exemples du quotidien ne manquent pas pour permettre d'accéder à leur compréhension intuitive[3], et en premier lieu à un repérage correct de l'unité statistique, à l'interprétation de la précision d'un sondage, à l'appréciation pratique de la méthode des quotas ou des redressements.

Mots-clés : *unité statistique, précision, quotas, redressement, sondage.*

ABSTRACT

Most French people misunderstand sampling. Mathematical formula cannot help them, but daily events can do: they can help in defining truly units, understanding sampling precision, accessing an empirical notion of quota sampling and sample correction.

Keywords : *statistical unit, precision, quotas, weighting, sampling.*

1 Introduction

Le sondage est aujourd'hui l'une des introductions de la statistique et des probabilités la plus proche du public et des jeunes, plus sans doute encore que la démographie. Il prend souvent la relève du rôle historique des jeux de hasard et, à ce titre, le sondage tient une place irremplaçable dans la vulgarisation de la statistique. La mécompréhension habituelle des concepts fondateurs de la théorie des sondages témoigne de l'ampleur du pas intellectuel à franchir pour accéder aux résultats de la statistique mathématique.

Nous avons eu l'occasion d'enseigner les méthodes de sondage à des publics variés allant de statisticiens démographes aux personnels administratifs de l'INED, en passant par les étudiants en sociologie ou en journalisme. Les objectifs étaient variés : pour les uns, il s'agissait de comprendre le principe du sondage afin de pouvoir en parler raisonnablement ; pour d'autres, il fallait acquérir un savoir-faire permettant une pratique professionnelle individuelle ou collective. Aussi devons-nous souvent adapter notre pédagogie aux objectifs et aux publics : ceux qui ont une sensibilité de matheux veulent comprendre l'origine des

[1] INED, Riandey@ined.fr

[2] Université de Provence, Département de sociologie, isabelle.widmer@univ-provence.fr

[3] Nous remercions les rédacteurs en chef et les relecteurs qui nous ont incités à expliciter davantage le cadre probabiliste de nos exemples, quitte à nous écarter, malgré notre intention, du langage accessible au plus grand nombre ; veuillez nous en excuser.

formules ; les autres préfèrent qu'on ne les montre pas trop. Et si une formule bien expliquée parle d'elle-même, les fichiers pédagogiques et les simulations demeurent les meilleurs soutiens à l'enseignement. D'ailleurs, pour un praticien expérimenté, une formule est bien souvent liée à l'histoire d'une enquête et à un choix opéré, tout comme à une erreur ou une catastrophe statistique. Rien de tel que les erreurs passées pour comprendre et faire comprendre les choses. Certaines notions théoriques seront en effet mieux assimilées si l'on fait appel à l'imaginaire des étudiants et à des situations très concrètes. A observer la vie quotidienne, chacun fait (ou veut faire) du sondage sans trop se questionner et les risques d'erreur d'observation sont alors nombreux. Nous pensons essentiellement aux erreurs de champ, à la mauvaise perception de l'unité de sondage et donc aux conditions d'inégales probabilités d'inclusion dans l'échantillon. Des paradoxes résultent de ces confusions et nous pensons qu'une réflexion en forme de devinette sur ces questions est toujours une bonne introduction aux sondages.

D'autres notions sont aussi rudes « à faire passer ». La précision d'un sondage, par exemple, repose sur une notion d'intervalle de confiance bien abstraite, voire même « anti-scientifique » pour certains auditoires, puisqu'il ne peut certifier à 100 % la fourchette de variation annoncée. Le recours à une métaphore très imagée nous aidera à présenter cette notion à un auditoire encore sceptique. Quant à la notion de « redressement », elle achève souvent d'introduire confusion tant elle comprend d'opérations miraculeuses ou placébo.

Mais ces efforts ne sont-ils pas « peine perdue » puisque la pratique médiatisée des sondages ignore les principes de la théorie ? Nous sommes convaincus du contraire et nous nous employons à en convaincre nos publics à chaque occasion. En effet, si le plus souvent les sondeurs travaillent sur quotas, c'est pourtant bien à la lumière de la théorie que l'on peut jauger de la qualité des sondages empiriques. C'est pourquoi nous confrontons aisément nos connaissances théoriques aux réflexions sommaires et inexactes que nous entendons souvent sur les sondages dont voici un échantillon :

- « *Un million de réponses volontaires valent mieux qu'un sondage* ».
- « *Sur 40 patients, j'en ai enquêté 20 : un taux de sondage de 50 %, c'est bien* ».
- « *Comment se fait-il que les statistiques prétendent que les autobus sont seulement à moitié pleins alors que vous et moi montons toujours dans des bus bondés* » ?
- « *Les sondages politiques sont bons à 3 % près* ».
- « *Mon échantillon est bon : mes quotas sont parfaitement respectés* ».
- « *La méthode des quotas est meilleure car au moins les variables de quotas sont mesurées sans aucune erreur d'échantillonnage* ».
- « *Pas représentatif mon échantillon ? Pas grave, on redresse* ».

Nous emprunterons donc, par l'anecdote et l'expérience, le chemin de l'enseignement du sondage à l'usage du plus grand nombre pour en présenter les principaux écueils.

2 La difficile pédagogie du hasard

La théorie des sondages s'impose par une idée bien abstraite : hors de la collecte exhaustive, hors du recensement donc, seul le tirage au hasard mathématique nous protège du biais de sélection. La bonne volonté ne peut pas remplacer le hasard, si difficile à simuler. Qui parmi nous n'a joué un jour à écrire une suite fictive de piles ou faces, à compter le nombre de séries longues de piles pour enfin constater qu'il n'a pas osé en écrire un nombre suffisant ? Nous sommes de bien faibles simulateurs du hasard[4]. S'en abstraire présente des risques.

2.1 Chaluter n'est pas sonder

Au printemps 1994, le premier ministre Edouard Balladur organise une vaste consultation des jeunes âgés de 15 à 25 ans et recueille un million cinq cent trente neuf mille questionnaires. Suite à ce succès quantitatif, en meilleur politique que statisticien, il déclare que « *Un million cinq cent mille réponses volontaires valent mieux qu'une enquête par sondage* ». Au contraire, les spécialistes nient toute valeur scientifique à ce vrac de questionnaires et la Société Française de Statistique (SFdS) demande dans un communiqué à la presse que, pour disposer d'une information scientifiquement fondée, ce questionnaire soit soumis à un échantillon représentatif. Ainsi, la première consultation estimait à 33 % la proportion de jeunes « *pas du tout confiants dans l'avenir* », tandis que quelques mois plus tard SCP Communication estimait cette proportion à seulement 13 %. Le décalage temporel ne rend certes pas directement comparables ces proportions, mais il conforte l'idée que l'inquiétude pour l'avenir aurait été une très forte motivation pour répondre à la consultation organisée par le Premier Ministre. Le succès de cette consultation postale était alors l'expression du mécontentement politique des jeunes, ce qui n'est guère un critère de représentativité.

De la même façon, les statisticiens ont repéré que le biais sélectif des réponses spontanées, notamment des enquêtes par voie de presse, est souvent motivé par une réaction relevant de l'esprit de contradiction ; ce moteur psychologique s'avère facilement activé mais, nous en conviendrons, il est, lui aussi, très peu favorable à la représentativité.

