

Les lois du hasard

Enjeux mathématiques, historiques, citoyens.

Journée d'étude du 20 janvier 2017

Livret 2 - partie pédagogique et professionnelle

Projet de recherche financé en 2016-17 sur l'appel d'offre "recherches pour l'éducation" de la mission de recherche de l'ESPE de Créteil, qui associe les quatre universités de l'académie et l'ESPE.

Chercheurs, instituts et laboratoires partenaires :

- Centre A. Koyré, UMR 8560 (A. Bernard)
- Centre de Recherches Historiques, Univ. Paris 8 (C. Ehrhardt)
- LAGA, UMR 7539, CNRS & SPC-Univ. Paris 13 (I. Gaudron)
- LIPN, UMR 7030, CNRS & SPC-Univ. Paris 13 (S. Schwer)
- IREM Paris Nord



Présentation du livret : objectifs et usages recommandés.

Ce livret est un document de travail, destiné à accompagner le projet de recherche "les lois du hasard, enjeux mathématiques, historiques et citoyens" aux trois niveaux suivants :

- a) lors de la journée d'étude du 20 janvier 2017, il permettra d'illustrer les exposés et discussions de la matinée, destinés à faire découvrir le volet "histoire" du projet.
- b) par la suite, en association avec les enregistrements vidéos de la journée et des transparents des exposés, il constituera une ressource de travail pour les participants volontaires au projet.
- c) après les corrections et compléments éventuels, suggérés par les intervenants et participants de la journée, il constituera éventuellement une brochure qui sera éditée par l'IREM de Paris Nord et l'ESPE de l'académie de Créteil.

En fonction de ces trois objectifs, nous recommandons aux participants :

- avant la journée, de relire le descriptif du projet diffusé depuis octobre (partie A), ainsi que le texte qui clôt le livret et servira de support à la discussion finale (partie C) ;
- pendant et après la journée, de faire part de toutes les suggestions utiles pour améliorer, clarifier ou compléter ce livret (partie B), en vue de son éventuelle publication comme brochure ESPE-IREM ;
- d'exploiter librement l'ensemble de ces textes, à l'instar d'un livret de formation continue ; ils peuvent être communiqués à des collègues que vous souhaiteriez associer.

Un premier livret sur le volet "histoire" du projet complète la documentation préparatoire à la journée d'étude.

Sommaire de la brochure

A. PROGRAMME DES EXPOSÉS DU 20 JANVIER APRÈS-MIDI	3
B. DOCUMENTS ILLUSTRANT LES EXPOSÉS DU 20 JANVIER AM	5
1. Exposé d'I. Gaudron-Trouvé. Un exemple d'activité pédagogique au niveau supérieur : exploiter les données de Maurice Halbwachs pour discuter du sex-ratio et de ses enjeux.....	5
2. Exposé de P. Dutarte. Statistique et citoyenneté :une démarche interdisciplinaire soutenue par les programmes de mathématiques	7
3. Exposé de F. Delzongle. L'affaire Castaneda vs Partida: un exemple d'activité dans l'esprit de cette démarche interdisciplinaire.....	17
4. Exposé d'A.M. Hazard-Tourillon. Enjeux actuels et interdisciplinaires de l'éducation à la citoyenneté.....	17
C. PRÉSENTATION DE LA PARTIE PÉDAGOGIQUE ET PROFESSIONNELLE DU PROJET DE RECHERCHE (A. BERNARD)	19
D. BIBLIOGRAPHIE.....	23
E. ANNUAIRE DE LA JOURNÉE	24
ANNEXE: EXTRAITS DU TEXTE DE ZEISEL 1997.....	25

A. PROGRAMME DES EXPOSÉS DU 20 JANVIER APRÈS-MIDI

13h45: Sylviane Schwer, Alain Bernard: **introduction**

1. 14h-14h30 Isabelle Gaudron (Univ. Paris 13): **Un exemple d'activité pédagogique au niveau supérieur : exploiter les données de Maurice Halbwachs pour discuter du sex-ratio et de ses enjeux.**

2. 14h30-15h Philippe Dutarte (Inspection Régionale, mathématiques): **Statistique et citoyenneté :une démarche interdisciplinaire soutenue par les programmes de mathématiques**

3. 15h-15h30 Françoise Delzongle (prof de mathématiques retraitée, ex-Lycée G. Eiffel et IUFM): **L'affaire Castaneda vs Partida: un exemple d'activité dans l'esprit de cette démarche interdisciplinaire**

pause 15h30-16h

4. 16h-16h30 Anne-Marie Hazard Tourillon (Inspection générale, histoire-géographie): **Enjeux actuels et interdisciplinaires de l'éducation à la citoyenneté.**

5. 16h30-17h15 Table ronde introduite par A. Bernard: **Discussion sur l'argumentaire du volet "pédagogique et professionnel" du projet = présentation de l'argumentaire de la partie C de ce livret et discussion**

B. DOCUMENTS ILLUSTRANT LES EXPOSÉS DU 20 JANVIER AM

1. Exposé d'I. Gaudron-Trouvé. Un exemple d'activité pédagogique au niveau supérieur : exploiter les données de Maurice Halbwachs pour discuter du sex-ratio et de ses enjeux.