L'abondance de réponses sans méthodologie ne suffit ainsi aucunement à produire la qualité. Le hasard mathématique est nécessaire pour protéger des biais de sélection. Ramasser toutes les réponses qui se présentent, comme un chalutier drague les fonds, semble au premier abord très efficace, mais on vient d'en voir le résultat. C'est ce que risquent de produire les enquêtes par Internet les plus sommaires : placer un questionnaire sur un site Internet pour solliciter la réponse de tous les internautes de passage relève de cette logique. Une pratique plus professionnelle des enquêtes par Internet aboutit à d'autres résultats. Nous en parlerons.

De nombreuses situations induisent des biais de sélection : « comme par hasard », les femmes décrochent plus souvent le téléphone lors d'enquêtes téléphoniques. En se rendant au domicile des enquêtés, les enquêteurs préfèrent interroger une personne présente, donc probablement souvent à la maison, plutôt que de revenir interroger une personne absente lors du premier passage. Bref, on a toujours de bonnes raisons de « choisir » et les autres font « comme par hasard » le plus souvent les mêmes choix. Le hasard des opportunités à saisir n'est que source de biais… mais aussi de réduction des coûts. Cruel et permanent dilemme !

Choisir ses enquêtés permet « d'équilibrer l'échantillon ». Mais tout dépend des algorithmes de choix. A témoin, cet exemple caricatural d'une première rencontre avec un enquêteur recruté pour des enquêtes sur quotas : « *Pour vos enquêtes, je vous sélectionne des enquêtés sérieux, pas ces gens qui racontent n'importe quoi. D'ailleurs j'ai remarqué qu'en discutant du thème du questionnaire avant l'interview avec l'enquêté, il répond beaucoup mieux au questionnaire* ». On n'insistera jamais assez sur l'importance de la formation des

[4] On pourra consulter à ce propos le site Statistix : http://www.statistix.fr/spip.php?article50 .

enquêteurs à la notion du hasard mathématique, comme à celle de la neutralité de l'enquêteur. Néanmoins, nous vanterons l'efficacité des échantillons aléatoires équilibrés[5].

2.2 Univers et unités statistiques : bizarreries et paradoxes

Un échantillon aléatoire requiert d'abord l'existence d'une base de sondage dont on scrutera l'exhaustivité, la mise à jour, la richesse informative et la fréquence dans cette base de la sous-population étudiée. Il ne s'agit là que d'une transposition de la définition démographique d'une population avec ses trois critères d'inclusion : le concept définissant la population, ses délimitations temporelle et géographique.

Mais si les démographes se penchent sur une population de personnes, les sondeurs sont amenés à distinguer toutes sortes d'unités pouvant prêter à confusion. Retenons l'impérieuse nécessité de sensibiliser les étudiants à l'identification de l'unité de sondage, parfois distincte de l'unité statistique d'analyse. Le sondage à plusieurs degrés familiarise les étudiants avec la question, mais certainement insuffisamment tant on voit d'erreurs de professionnels dans l'usage des pondérations.

Le plus souvent, la question de la définition de l'unité statistique à enquêter ne se pose pas : un sondage électoral d'électeurs traite d'individus ; une enquête « budget de famille » fait appel à un sondage de ménages. Mais cette simplicité habituelle peut être trompeuse et fort heureusement, quelques paradoxes remarquables viennent attirer notre attention et maintenir notre vigilance *: « Dans l'après-guerre, un quart des mères avaient quatre enfants ou plus. Pourtant dans ma classe, comme dans les autres, nous étions la moitié à appartenir à une famille d'au moins quatre enfants* ».

Avec pédagogie, Laurent Toulemon démonte l'erreur tendant à confondre un sondage de mères, pour l'étude des descendances, avec un sondage d'enfants[6] pour la mesure des fratries. Certes un sondage d'enfants permet l'estimation de la descendance de leurs mères, mais en tenant compte des probabilités d'observation inégales introduites au niveau des mères précisément par ces descendances inégales, donc en pondérant chaque enfant par l'inverse de la taille de sa fratrie. Sinon l'erreur serait conséquente. Ne confondons pas unité de sondage et unité d'analyse[7].

Le paradoxe n'était pas différent pour les statistiques (un peu anciennes) de remplissage des avions : « *Comment les compagnies aériennes s'accordent-elles à déclarer un taux de remplissage moyen de 50 % alors que quand vous ou moi prenons l'avion, il est presque toujours plein* »[8] ? Dans ce sondage de passagers, dont nous sommes les observateurs, chaque

[5] Par un algorithme complexe, le sondage équilibré construit progressivement un échantillon aléatoire en s'imposant de respecter exactement certaines contraintes statistiques de distributions ou moyennes. Il cumule l'avantage du respect de quotas avec la qualité de l'échantillon probabiliste (cf. infra 5.3).

[6] 3,1 enfants par mère de la génération 1930, mais des fratries moyennes de 4,5 pour les enfants de ces mères. Cf. *Population et Sociétés*, 374 (www.ined.fr/publications). Voir l'annexe pédagogique 1 ci-jointe. Un fichier pédagogique de cette application est disponible à l'adresse www-ehf.ined.fr .

[7] Rappelons qu'il est parfaitement fondé de réaliser des sondages à probabilités inégales dès lors qu'aucune fraction de cet univers fini n'a une probabilité nulle d'être échantillonnée (ce qui l'exclurait de la population étudiée). Mais l'estimation doit compenser cette inégale probabilité dite d'inclusion (à l'échantillon) en pondérant chaque unité de l'échantillon par l'inverse de sa probabilité d'inclusion. On compense donc le sous-échantillonnage par une pondération proportionnellement accrue. On peut cependant se dispenser de cette pondération lorsqu'il y a indépendance entre la probabilité d'inclusion et la variable étudiée. Cette situation sera évoquée dans l'exemple du sondage dans les archives et dans la justification des enquêtes par quotas.

[8] Voir aussi http://www.statistix.fr/spip.php?article25 .

vol est observé avec une probabilité proportionnelle à son nombre de passagers. Pondérons donc nos observations par l'inverse de ce nombre de passagers et nous tomberons d'accord avec les compagnies.

Cet exemple illustre comme notre vision du monde ambiant peut être biaisée par l'ignorance de nos probabilités d'observation et nous oriente vers une nouvelle disposition mentale pour reconstituer cet univers ; par exemple, un automobiliste roulant à 125km/h sur l'autoroute et se voyant dépassé par une multitude de Mercedes et de BMW en conclura spontanément que le parc automobile français est essentiellement constitué de ces voitures. Au contraire, le sondeur, pondérant chaque voiture qui le dépasse (ou qu'il dépasse) par l'inverse de la différence – absolue – de vitesse des deux véhicules, aura une vision juste du parc automobile empruntant l'autoroute. Mais, en bon statisticien, pour observer les voitures roulant comme lui à 125 km/h, il n'omettra pas de changer d'allure et corrigera ainsi cette erreur de couverture[9].

Par un raisonnement trop rapide, l'application précédente semble expliquer pourquoi telle enquête sur la fréquentation des autobus indique des autobus bondés alors que là encore les entreprises de transports annoncent un taux de remplissage de 50 %. En fait, il s'agissait bien d'un sondage d'autobus et non de passagers : il avait été décidé de tirer le prochain autobus se présentant à un arrêt intermédiaire donné. Pourquoi donc, malgré cette précaution de sondage, rencontrer la même estimation biaisée ? Nous présentant au hasard à l'arrêt de bus un instant donné, nous tirons un autobus avec une probabilité proportionnelle à l'intervalle inter-bus[10]. Nous tirons donc avec une probabilité plus grande les bus qui se font attendre et que nous savons bondés. Bref nous avons tiré des secondes d'attente et non directement des bus, d'où l'estimation biaisée. Tirons donc nos bus dans la liste de la compagnie, ou pondérons par l'inverse du temps d'attente… en laissant passer le premier bus pour observer l'intervalle d'attente complet[11].

Cette anecdote nous rend apte à l'exercice pédagogique suivant, histoire vraie ou inventée dont on ne connaît pas l'auteur. Avant le développement de la gestion informatique, une société de logements sociaux s'inquiétait de la fréquence apparente des impayés. Le responsable recrute un stagiaire et lui dit : « *Tu vas me tirer cent dossiers au hasard et me les regarder de près* ». En l'absence de liste des dossiers, le stagiaire, observant 75 mètres linéaires de dossiers classés par ordre alphabétique, mesure 75 centimètres sur une ficelle (ou avec un mètre de couturière) et sort un dossier par longueur de ficelle. Mais à la présentation des résultats de l'étude par le stagiaire, le responsable s'étonne : « *Une telle proportion de dossiers d'impayés, ce n'est pas possible !* ».

L'erreur est simple en effet : le stagiaire n'a pas réalisé un sondage uniforme entre les dossiers, mais a tiré des centimètres et donc des dossiers proportionnellement à leur épaisseur, ce qui entraîne la sélection d'un plus grand nombre de dossiers avec impayés car ceux-ci sont plus épais. Pour bien faire, notre stagiaire aurait dû penser à une astuce : ayant mesuré 75 centimètres, il aurait suffi de prendre le dossier suivant et de sortir ainsi 100 dossiers. Les

---

[9] Ce n'est pas l'objet de cet article de préciser la configuration autoroutière et de flux automobile fondant cet exemple. Pour les spécialistes, nous avons raisonné en modèle de super-population et non en population finie, à l'image des flux de particules des accélérateurs linéaires de la physique nucléaire. Tout exemple appellerait spécification des hypothèses. Excusez ce langage.

[10] Si on partage la plage de circulation des autobus en micro-périodes, la probabilité de sélectionner un autobus donné sera proportionnelle au nombre de micro-périodes séparant le passage de l'autobus précédent et celui de l'autobus tiré.

[11] Nicolas Gauvrit (2007), *Statistique, méfiez- vous*, Ellipse, Paris, p.192.

dossiers successifs sont d'épaisseurs indépendantes, classés par ordre alphabétique, c'est-à-dire dans un ordre a priori non signifiant. Une fois ces 100 dossiers sortis sur la table, le stagiaire reprend son mètre de couturière, mesure l'épaisseur de chaque dossier et le pondère par l'inverse de son épaisseur[12].

L'erreur ne sera plus refaite mais cette méthode ne porte alors plus le nom de sondage au hasard mais bien celui de sondage systématique puisque la première unité tirée détermine toutes les autres par simple progression arithmétique !

2.3 Eloge du sondage systématique

C'est ainsi que les qualités des sondages systématiques méritent d'être vantées auprès des étudiants : faciles à mettre en œuvre, ils offrent de multiples usages. D'ailleurs nous faisons régulièrement des tirages systématiques sans le vouloir, par exemple en distribuant les cartes à la ronde, comme au bridge notamment[13]. Dès avant la distribution, les quatre jeux de 13 cartes sont constitués mais pas attribués. En coupant le jeu (soit en tirant au hasard un nombre, demeuré inconnu, entre 1 et 4 ou entre 0 et 3), on nous affecte notre échantillon de treize cartes au taux de sondage 1/4 selon la règle de la progression arithmétique. Si le jeu a été bien battu, on se convainc facilement de ce qu'un tirage systématique est alors équivalent à un sondage aléatoire simple.

Si le jeu est trié selon une catégorie qualitative ou quantitative, le tirage systématique équivaut à une quasi stratification selon ce critère[14]. C'est précisément ce qui se passe avec un jeu neuf dont les cartes sont triées par valeur décroissante de l'as (la plus forte) au 2 (la plus faible), les quatre cartes de la strate respectant l'ordre pique, cœur, carreau, trèfle. Quelle que soit la coupe, le joueur recueillera un jeu parfaitement équilibré entre les treize strates (valant toujours 10 points d'honneur au bridge). Sur un fichier opportunément trié, le sondage systématique est meilleur que le sondage aléatoire simple.

Certes, le tirage systématique est sujet à une régularité bizarre quand le pas de sondage est multiple d'un cycle présent dans la base. C'est le cas qui nous occupe puisque avec le jeu neuf, un joueur recueillera treize cartes de la même couleur et serait ainsi porté à dire : « *d'après mon échantillon, ce jeu ne comporte que des piques* ». Cette bizarrerie se retrouve lorsqu'à un mariage, dans une farandole bien réglée alternant garçons et filles, vous tirez un échantillon de danseurs avec un pas de sondage pair : à la seule vue de l'échantillon, vous jugeriez la noce entièrement masculine ou féminine. Toutefois, cette situation a peu de chance de se manifester dans la pratique professionnelle où le pas de la progression est généralement décimal. Nous avons dû nous en prémunir une seule fois dans toute notre carrière.

---

[12] L'astuce demande justification : la probabilité de sélection « astucieuse » d'un dossier est donc proportionnelle à l'épaisseur du dossier précédent, épaisseur totalement indépendante de ce qu'on va mesurer dans le dossier tiré. On est donc dispensé de pondérer par l'inverse de cette épaisseur. Par ailleurs, sous l'hypothèse théorique faite d'un ordre non significatif des dossiers, on pourrait estimer que les 100 premiers dossiers constituent un échantillon aléatoire simple de cette population de dossiers et, simplement, les sélectionner tous les cent. Mais cette stratégie de sondage est à ce point dépendante de la perfection de l'hypothèse faite qu'aucun praticien ne s'y risquerait.

[13] Le premier joueur se voit attribuer les cartes 1,5, 9,…,49 ; le second, les cartes 2, 6, 10,…,50, etc. Au tarot, la distribution des jeux n'est pas systématique du fait de la constitution du chien qui rompt la cadence.

[14] Le « quasi » tient au fait que les strates, si elles étaient définies par un découpage en classes de la variable de tri, n'auraient pas un effectif multiple du pas de tirage.

L'exemple des cartes ne saurait illustrer une performance majeure du sondage systématique : cette technique permet de tirer des communes (des unités primaires) proportionnellement à leur population. Il suffit de cumuler les populations des communes – opportunément classées par population croissante – et de sélectionner la commune d'un individu fictif tous les *n* individus. Ce glissement entre unités statistiques peut prendre à défaut le sondeur s'il n'y a pris garde, comme l'illustre la mésaventure réelle suivante.