Introduction

Le sex-ratio, à savoir le rapport entre le nombre de garçons et le nombre de filles, dans une population donnée, et à âge donné, est une quantité qui intéresse démographes, géographes, sociologues, biologistes, et statisticiens. Nous regarderons ici le sex-ratio à la naissance, à savoir le rapport entre le nombre de garçons nés vivants et le nombre de filles nées vivantes. Au cours de recherches de données réelles susceptibles de me fournir des exercices d'examens en licence de sciences économiques, j'avais découvert, dans les quelques pages du livre « Le point de vue du nombre » de Maurice Halbwachs et d'Alfred Sauvy disponibles sur internet, une partie des données décrivant les nombres de garçons nés vivant dans différents pays d'Europe et à différentes époques. Ce livre, publié en 1936, a été republié dans une édition critique augmentée de différents articles par l'INED en 2005.

J'avais à l'époque uniquement regardé les sex-ratios à Paris et hors Paris au XIX^{ème} siècle, ainsi que les sex-ratios en France et en Allemagne en 1930 (tableaux pages 194 et 195 de [HS]). Les intervalles de confiance, très précis du fait des tailles très importantes des échantillons, permettaient de conclure dans les deux cas que la proportion de garçons nés vivant à Paris étaient sensiblement inférieure à celle des garçons nés hors Paris, et que la proportion de garçons nés en Allemagne était supérieure à celle des garçons nés en France.

Une première chose remarquable au sujet du sex-ratio est sa grande stabilité au cours du temps et dans les différentes régions du monde, qui rendent suspects tous les écarts à la norme. Dans l'introduction de son article « Recherches statistiques sur la détermination du sexe à la naissance » de 1933, M. Halbwachs indique que c'est un anglais, Graunt, qui dès 1660, à la suite de recherches poursuivies pendant 30 ans dans les registres de naissances, affirme qu'il naît en moyenne 105,8 garçons pour 100 filles. C'est à peu près la moyenne trouvée dans les années 30, dans une bonne vingtaine de pays, ce qui fait que ce sex-ratio apparaît comme à peu près constant.

Les proportions de garçons et filles à la naissance ont régulièrement intéressé les mathématiciens aux XVIII^{ème} et XIX^{ème} siècles (Laplace, De Moivre et Poisson). Ces derniers ont modélisé ces proportions à l'aide de la loi binomiale, et ont déterminé si des écarts dans ces proportions pouvaient être dus ou non au hasard. Les registres de naissance ont ainsi été une formidable base de données pour l'avancée du calcul des probabilités et de la statistique mathématique.

Présentation du livre « Le point de vue du nombre » de M. Halbwachs et A. Sauvy :

Ce livre est publié en 1936, et est en fait le troisième chapitre du septième tome d'un vaste projet éditorial qui est d'écrire l'Encyclopédie Française. Ce projet naît en 1932, dans le contexte de la crise de 1929, et de la montée du nazisme et du fascisme. Il est initié en partie par le ministre radical-socialiste de l'éducation Anatole De Monzie, et dirigé par Lucien Febvre. Parmi les encyclopédistes, on peut citer Emile Borel, Jules Romains, Georges Duhamel, Henry De Jouvenel, Paul Langevin, Paul Valéry. Les idées de réforme de l'éducation, du développement au sein de la société des nations, du rayonnement de la science y sont à l'oeuvre. Vingt et un volumes paraissent jusqu'en 1966.

Le point de vue du nombre participe donc de cette encyclopédie. Il est écrit par le sociologue M. Halbwachs, né en 1877, et le statisticien Alfred Sauvy, né en 1899. Outre ses travaux en sociologie, M. Halbwachs a publié avec le mathématicien Maurice Fréchet « Le calcul des probabilités à la portée de tous » en 1924. Très connu pour son travail sur la classe ouvrière (thèse de 1912), il est un acteur essentiel de l'introduction des statistiques (et de leur traitement rigoureux) en sociologie. Gendre de Victor Basch, il sera déporté à Buchenwald et y mourra en 1945.

Alfred Sauvy est un jeune économiste et statisticien issu de l'école Polytechnique. Il sera le directeur emblématique de l'INED de sa création en 1945 jusqu'en 1962.

Les données concernant le sex-ratio sont présentées dans le chapitre II de la section B du Point de vue du nombre. Les auteurs cherchent à expliquer les écarts non explicables par le pur hasard en cherchant un lien entre la proportion de garçons nés vivant et la différence d'âge des parents. A cet effet des données sont recueillies par M. Halbwachs dans le département du Bas-Rhin en 1925. Cette étude sera utilisée par Richard Korherr, statisticien nazi, qui mènera des enquêtes dans un but de contrôle des naissances.

Idées d'activités :

1) Intervalles de confiance :

Dans la mesure où les données présentes dans le livre portent sur des échantillons de grande taille, l'approximation gaussienne des proportions empiriques est tout à fait justifiée. Les intervalles de confiance permettent de comparer les différents sex-ratio en différents lieux, ou en différentes questions, et de poser un certain nombre de questions quant aux causes de ces disparités.

Une autre possibilité est de calculer les « P-value » pour regarder si les différentes données sont en accord avec le sex-ratio standard, et quelles sont celles qui sortent justement du cas standard.

Des simulations de différents tableaux de naissance peuvent aussi être effectuées à partir du sex-ratio standard.

2) Retour sur l'approximation gaussienne d'une proportion empirique :

L'apport des ordinateurs et logiciels de statistique (comme R) permet justement de regarder dans quelle mesure

cette approximation gaussienne est correcte dans le cas d'échantillons de taille plus réduite (par exemple de l'ordre de 1000), et pour une proportion théorique se situant autour de 0.5.

Des tableaux de simulation permettront de comparer les valeurs « exactes » des intervalles de confiance calculées avec la loi binomiale, et les valeurs approchées gaussiennes.

3) Sex-ratio versus proportion de garçons

De nombreuses études démographiques présentent les sex-ratios dans différentes régions du monde ou à différentes époques, M. Halbwachs, suite à l'ouvrage sur les probabilités écrit avec Fréchet [HF], indique et justifie que l'estimateur de la proportion de garçons, est plus précis que celui du sex-ratio. Le calcul du risque quadratique de ces deux estimateurs pourrait être effectué.

4) Lien entre le sex-ratio et l'âge des parents

Une activité, à un niveau un peu plus avancé que les précédentes, d'analyse des correspondances, pourrait peut-être permettre de regarder plus finement s'il y a un lien entre ces deux caractères.

2. Exposé de P. Dutarte. Statistique et citoyenneté : une démarche interdisciplinaire soutenue par les programmes de mathématiques

1. La genèse du groupe « mathématiques et citoyenneté » de l'IREM de Paris Nord

Le groupe « mathématiques et citoyenneté » de l'IREM de Paris Nord a été actif de 2002 à 2009 environ. Il est issu du groupe « lycées technologiques » (maintenant groupe « enseignements technologiques ») de cet IREM, lequel a une grande expérience des mathématiques appliquées, de l'interdisciplinarité et de l'enseignement de la statistique, notamment de la statistique inférentielle en sections de techniciens supérieurs. Trois conditions nous semblent avoir été propices à son apparition :

- une expertise de l'enseignement de la statistique inférentielle en sections de techniciens supérieurs ;
- une pratique régulière de l'interdisciplinarité propre aux sections technologiques ;
- une demande institutionnelle liée d'abord aux programmes de 2000 introduisant des éléments de statistique inférentielle dès la classe de seconde générale et technologique, puis aux résultats du premier tour des élections présidentielles de 2002 et au rôle des sondages.

Tout d'abord désigné comme « groupe statistique » et formé de quelques membres, il est sollicité dans le cadre de la mise en œuvre des programmes de lycées généraux et technologiques de 2000 qui introduisent, pour la première fois, des éléments de statistique inférentielle. À partir de 2005, le groupe s'étoffe et devient groupe « statistique et citoyenneté » avec la perspective de proposer aux professeurs de mathématiques des activités de classe sur cette thématique. Une proposition de stage, dont voici le descriptif, est faite pour le plan académique de formation 2005/2006 de l'académie de Créteil.

Titre	Durée	Lieu	Dates	Horaire
INDUCTION STATISTIQUE ET CITOYENNETE	9 h	Lycée F. Villon 14 av. Marc Sangnier 75014 PARIS	24/02/06 10/03/06 31/03/06	14 h – 17 h

Descriptif :

« Complément d'information sur un thème récent des programmes de mathématiques des lycées et amené à se développer. Sensibilisation des professeurs de mathématiques sur les spécificités du raisonnement inductif utilisé dans certaines méthodes statistiques et sur ses nombreuses implications dans la formation des futurs citoyens.

Echantillonnage, sondages et estimation, classements, validation statistique d'un modèle, prise de décision dans un environnement incertain. Enjeux démocratiques des méthodes d'induction statistique. »

Cette proposition n'aboutira pas, mais le champ est élargi l'année suivante à « mathématiques et citoyenneté » avec une ouverture interdisciplinaire vers l'histoire, le français et la philosophie. Sous cette forme, le stage « mathématiques et citoyenneté » sera inscrit au PAF de l'académie de Créteil en 2006/2007 et 2007/2008 à destination des professeurs de mathématiques.

Le groupe a édité une brochure fin 2007 : brochure n° 135 de la Commission inter-IREM Lycées technologiques « *Statistique et citoyenneté : le citoyen face au chiffre* ».

2. Trois exemples emblématiques

2.1 Peut-on encore croire un sondage après avril 2002 ?

Lors du premier tour des élections présidentielles de 2002, le dernier sondage publié par l'institut B.V.A. , effectué sur 1 000 électeurs le vendredi 19/04/02, prévoyait :

Jacques Chirac	19 %
Lionel Jospin	18 %

La surprise a été grande le dimanche 21/04/02 au vu des résultats, puisque Jean-Marie Le Pen figurait au second tour :

Jacques Chirac	19,88 %
Lionel Jospin	16,18 %

Doit-on considérer que le dernier sondage B.V.A. était « faux » ?