La **méthode des itinéraires** désigne une application classique du sondage systématique. Il s'agit d'opérer un sondage de logements en l'absence de fichier, c'est-à-dire de base de sondage des logements. Le tirage est alors opéré directement sur le terrain : on définit un balayage de l'espace sondé, par exemple le quartier – l'Iris dirait l'INSEE –, qui ne passe qu'une fois à chaque adresse ; puis l'enquêteur retient un logement tous les *n*. Cette méthode est très employée aux Etats-Unis dans un habitat pavillonnaire. En France, elle a donné lieu à quelques belles catastrophes statistiques. Ainsi tel étudiant géo-démographe, enquêtant en habitat mixte, sélectionnait une adresse sur 7 et, à sa propre surprise, conclut de son enquête que la commune populaire enquêtée était bien loin d'atteindre le seuil des 20 % de logements sociaux requis par la loi SRU relative au logement social. Grossière erreur méthodologique, mais laquelle ?

Ce tirage systématique a été mené dans un univers d'adresses et non de logements[15]. Ainsi, un pavillon avait une chance sur 7 d'être tiré ; un logement social d'un HLM de 30 logements avait une chance sur 210 de figurer dans l'échantillon. N'ayant pas cumulé le nombre de logements à chaque adresse pour tirer des logements et non des adresses au cours de l'itinéraire, l'erreur était prévisible. Ainsi, dans ce HLM, on aurait, selon le hasard, sélectionné 4 ou 5 logements. L'étudiant pouvait cependant réparer son erreur en pondérant chaque observation par la variable « *Nombre de logements dans l'immeuble* » pour mesurer ensuite correctement la variance de l'estimateur.

La leçon a été bien comprise mais, malheureusement, la dimension des immeubles pavillonnaires et sociaux est tellement disproportionnée que la variance se met alors à exploser. Nous conseillons à l'étudiant de retourner dans le quartier pour faire un sondage complémentaire dans les immeubles collectifs, et fusionner ensuite correctement ses deux sous-échantillons. Pour bien faire, il aurait fallu cumuler le nombre de logements à chaque adresse en parcourant le quartier et retenir un logement tous les *n* ; à décharge pour l'étudiant, cette opération n'est pas du tout commode à faire, même lorsque l'on connaît *a priori* l'effectif des logements de l'immeuble.

Dans la pratique, en France, les communes de plus de 10 000 habitants tiennent à jour un répertoire des immeubles localisés, le RIL du *Recensement rénové*. Ce répertoire enregistre le nombre de logements à chaque adresse. Il n'y aurait aucune difficulté technique à procéder à un sondage de logements dans ce fichier anonyme RIL et nous ne comprenons pas pourquoi il y est encore fait obstacle. De même, le « bulletin immeuble » du recensement des petites communes dénombre les logements de l'adresse. Si ces bases de sondage étaient plus accessibles en France, la méthode des itinéraires ne serait plus utile pour les sondages en habitat aggloméré.

[15] La méthode des itinéraires a été conçue pour un univers de pavillons dans laquelle les unités « immeuble » et « logement » sont confondues. Dans un habitat mixte de pavillons et d'immeubles collectifs, le fait réel cité, consistant à sélectionner des adresses puis un seul logement par adresse, correspond à un sondage à deux degrés : systématique équiprobable pour les immeubles, à taux de sondage variable pour le logement de chaque immeuble.

3 L'incertaine précision d'un sondage

L'idée (numériquement fausse) la plus répandue sur la précision des sondages est qu'elle est fonction du taux de sondage ($f = n/N$). La précision des sondages ne dépend habituellement pas du taux de sondage[16], mais bien de la seule taille n de l'échantillon[17]. Ainsi, la formule *approchée* de la précision d'une proportion parle d'elle-même : l'expression $\pm 2\sqrt{pq/n}$ dit que le taux de sondage f n'influe pas et que seule n, la taille de l'échantillon, pèse. Nous venons de chasser l'opinion rassurante : « *Un taux de sondage de 50 % ou de 10 %… c'est bien* ».

3.1 Se familiariser avec la précision des proportions

La découverte du joli tableau de Anne-Marie Dussaix et Jean-Marie Grosbras[18] (reproduit en annexe 2) débarrasse ensuite nos auditeurs des simplifications médiatiques statuant sur la précision des sondages politiques, du type : « *Avec 20 % d'intentions de vote, on connaît le pronostic à 3 % près pour ce candidat* ». En fait, avec un échantillon de 1 000 répondants, la demi-largeur de l'intervalle de confiance n'est pas toujours de 3 %, mais au maximum de 3 %, et, à ce niveau de fréquence observée de 20 %, la demi-largeur de l'intervalle de confiance vaut environ 2,5 %. Voilà qui serait compliqué à mentionner dans la grande presse. On insistera aussi sur le fait que la précision est coûteuse puisque : « *Pour avoir une enquête deux fois plus précise, il faut un échantillon quatre fois plus grand. Pour gagner une décimale de précision, il faut un échantillon cent fois plus grand, cent fois plus cher. Pour estimer avec une précision de 2 % (0,02) une proportion tournant autour de 10 % (0,10), l'échantillon doit être de taille 900* ». Et l'on n'omettra pas de passer de la précision absolue à la précision relative, pour que nos auditoires comprennent pourquoi les enquêtes épidémiologiques recourent à de grands effectifs pour mesurer la prévalence de maladies rares.

Nos collègues objecteront ensuite, à juste titre, que cette présentation de l'intervalle de confiance n'est pas exacte : la formule présentée relève du sondage aléatoire simple *avec* remise, alors que cette notion de remise n'a pas de sens pratique dans les sondages puisqu'on ne va pas poser deux fois le même questionnaire à une personne. Dans la situation habituelle des sondages nationaux, la formule rigoureuse $1{,}96\sqrt{1-f}\sqrt{pq/n}$ pour la demi-largeur de l'intervalle de confiance est tellement numériquement identique à celle relative au sondage avec remise qu'on évitera aux non mathématiciens d'entrer dans cette distinction troublante.

Cette première étape passée, on introduira la formule exacte comme une correction nécessaire fondée sur le contre-exemple flagrant du recensement, puis on examinera jusqu'à quel taux cette correction s'impose. On éveillera ainsi l'esprit critique de nos auditeurs sur ce fait : par nature, l'incertitude d'échantillonnage des estimations d'un recensement est nulle ; or, la formule approximative $2\sqrt{pq/n}$ ne le montre pas. Il y a donc lieu d'introduire le facteur correctif $\sqrt{1-f}$ qui précisément s'annule pour la valeur 1 du taux de sondage f aux recensements.

---

[16] ou habituellement si peu (cf. ci-après).

[17] Pour une population donnée de taille définie connue, préciser le taux de sondage ou la taille de l'échantillon sont synonymes, mais cette situation tautologique est tout autre.

[18] Dussaix, A.-M. et J.-M. Grosbras (1993), *Les sondages. Principes et méthodes*, Que sais-je ?, n° 701, p. 27.

Ce facteur correctif n'est aucunement négligeable quand les taux de sondage sont inhabituellement très élevés (1/10ème, 1/4, 1/2). Un auditoire sceptique en prendra conscience avec l'exemple suivant : une administration en charge des populations étrangères lance un appel d'offre pour des études ne devant pas dépasser 100 interviews. Un chercheur remporte l'appel d'offre avec un projet concernant un corps professionnel à enquêter formé de 130 personnes étrangères. Le chercheur est donc contraint de constituer un échantillon de taille 100 issu d'une population de taille 130. Malgré le taux de sondage de 3/4 environ[19], l'incertitude d'échantillonnage n'est réduite que de moitié par rapport à un échantillon de même taille en population générale[20]. Le statisticien conclura que le chercheur est en bonne voie mais qu'il doit absolument terminer le travail en enquêtant les 30 derniers individus. Dans les petites populations, le seul bon sondage est l'enquête exhaustive, le recensement de tout le champ. Au jeune médecin satisfait d'avoir enquêté 20 de ses 40 patients, nous répondrons toujours : « *Terminez le travail en enquêtant les 20 autres* ».

3.2 Complexe la notion d'intervalle de confiance ?

La table de A.-M. Dussaix et J.-M. Grosbras nous indique donc « *qu'un échantillon de 1 000 enquêtés estime à 2,5 % près une intention de vote voisine de 20 %* ». Toutefois, un collègue rigoureux complètera que, dans cinq cas sur 100, la valeur à estimer ne figurera pas dans cette fourchette. Qui peut donc comprendre que l'intervalle annoncé n'est pas le bon pour 5 échantillons sur 100 ?

Effectivement, en sélectionnant un échantillon au hasard dans la population, même d'une taille importante de 1 000 personnes, on peut avoir la malchance insigne de sélectionner uniquement des électeurs au comportement électoral atypique et l'on comprend alors que les sondages sont susceptibles de se tromper dans l'évaluation. Evidemment, l'explication ne peut en rester là et la question suivante est alors de savoir d'où sortent ces 5 % de « mauvais échantillons ». Nous proposons souvent d'illustrer cette énigme par une métaphore de chasse aux papillons, considérant qu'une estimation d'enquête peut être assimilée au coup de filet d'un chasseur de papillons qui présente diverses chances de rater sa cible ou de capturer sa proie :

Une estimation d'enquête
=
Un coup de filet à papillons

avec - 1 chance sur 10 de rater la cible (estimation *) ;
- 1 chance sur 20 de rater la cible (estimation **) ;
- 1 chance sur 100 de rater la cible (estimation ***).

La valeur à estimer est ici notre papillon. Un filet à papillons est d'abord doté d'un point central et d'une largeur autour de ce centre (d'un rayon s'il est circulaire). De plus, le filet à papillons rate d'autant plus souvent sa cible qu'il est plus étroit. Le filet à papillons présente donc toutes les caractéristiques d'un intervalle de confiance. Le statisticien qui a collecté son enquête d'une taille d'échantillon fixée dispose d'un jeu de filets à papillons d'une, deux ou trois étoiles, de largeurs différentes. Ils manquent le papillon 1 fois sur 10, 1 fois sur 20 ou 1

---

19 100/130, soit 77 %.

20 $\sqrt{1-f} = \sqrt{1-3/4} = \sqrt{1/4} = 1/2$

fois sur 100, respectivement. Le filet trois étoiles, beaucoup plus sûr, avec seulement un échec sur cent, est nécessairement beaucoup plus large.

Une estimation d'enquête s'apparente donc à un coup de filet à papillons qui prendra sa proie… sauf x fois sur cent. Pour ne pas se compliquer la vie, les statisticiens prennent, sauf mention contraire, un filet deux étoiles, c'est-à-dire qu'ils s'arrogent le droit à l'erreur 1 fois sur 20, et voici nos 5 %.

4 Les redressements : solutions à tous les maux ?

« *Pas représentatif mon échantillon ? Pas d'importance, on redresse* ». Cette réflexion ancienne d'une responsable d'un grand institut de sondage français n'avait pas manqué de choquer son interlocuteur, mais elle pose une question théorique et pédagogique importante et complexe. Elle comprend sa part de vérité comme d'erreur.

C'est vrai qu'il est facile de remettre à niveau la proportion d'hommes ou de Franciliens dans un échantillon. Mais est-il indifférent que l'algorithme de redressement (expression déplorable) corrige judicieusement par post-stratification une fluctuation aléatoire ou masque un biais issu d'un phénomène de sélection ? La réponse dépend du processus sélectif. Si spontanément les femmes répondent davantage aux enquêtes, c'est pour une part parce qu'elles sont plus souvent présentes au logement. Cette cause signifie que les femmes au foyer seront surreprésentées par rapport aux femmes actives. Et que ce soit par redressement ou par quotas que le sous-échantillon féminin est remis au bon niveau ne change rien à son biais de structure.

Ainsi, même redressés de leur sous-représentation des catégories ouvrières, les panels postaux des instituts privés conservent le biais intrinsèque du mode de collecte à l'encontre des personnes peu instruites, mal à l'aise avec l'écriture. « *Il n'y a qu'à redresser en fonction du diplôme* » ! Hélas, le diplôme est la variable la plus mal renseignée au recensement, même pour les Français de naissance. L'importance de ce biais ne peut qu'être approchée faute de statistique fiable, si ce n'est au niveau national par l'enquête Formation Qualification Professionnelle de l'INSEE (Thélot, 2004).

Un excellent taux de réponse est sensé garantir l'absence de processus sélectif et l'INED s'abstenait de redresser les échantillons de ses enquêtes collectées par l'INSEE avec un taux de succès de 90 %. Néanmoins, l'ancienne enquête Emploi, du haut de ses 94 % de succès, enregistrait cependant une sous-représentation de certaines catégories de jeunes peu présents au domicile. La meilleure des enquêtes en France demeurait donc imparfaite.

Revenons au mot **redressement**. Il décrivait l'enchaînement de « vilaines règles de trois » qui redonnait une allure propre aux échantillons des instituts privés… ou de l'INED, tandis que l'INSEE repondérait ses échantillons grâce à une superbe inversion de matrice. Nous devons à Jean-Claude Deville d'avoir effacé cette représentation : l'algorithme très simple de ratissage alternatif inventé par Deming est une optimisation très légitime des pondérations sous la contrainte du respect de distributions statistiques. « *Pourquoi faire compliqué quand on peut faire aussi bien simplement* ? » nous taquinait-il avant qu'Olivier Sautory conçoive son programme de redressement Calmar. Il reste que l'appréciation d'un sondage atteint d'une forte proportion de non-réponses suscite un malaise du professionnel, le redressement des non-réponses ne fournissant une estimation solide que si les hypothèses d'équivalence locale à la base du redressement ne sont pas trop fausses.

Précisons cette notion d'équivalence locale : le redressement des non-réponses à une enquête (refus de l'enquête, personne absente ou impossible à enquêter…) repose sur l'idée que l'on va transférer le poids des non-répondants à des personnes qui leur sont équivalentes. Ne pas faire de redressement consisterait à faire comme si les répondants étaient représentatifs des non-répondants. Or, les répondants et les non-répondants n'ont souvent pas la même structure par sexe, âge et profession. Choisir de redresser par rapport à ces variables sous-entend que les répondants d'un sexe, d'un âge et d'une profession donnés sont représentatifs des non-répondants de ce même profil. Ce n'est peut-être pas vrai, mais c'est déjà moins inexact à ce niveau « local » qu'au niveau de l'ensemble de l'échantillon. Cette hypothèse d'équivalence locale, si elle est raisonnablement tenable, permet de réduire le biais.