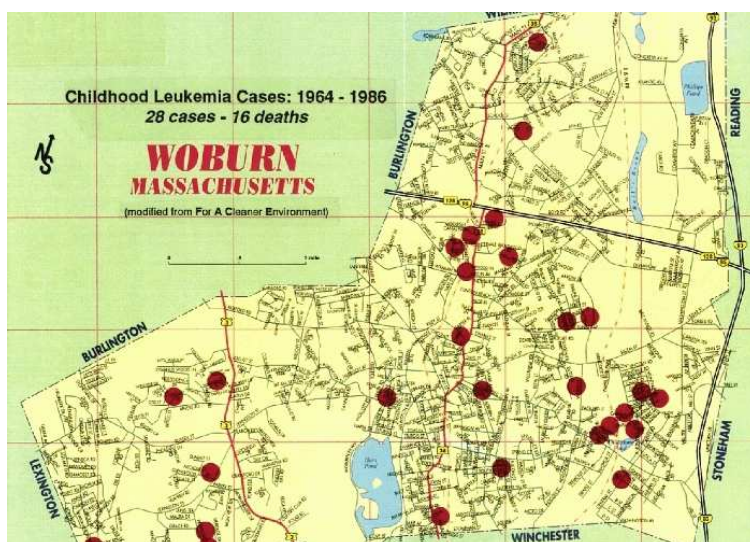


2.2 Cas de leucémie à Woburn : hasard ou pollution ?

Woburn est une petite ville industrielle du Massachusetts, au Nord-Est des Etats-Unis. Du milieu à la fin des années 1970, la communauté locale s'émeut d'un grand nombre de leucémies infantiles survenant dans certains quartiers de la ville. Les familles se lancent alors dans l'exploration des causes et constatent la présence de décharges et de friches industrielles ainsi que l'existence de polluants. Dans un premier temps, les experts gouvernementaux concluent qu'il n'y a rien d'étrange. Mais les familles s'obstinent et saisissent leurs propres experts.

Une étude statistique montre qu'il se passe sans doute quelque chose « d'étrange ».

Le tableau suivant résume les données statistiques concernant les enfants de Woburn de moins de 15 ans, pour la période 1969-1979 (Source : *Massachusetts Department of Public Health et Harvard University*).



Enfants entre 0 et 14 ans	Population de Woburn selon le recensement de 1970 n	Nombre de cas de leucémie infantile observés à Woburn entre 1969 et 1979	Fréquence des leucémies à Woburn f	Fréquence des leucémies aux États-Unis p
Garçons	5 969	9	0,001 51	0,000 52
Filles	5 779	3	0,000 52	0,000 38
Total	11 748	12	0,001 02	0,000 45

Compte-tenu de ces données, le hasard seul peut-il raisonnablement expliquer les fréquences observées à Woburn, considérées comme résultant d'un échantillon prélevé dans la population américaine ?

2.3 Preuve statistique de discrimination : l'affaire Castaneda contre Partida

En 1976 au Texas, un accusé d'origine mexicaine conteste le jugement du tribunal au motif que la désignation des jurés est discriminatoire envers les Américains d'origine mexicaine. On analyse ici les arguments statistiques et probabilistes qui apparaissent dans l'attendu de la Cour Suprême des États-Unis.



Attendu de la Cour Suprême des États-Unis (affaire Castaneda contre Partida)¹ :

« Si les jurés étaient tirés au hasard dans l'ensemble de la population, le nombre d'américains mexicains dans l'échantillon pourrait alors être modélisé par une **distribution binomiale**... Etant donné que **79,1 %** de la population est mexico-américaine, le nombre attendu d'américains mexicains parmi les **870** personnes convoquées en tant que grands jurés pendant la période de 11 ans est approximativement **688**. Le nombre observé est **339**. Bien sûr, dans n'importe quel tirage considéré, une certaine fluctuation par rapport au nombre attendu est prévisible. Le point essentiel cependant, est que le modèle statistique montre que les résultats d'un tirage au sort tombent vraisemblablement dans le voisinage de la valeur attendue... La mesure des fluctuations prévues par rapport à la valeur attendue est l'**écart type**, défini pour la distribution binomiale comme la racine carrée de la taille de l'échantillon (ici 870) multiplié par la probabilité de sélectionner un américain mexicain (ici 0,791) et par la probabilité de sélectionner un non américain mexicain (ici 0,209)... Ainsi, dans ce cas, l'écart type est approximativement de **12**. En règle générale pour de si grands échantillons, si la différence entre la valeur attendue et le nombre observé est plus grand que deux ou trois écarts types, alors l'hypothèse que le tirage du jury était au hasard serait suspect à un spécialiste des sciences humaines. Les données sur 11 années reflètent ici une différence d'environ **29** écarts types. Un calcul détaillé révèle

¹ Source : "Prove It with Figures (Statistics for Social Science and Behavioural Sciences)" - Hans Zeisel, D. H. et D. Kaye - Springer 2006.

qu'un éloignement aussi important de la valeur attendue se produirait avec moins d'une chance sur 10¹⁴⁰. »

La constitution des jurys est-elle faite au hasard ?

3. La demande institutionnelle et son évolution

La demande institutionnelle des programmes, dont ceux de mathématiques, pour une éducation du futur citoyen va croissante.