Les redressements ne corrigent pas tous les maux d'un trop mauvais échantillon. Les sondages électoraux français illustrent bien les écueils que peut rencontrer la pratique des redressements. Les échantillons électoraux bruts sont très fortement biaisés, mais ils profitent d'une circonstance favorable rare, la disposition d'une statistique exhaustive très corrélée à la variable d'intérêt : le résultat d'un vote antérieur ou bien parfois le premier tour de l'élection en cours. Grâce à cette variable de calage, la précision des sondages électoraux s'annonce donc théoriquement exceptionnelle, ce qui est régulièrement vérifié pour les estimations de second tour des élections présidentielles, mais sauvagement invalidé pour bien des premiers tours. Le redressement de l'échantillon sur la base du vote antérieur ne sauve pas les estimations d'intention de vote. La mauvaise qualité de réponse sur le vote passé biaise le redressement. On ne peut donc faire la part entre le biais d'échantillonnage et l'erreur de déclaration, volontaire ou non.

On ne saurait conclure sur les redressements problématiques des sondages politiques sans citer deux faits nouveaux : d'une part, l'incitation à davantage de méthode statistique auquel nous convie Daniel Bachelet (2007), d'autre part l'expérience de Jean Chiche (2009) d'un sondage politique en ligne. Loin de pratiquer le chalutage de bien des sondages non professionnels par Internet, il a constitué son panel des présidentielles 2007 dans un vivier d'internautes volontaires qui avaient accepté le principe d'une enquête (de sujet indéterminé), ce que les instituts appellent du vilain mot d'« access panel ». Son échantillon brut a directement fourni des pronostics fiables des votes, ce qu'aucune enquête téléphonique n'a permis. Il est néanmoins clair qu'un vivier d'internautes volontaires ne constitue pas un échantillon représentatif de la population (en particulier pas des sous-populations de non internautes ou d'internautes refusant le principe de répondre à une enquête). Mais cet exemple illustre qu'au delà de la mécanique du redressement, ce qui compte c'est le modèle sociologique de comportement d'enquêté de la population cible. La pratique d'un redressement des non-réponses est toujours – quoique le plus souvent implicitement – la mise en œuvre d'un modèle statistique reposant sur des hypothèses de comportement. Au contraire, s'il s'agit seulement de neutraliser la fluctuation aléatoire sur une ou plusieurs variables par post-stratification ou calage, les estimations sont libres de toute hypothèse sociologique, de tout modèle de comportement.

Enfin, ne rejetons pas les méthodes empiriques avant d'aller voir ce qu'elles donnent. Convaincus de l'apport du hasard, il nous faut cependant réaliser que la mise en œuvre de sondages sur échantillon aléatoire est plutôt limitée à la statistique publique et à la recherche et que la pratique en instituts privés est plutôt à la réalisation d'enquêtes par choix raisonnés sur des échantillons non probabilistes. Malgré tout, la théorie probabiliste sert de cadre de référence et d'évaluation à ces enquêtes empiriques, notamment aux enquêtes sur quotas.

5 Quelle juste appréciation des enquêtes par quotas ?

Le mot quota est connu de longue date dans les médias et a d'abord renvoyé aux fameux « quotas laitiers ». Il évoque seulement un effectif de référence et aucune nuance ne permet d'y associer un principe de sélection des unités constituant cet effectif. Ce mot est potentiellement une source de malentendus dans les sondages, comme le mot « hasard », ou le mot « sondage » employé pour enquête d'opinion par *sondage*. Un travail sur les définitions et le vocabulaire s'impose à nouveau avant tout contenu technique. Une enquête par quotas est d'abord une enquête dont le choix des enquêtés est laissé à la bonne fortune des enquêteurs même si, pour éviter le désastre statistique du chalutage, les instituts ont soumis les enquêteurs au respect de fameux quotas (le plus souvent par sexe, âge et profession). Cette simplification notoire du travail jette *a priori* un doute sur le résultat.

Comment apprécier justement cette pratique si éloignée des principes théoriques actuels ? D'ailleurs, si elle s'avérait si néfaste, ne l'aurait-on pas abandonnée ? Argument insuffisant, mais incitatif à la réflexion.

« *Mes quotas sont vérifiés, j'ai un bon échantillon* ». Cette remarque avait longtemps été jugée pertinente au début du XX$^{\text{ème}}$ siècle tant elle relève du bon sens, si on a sélectionné l'échantillon par un « choix judicieux », selon l'expression de l'époque. Alors le modèle réduit représente la totalité à la manière d'un plan ou d'une carte. Ressembler suffit-il pour représenter ? L'échantillon risque-t-il de n'avoir que l'apparence de l'univers représenté ?

Le choix judicieux semble une garantie ; en fait, il n'est que relatif au point de vue par lequel on aborde l'univers. La mécanique – de précision imparfaite – du sondage aléatoire présente une qualité démographique d'universalité : elle autorise des estimations sans biais – d'échantillonnage –, quel que soit le sujet traité. Elle induit un rapport direct entre les deux univers de la population et de l'échantillon, d'où notre qualificatif de propriété démographique.

Mais si finalement le sondage aléatoire est une condition regrettable mais nécessaire à la remarquable qualité d'universalité de l'échantillon, l'intuition du choix judicieux s'est finalement déplacée vers la stratification (Didier, 2009)… au point que l'on entend des professionnels confondre quotas et stratification géographique. Plus généralement, à l'intuition du choix judicieux s'est substituée la volonté d'introduire dans le sondage notre connaissance de l'univers sans renoncer à la sélection aléatoire des unités non exhaustivement retenues.

Il est exact que des quotas corrects améliorent des estimations souvent très corrélées aux variables sous quotas, comme par exemple le risque de chômage en fonction de la pyramide des âges. Mais les variables des quotas sont-elles aussi bonnes qu'elles paraissent ? Souvent les ouvriers manquent dans l'échantillon, mais dans le cas contraire, ne recueille-t-on pas une proportion excessive d'ouvriers qualifiés ? Que dire des cadres supérieurs ? Ne sont-ils pas en trop grande proportion des personnels du secteur public, notamment des professeurs ? Il nous faut donc intégrer que, même si l'on suppose que les variables des quotas sont intrinsèquement satisfaisantes, elles n'apportent de garantie que pour elles-mêmes. Que dire d'un échantillon de mères fondé sur des quotas strictement respectés dans une enquête sur le thème de la garde des enfants, mais collectée, par facilité, à la sortie des crèches ? « *100 % des nouveau-nés sont gardés en crèches prouve cette belle enquête sur quotas* » ! Des quotas apparemment bien respectés n'apportent aucune garantie à l'encontre des biais d'échantillonnage (probabilité forte des mères des bébés gardés en crèche d'être dans l'échantillon et nulle ou négligeable pour les autres mères). Les biais d'échantillonnage les plus graves sont ceux liés à la variable d'intérêt car, le plus souvent, on ne dispose d'aucune variable adaptée pour les redresser, ce manque d'information disponible étant la raison même de mener cette enquête.