Parmi les programmes actuellement en vigueur, ceux du lycée professionnel² sont particulièrement explicites. La première phrase des programmes de mathématiques, sciences physiques et chimiques est la suivante :

« L'enseignement des mathématiques et des sciences physiques et chimiques concourt à la formation intellectuelle, professionnelle et citoyenne des élèves. »

Le domaine « statistique et probabilité » du programme contribue spécifiquement à cet objectif, notamment par son caractère interdisciplinaire. On lit ainsi dans le programme de la classe de seconde professionnelle :

« Ce domaine [statistique et probabilités] constitue un enjeu essentiel de formation du citoyen. Il s'agit de fournir des outils pour comprendre le monde, décider et agir dans la vie quotidienne. La plupart d'entre eux ont déjà été introduits au collège. Leur enseignement facilite, souvent de façon privilégiée, les interactions entre diverses parties du programme de mathématiques (traitements numériques et graphiques) et les liaisons entre les enseignements de différentes disciplines. »

Le programme de mathématiques de la classe de seconde générale et technologique³ indique les finalités suivantes, au premier rang desquelles un objectif plutôt citoyen :

« Le programme de mathématiques a pour fonction :
– *de conforter l'acquisition par chaque élève de la culture mathématique nécessaire à la vie en société et à la compréhension du monde ;*
– *d'assurer et de consolider les bases de mathématiques nécessaires aux poursuites d'étude du lycée ;*
– *d'aider l'élève à construire son parcours de formation. »*

Au collège, le socle commun de connaissances, de compétences et de culture, paru au BO n°17 du 23/04/2015, s'inscrit dans le cadre de la loi d'orientation du 8 juillet 2013 qui, en son article 13, pose le principe du socle commun :

« La scolarité obligatoire doit garantir à chaque élève les moyens nécessaires à l'acquisition d'un socle commun de connaissances, de compétences et de culture, auquel contribue l'ensemble des enseignements dispensés au cours de la scolarité. Le socle doit permettre la poursuite d'études, la construction d'un avenir personnel et professionnel et préparer à l'exercice de la citoyenneté. Les éléments de ce socle commun et les modalités de son acquisition progressive sont fixés par décret, après avis du Conseil supérieur des programmes. »

L'article 4 précise :

² BO spécial n° 2 du 19/02/2009.

³ BO 30 du 23/07/2009.

« (la formation scolaire) développe les connaissances, les compétences et la culture nécessaires à l'exercice de la citoyenneté dans la société contemporaine de l'information et de la communication. »

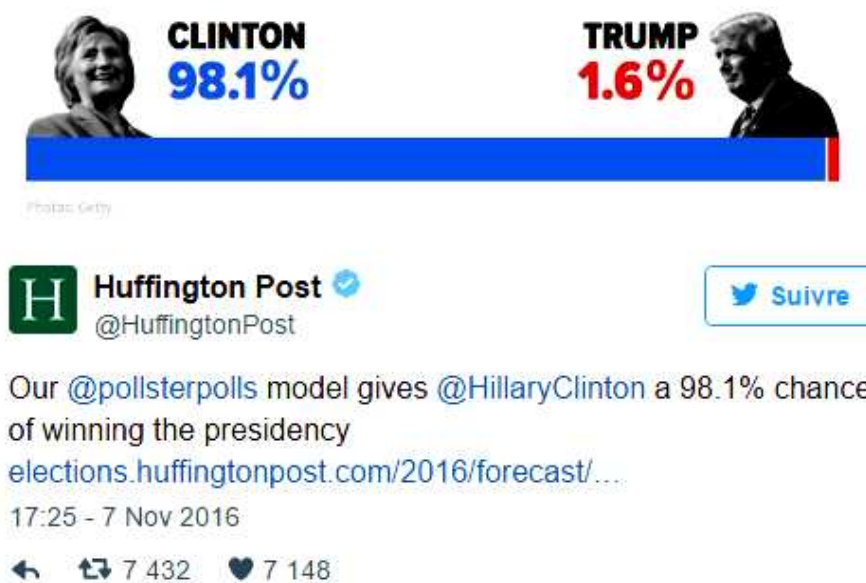
Le socle donne à la scolarité obligatoire l'objectif suivant :

« la scolarité obligatoire poursuit un double objectif de formation et de socialisation. Elle donne aux élèves une culture commune, fondée sur les connaissances et compétences indispensables, qui leur permettra de s'épanouir personnellement, de développer leur sociabilité, de réussir la suite de leur parcours de formation, de s'insérer dans la société où ils vivront et de participer, comme citoyens, à son évolution. »

Le domaine 3 du socle est celui de « la formation de la personne et du citoyen ». Il est affirmé que « L'École a une responsabilité particulière dans la formation de l'élève en tant que personne et futur citoyen. »

4. Un exemple d'actualité

Un tweet malheureux du *Huffington Post* le 7 novembre 2016, veille de l'élection américaine :



Une réponse d'un internaute dépité le 9 novembre :



La tendance des médias a été assez générale, même si certains ont été un peu plus prudents.



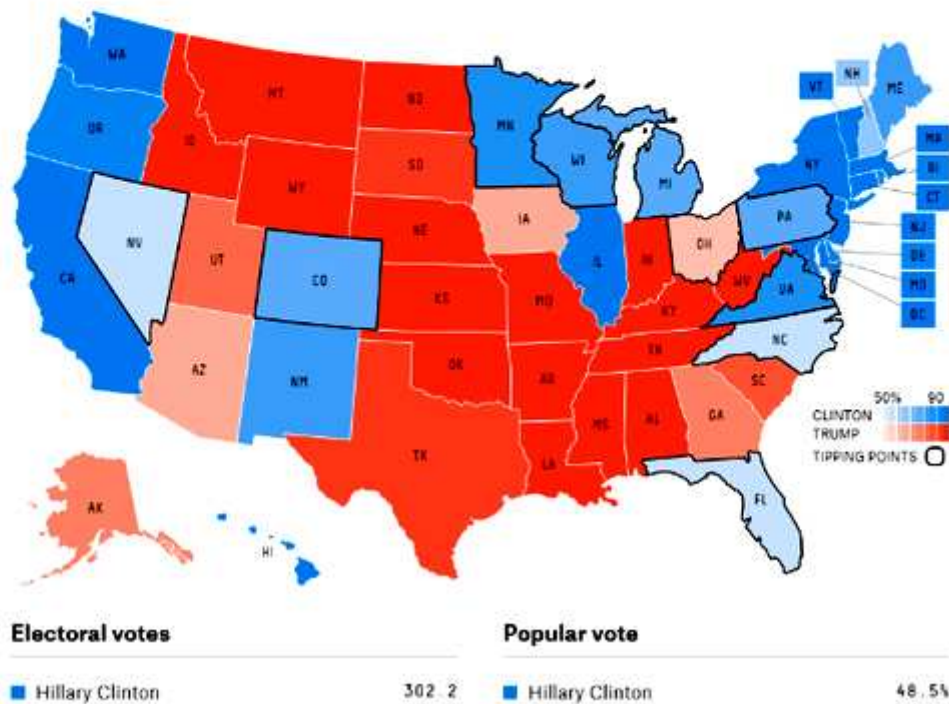
Reuters: Clinton has a 90% chance of winning

Democratic presidential candidate has about a 90 percent chance of defeating Trump, final Reuters/Ipsos States of the Nation project finds.

Reuters/Ipsos

Who will win the presidency?