5.1 Pourquoi les enquêtes par quotas, ça marche souvent ?

La force des sondages empiriques est de profiter de l'expérience passée pour en repérer les biais et leur trouver une parade par une consigne judicieuse : ne pas enquêter la même personne plus d'une fois par an ; ne pas enquêter plus d'une personne par cage d'ascenseur, imposer des horaires de collecte, notamment par téléphone… Les quotas de sexe et d'âge évitent d'enquêter trop de femmes ou trop de retraités. Ainsi on rétablit une équiprobabilité par sexe et par âge.

Néanmoins, il ne faut pas se faire d'illusion. Compte tenu de leur activité, de leur quartier de résidence, les individus n'ont aucunement des probabilités égales d'être enquêtés par quotas. La rigueur voudrait que leur observation soit pondérée par l'inverse de leur probabilité d'inclusion… qui nous est bien inconnue. Mais le biais constitutif des échantillons sur quotas n'entraîne pas nécessairement des estimations biaisées : le théorème de Horwitz-Thomson indique qu'il suffit que les probabilités d'inclusion – inconnues – soient indépendantes de la variable mesurée pour que l'estimation soit non biaisée. Dans ce cas, le biais sélectif n'est pas perturbant. La théorie vient donc au secours des sondages empiriques, mais sous une hypothèse qui n'est vérifiable que sur des variables déjà mesurées par ailleurs. Sinon l'expérience du terrain et le bon sens suppléent plus ou moins bien au contrôle de l'hypothèse. Notre chance est que la statistique est complaisante. Cette hypothèse hasardeuse se vérifie assez souvent pas trop mal.

Cette réflexion théorique sur les estimations correctes à partir d'échantillons biaisés s'adresse à des spécialistes. Après avoir convaincu les étudiants qu'il faut solliciter l'action du hasard, leur laisser entendre qu'on peut très bien réaliser de bonnes enquêtes en ne tenant aucun compte de ce principe fondamental devrait dérouter les étudiants autant que leur malheureux enseignant. Enseigner une chose et son contraire n'est pas confortable. L'argument selon lequel « avec de petits échantillons, le sondage aléatoire n'est pas très bon » vient à la rescousse de l'enseignant qui complètera en ajoutant que « rien ne garantit que la qualité des échantillons par quotas s'améliore en accroissant son échantillon exposé au biais ».

5.2 Pourquoi les quotas ne marchent pas toujours ?

Les biais de sélection sont souvent liés à la variable mesurée. Aussi est-on attentif à éviter les refus thématiques. Ainsi, lors d'une enquête par quotas sur la pratique sportive, les enquêteurs recevaient une première consigne absolue : ne pas prononcer le mot « sport » avant l'acceptation de l'enquêté. On visait par là à éviter des désistements sélectifs : *« Posez votre questionnaire à mon frère ; il répondra beaucoup mieux que moi* ». Malgré ces précautions, cette enquête a fourni une estimation très élevée de la pratique sportive, et l'on n'a jamais su si elle était due au non respect de cette consigne, au recrutement des enquêtés à la sortie de salles sportives, à un recrutement d'enquêteurs sportifs, à une définition large non contrôlée de la pratique sportive. Personne ne pourra jamais le dire.

C’est pourquoi nous conseillons toujours la généralisation des introductions peu inductives utilisées dans les sondages omnibus : « *Voulez-vous bien répondre à un questionnaire sur des thèmes variés* » ?

Une expérience ancienne de l’INED engage à une certaine précaution : dans la décennie soixante-dix, ses enquêtes ont montré l’émergence, puis l’extension de la cohabitation juvénile. Fort intéressantes, elles ont eu un retentissement certain, en particulier l’enquête de 1977. Mais, l’année suivante, à l’occasion du volet français de l’enquête mondiale sur la fécondité, les démographes, mesurant le « risque de concevoir », ont mesuré la fréquence des unions de toute nature avec un échantillon aléatoire. Leur estimation de la cohabitation, phénomène en croissance rapide, s’est trouvée de 25 % inférieure à celle de l’enquête par quotas de l’année antérieure. Qu’en conclure ? Il ne faut pas fixer à une enquête par quotas les deux objectifs joints uniques : mesurer la fréquence d’un phénomène et étudier les caractéristiques des personnes concernées. Des jeunes cohabitants pouvaient paraître de très bons enquêtés pour une enquête sur la cohabitation. Les enquêteurs les avaient donc sur-sélectionnés, probablement intentionnellement. La mesure de la prévalence de ce comportement matrimonial aurait fort bien et à bas prix été préalablement mesurée dans une enquête omnibus et aurait permis d’en établir un quota pour l’enquête sociologique.

5.3 Les quotas plus précis que le sondage aléatoire ?

Des quotas bien choisis réduisent la variance de l’estimateur d’un paramètre, c’est-à-dire l’amplitude de son intervalle de confiance. Si l’échantillon est sans biais, l’échantillon sur quotas sera plutôt plus précis qu’un sondage aléatoire simple de même taille et sans calage. Reste à espérer qu’une main invisible ait bien guidé les enquêteurs. L’idéal serait d’avoir un échantillon aléatoire avec quotas. C’est ce qu’on appelle le **sondage équilibré**. Il bénéficie des avantages de l’aléatoire et des quotas, mais bien sûr ce ne sont pas les enquêteurs qui désignent sans règle les enquêtés. Il s’agit simplement d’équilibrer l’échantillon sans laisser aux enquêteurs le libre choix des enquêtés. La condition préalable à des sondages équilibrés efficaces est de disposer d’une base de sondage riche de variables bien corrélées avec la variable d’intérêt. C’est le cas des sondages répétés visant à mettre à jour leur base de sondage, typiquement des enquêtes de conjoncture : quel meilleur sondage sur les revenus fiscaux qu’un sondage équilibré issu des déclarations fiscales de l’année antérieure ?

En définitive, on ne peut échapper à s’interroger sur les différences entre un échantillon sur quotas, éventuellement redressé, et un échantillon aléatoire redressé à la suite de nombreuses non-réponses. Si le nombre de non-réponses de l’échantillon aléatoire voisine celui des contacts sans suite de l’enquête par quotas, l’existence de quotas demeure un avantage, sauf à dénombrer les non-répondants et à les remplacer par un sondage aléatoire au sein de strates très fines. Toutefois, l’enquête par quotas ajoute à l’indéfini comportement des enquêtés, celui de l’enquêteur, pas toujours très professionnel. C’est toujours un dilemme de juger à partir de quel niveau les non-réponses enlèvent leur légitimité aux échantillons aléatoires et la tendance à ne plus publier les taux de réponse soulève un réel doute.

6 Conclusion

En conclusion, il semble possible de comprendre les notions de base de la théorie des sondages sans entrer dans les développements techniques et les calculs spécifiques. Par ailleurs, les logiciels produisent des traitements d'enquêtes faites quasiment « en pilotage automatique », mais la pratique presse-bouton de ces logiciels comporte des risques… Si les développements mathématiques ne sont pas nécessaires pour tous, l'intuition des sondages doit passer par un riche rendu d'expériences**,** même pour les étudiants les meilleurs probabilistes.

Références

[1] Bachelet, D. (2007), 5ème colloque francophone sur les sondages, *Revue française de marketing*, rubrique actualité, novembre.

[2] Chiche, J. (2009), Une expérience de Sondage électoral en ligne, *Journée d'étude du groupe Enquêtes de la SFdS*, le 3 décembre 2009.

[3] Didier, E. (2009), *En quoi consiste l'Amérique. Les statistiques, le new deal et la démocratie*, La découverte, Paris.

[4] Dussaix, A.-M. et J.-M. Grosbras (1993), *Les sondages. Principes et méthodes*, Que sais-je ?, n° 701.

[5] Gauvrit, N. (2000), *Statistique, méfiez-vous*, Ellipse, Paris.

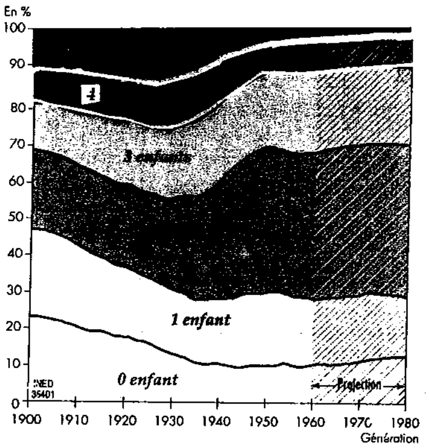
[6] Toulemon, C. (2001), Combien d'enfants, combien de frères et sœurs depuis cent ans ?, *Population et Sociétés*, INED, n°374, février.

[7] Thélot, C. (2004), *Tel père, tel fils*, Pluriel, nouvelle édition, Paris.

Annexe 1 Fratries ou descendance de la mère

Laurent Toulemon

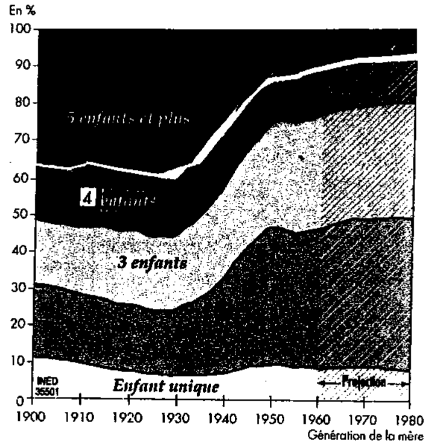
Figure 3 – Répartition des femmes selon le nombre final d'enfants



Note : La descendance par rang n'est connue de manière certaine que pour les femmes de plus de 50 ans, et avec une imprécision faible pour les femmes de plus de 40 ans (nées avant 1960). Pour les générations suivantes, la projection est fondée sur les taux de fécondité par rang de naissance des années 1996-1998 et suppose que la fécondité va rester stable à l'avenir.

Sources : Insee, état civil [2] [4], enquêtes Famille 1982, 1990 et 1999, recensement de 1946 et projections [5] [6] [9].

Figure 4 – Répartition des enfants selon le nombre d'enfants que compte leur fratrie (eux-mêmes compris)



Sources : voir figure 3.

Tableau 1 – Répartition de 100 femmes et de 100 enfants selon la taille de la fratrie, pour trois générations de femmes

Nombre d'enfants	Femmes nées en...			Taille de la fratrie (1)	Enfants dont les mères sont nées en...		
	1900	1930	1960		1900	1930	1960
0	23	13	10	0	0	0	0
1	24	18	18	1	11	7	9
2	22	26	40	2	21	19	38
3	13	18	22	3	18	20	32
4	7	10	7	4	14	16	12
5	4	6	2	5	10	11	5
6 et +	7	9	1	6 et +	26	27	4
Total	100	100	100	**Total**	100	100	100
Moyenne	2,1	2,7	2,1	*Moyenne*	4,2	4,5	2,8

* Moy G 2,7 3,1 2,3

(1) Nombre de frères et sœurs y compris l'individu de référence.

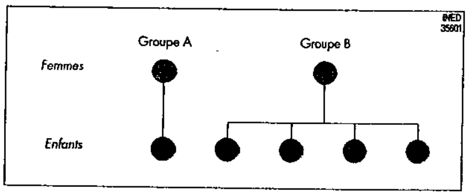
Sources : voir figure 3.

* Descendance moyenne des mères

Encadré 3

La taille de la fratrie : la famille vue par les enfants

Prenons 100 femmes A ayant chacune 1 enfant et 100 femmes B ayant 4 enfants. Les deux groupes de femmes sont numériquement à égalité, alors que l'on compte 4 fois plus d'enfants dans les fratries B que dans les fratries A (400 contre 100). De façon générale, le poids des familles nombreuses est plus élevé parmi les enfants qu'il ne l'est parmi les femmes. C'est pourquoi la taille de la fratrie (le nombre moyen de frères et sœurs y compris l'individu de référence) dans une population est supérieure au nombre moyen d'enfants chez les femmes ayant donné naissance à cette population.



	Nombre d'enfants			Nombre moyen d'enfants ou taille moyenne de la fratrie
	1 (groupe A)	4 (groupe B)	Ensemble	
Femmes	100	100	200	2,5
Enfants	100	400	500	3,4

Annexe 2 Dussaix, A.-M. et J.-M. Grosbras (1993), *Les sondages. Principes et méthodes*, Que sais-je ?, n° 701, p. 27.

Tableau 2.2. – Précision de l'estimation d'une proportion calculée à partir d'un échantillon

Proportion observée p / taille d'échantillon n	5 % ou 95 %	8 % ou 92 %	10 % ou 90 %	15 % ou 85 %	20 % ou 80 %	25 % ou 75 %	30 % ou 70 %	35 % ou 65 %	40 % ou 60 %	50 %
100					8	8,6	9,2	9,6	9,8	10
150				5,7	6,4	6,9	7,3	7,6	7,8	8
200			4,3	5,1	5,7	6,1	6,5	6,8	6,9	7,1
250	2,8	3,4	3,8	4,5	5	5,4	5,8	6	6,2	6,3
300	2,5	3,1	3,5	4,2	4,6	5	5,3	5,6	5,7	5,8
350	2,3	2,9	3,2	3,8	4,2	4,6	4,9	5,1	5,2	5,3
400	2,2	2,7	3	3,6	4	4,3	4,6	4,8	4,9	5
500	2	2,4	2,7	3,2	3,6	3,9	4,1	4,3	4,4	5
600	1,8	2,2	2,4	3	3,3	3,5	3,8	3,9	4	4,1
700	1,7	2,1	2,3	2,7	3	3,3	3,5	3,5	3,7	3,8
800	1,5	1,9	2,1	2,5	2,8	3	3,2	3,3	3,4	3,5
900	1,5	1,8	2	2,4	2,7	2,9	3	3,1	3,2	3,3
1 000	1,4	1,7	1,8	2,3	2,5	2,7	2,9	3	3	3,1
1 500	1,2	1,4	1,5	1,9	2,1	2,3	2,4	2,5	2,6	2,6
2 000	1	1,2	1,3	1,6	1,8	2	2,1	2,2	2,2	2,3
3 000	0,8	1	1,1	1,3	1,4	1,5	1,6	1,7	1,8	1,8
5 000	0,6	0,8	0,8	1	1,1	1,2	1,3	1,4	1,4	1,4
10 000	0,4	0,5	0,6	0,7	0,8	0,9	0,9	1	1	1