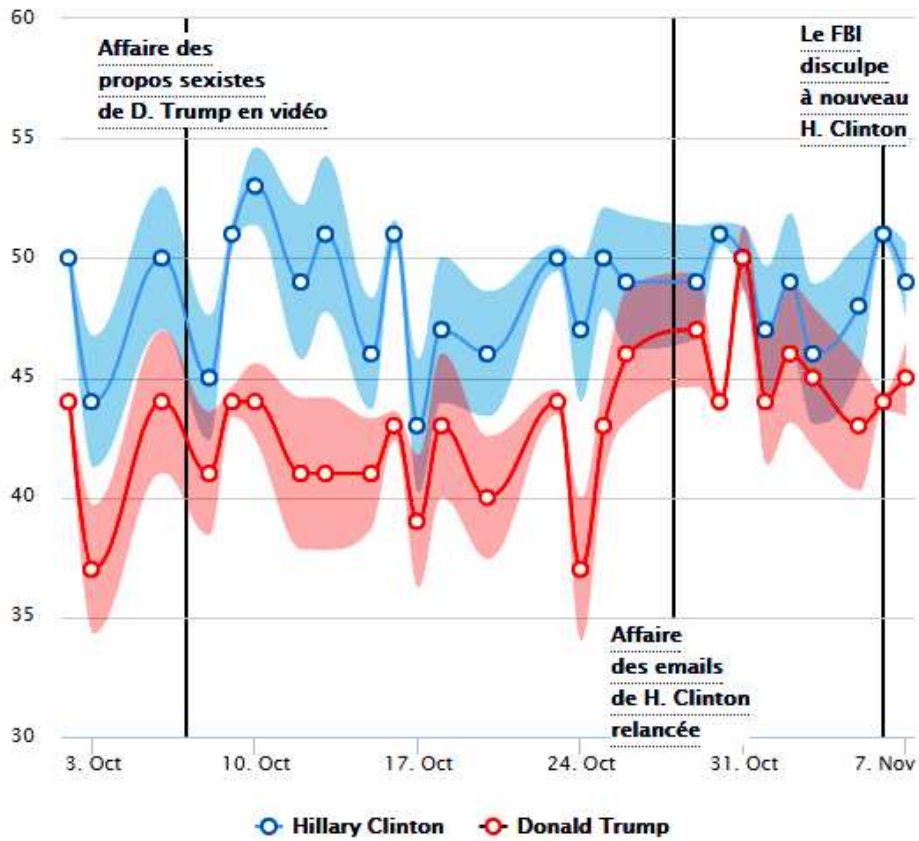
Chance of winning



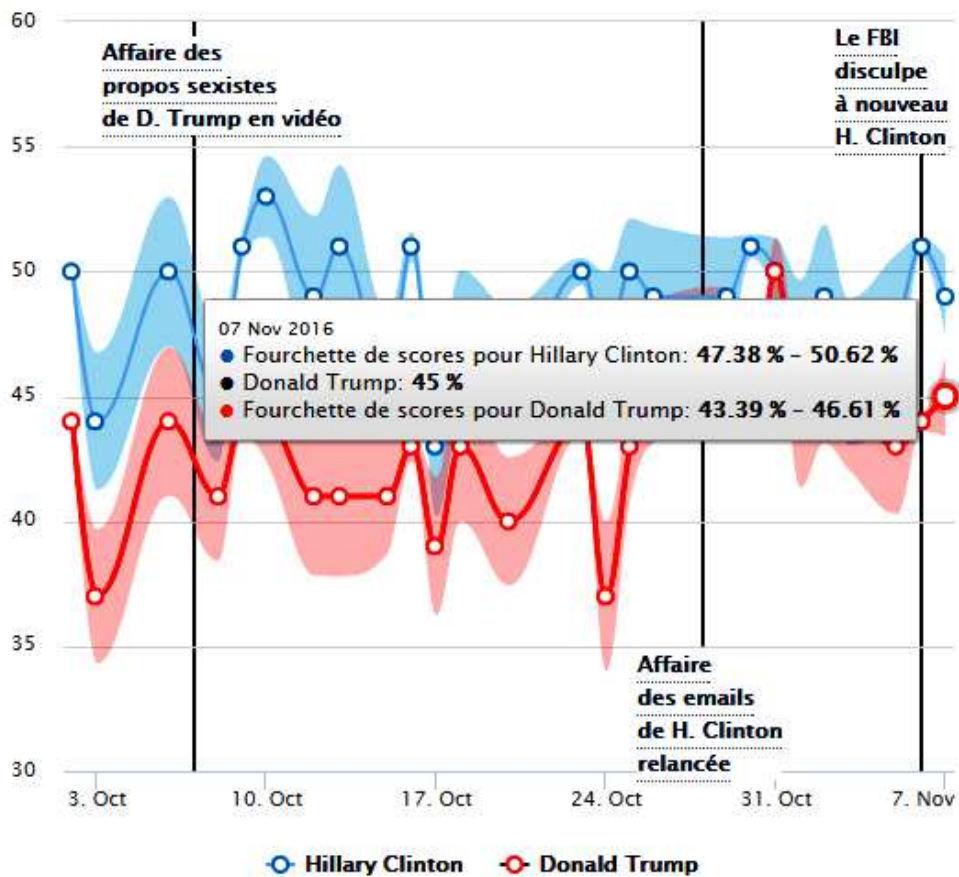
FiveThirtyEight.com

Si l'on prend cet exemple et qu'on le compare aux résultats du scrutin, on s'aperçoit que le résultat diffère dans assez peu de cas.

La prise en compte des marges d'erreurs permettait de relativiser la vision des sondages.



SOURCE : HUFFINGTON POST



SOURCE : HUFFINGTON POST

Sur le nombre de votes, les sondages ont peu failli puisqu'Hillary Clinton a obtenu la majorité des suffrages. La difficulté principale de l'estimation tient au système électoral américain et aux fameux « swing states » dont le basculement d'un côté ou de l'autre peut modifier, par le jeu des grands électeurs, le résultat de l'élection. Pour cette petite dizaine d'états, les sondages ont été assez nombreux et montraient bien une tendance très serrée la dernière semaine.

Nous indiquons ici les résultats des sondages, début novembre, de quatre de ces « swing states » : la Caroline du Nord, la Floride, la Pennsylvanie et le Wisconsin.

Source :

https://fr.m.wikipedia.org/wiki/Liste_de_sondages_sur_l'%C3%A9lection_pr%C3%A9sidentielle_am%C3%A9ricaine_de_2016

Caroline du Nord

15 grands électeurs

Démocrate en 2008 (50%-49%)

Républicain en 2012 (50%-48%)

Vainqueur: Donald Trump

Source	Date de réalisation	Donald Trump	Hillary Clinton	Différence	Échantillon	Marge d'erreur
New York Times/Siena College	4 au 6 novembre	44%	44%	Égalité	800	± 3.5 %
Quinnipiac University	3 au 6 novembre	45%	48%	3	870	± 3.3 %
Remington Research	1er au 2 novembre	48%	45%	3	2 596	± 1.9 %
Gravis Marketing	1er au 4 novembre	45%	46%	1	1 250	± 2.8 %
Trafalgar Group	27 octobre au 1er novembre	49%	44%	5	1 154	± 2.9 %
Quinnipiac University	27 octobre au 1er novembre	46%	48%	2	602	± 4.0 %

Floride

29 grands électeurs

Démocrate en 2008 (51%-48%)

Démocrate en 2012 (50%-49%)

Vainqueur: Donald Trump

Source	Date de réalisation	Donald Trump	Hillary Clinton	Différence	Échantillon	Marge d'erreur
Trafalgar Group	6 novembre	50%	46%	4	1 100	± 2.9 %
Opinion Savvy	5 au 6 novembre	46%	48%	2	853	± 3.4 %
Quinnipiac University	3 au 6 novembre	46%	46%	Égalité	884	± 3.3 %
CBS News/Yougov	2 au 4 novembre	45%	45%	Égalité	1 188	± 3.6 %
Remington Research	1er au 2 novembre	48%	45%	3	2 352	± 2.0 %
Gravis Marketing	1er au 2 novembre	45%	46%	1	1 220	± 2.8 %
Opinion Savvy	1er au 2 novembre	45%	49%	4	603	± 4.0 %
CNN/ORC	27 octobre au 1er novembre	49%	50%	1	773	± 3.5 %
Quinnipiac University	27 octobre au 1er novembre	45%	47%	2	626	± 3.9 %

Pennsylvanie

20 grands électeurs

Démocrate en 2008 (54%-44%)

Démocrate en 2012 (52%-47%)

Vainqueur: Donald Trump

Source	Date de réalisation	Donald Trump	Hillary Clinton	Différence	Échantillon	Marge d'erreur
Gravis Marketing	3 au 6 novembre	40%	46%	6	1 220	± 2.8 %
Trafalgar Group	3 au 5 novembre	48%	47%	1	1 300	± 2.7 %
Morning Call/Muhlenberg College	30 octobre au 4 novembre	42%	48%	6	405	± 5.5 %

Harper Polling	2 au 3 novembre	46%	46%	Égalité	504	± 4.4 %
Remington Research	1er au 2 novembre	45%	46%	1	2 683	± 1.9 %
Gravis Marketing	1er au 2 novembre	45%	47%	2	1 016	± 3.1 %
Susquehanna	31 octobre au 1er novembre	43%	45%	2	681	± 3.8 %
Monmouth University	29 octobre au 1er novembre	44%	48%	4	403	± 4.9 %
CNN/ORC	27 octobre au 1er novembre	46%	51%	5	799	± 3.5 %
Quinnipiac University	27 octobre au 1er novembre	44%	50%	6	612	± 4.0 %

Wisconsin

10 grands électeurs

Démocrate en 2008 (56%-42%)

Démocrate en 2012 (53%-46%)

Vainqueur: Donald Trump

Source	Date de réalisation	Donald Trump	Hillary Clinton	Différence	Échantillon	Marge d'erreur
Gravis Marketing	3 au 6 novembre	44%	47%	3	1 184	± 2.9 %
Remington Research	1er au 2 novembre	41%	49%	8	2 720	± 1.9 %
Loras College	31 octobre au 1er novembre	38%	44%	6	500	± 4.4 %

Résultat plus inattendu pour cet état mais qui se situe dans la marge d'erreur du dernier sondage.

Les sondages de la dizaine de « swing states » de 2016 montrent que les deux candidats étaient très proches lors de la dernière semaine, avec quasiment une chance sur deux, pour chacun, de l'emporter. Certains ont pu tenir le raisonnement selon lequel il était hautement improbable qu'un candidat remporte l'ensemble des « swing states » puisque

$\left(\frac{1}{2}\right)^{10} \approx 0,001$. C'était supposer qu'il y a indépendance de ces 10 événements, ce qui n'est bien sûr pas le cas.

3. Exposé de F. Delzongle. *L'affaire Castaneda vs Partida: un exemple d'activité dans l'esprit de cette démarche interdisciplinaire*

Le texte en anglais en annexe est un extrait du livre en anglais, qui a servi de source à l'activité publiée dans la brochure IREM 135.

4. Exposé d'A.M. Hazard-Tourillon. *Enjeux actuels et interdisciplinaires de l'éducation à la citoyenneté.*

Extraits des bulletins officiels:

MEN 2011 Travaux personnels encadrés - Classe de première des séries générales : mise en œuvre pédagogique à compter de la rentrée 2011. *Bulletin officiel* n° 17 du 23 avril 2015
<http://www.education.gouv.fr/cid56642/mene1116130n.html>

MEN 2015 Socle commun de connaissances, de compétences et de culture *Bulletin officiel* n° 17 du 23 avril 2015 http://www.education.gouv.fr/pid285/bulletin_officiel.html?cid_bo=87834

MEN 2015 Programme d'enseignement moral et civique, École élémentaire et collège, cycle 4. *Bulletin officiel* spécial n°6 du 25 juin 2015.
http://www.education.gouv.fr/pid25535/bulletin_officiel.html?cid_bo=90158

MEN 2015 Programme d'enseignement moral et civique, Classes de seconde générale et technologique, de première et terminale des séries générales. *Bulletin officiel* spécial n°6 du 25 juin 2015. http://www.education.gouv.fr/pid25535/bulletin_officiel.html?cid_bo=90243

MEN 2015 Enseignements au collège, Organisation. Voir §2 consacré aux EPI. Circulaire n° 2015-106 du 30-6-2015. http://www.education.gouv.fr/pid25535/bulletin_officiel.html?cid_bo=90913

MEN 2016 Enseignements primaire et secondaire - Actions éducatives - Le parcours citoyen de l'élève. *Bulletin officiel* n°25 du 23 juin 2016.
http://www.education.gouv.fr/pid285/bulletin_officiel.html?cid_bo=103533

MEN 2015 Programme d'enseignement du cycle des approfondissements (cycle 4) → voir entête et parties "mathématiques" et "histoire-géographie" *Bulletin officiel* spécial n°11 du 26 novembre 2015. http://www.education.gouv.fr/pid285/bulletin_officiel.html?cid_bo=95184 (annexe 3)

Ressources sur Eduscol:

- **Glossaire de l'EMC**, accessible depuis la page sur les principes et objectifs de l'EMC (<http://eduscol.education.fr/cid92403/l-emc-principes-et-objectifs.html>)
http://cache.media.eduscol.education.fr/file/EMC/02/5/Ress_emc_glossaire_464025.pdf
- **Introduction**, même page -
http://cache.media.eduscol.education.fr/file/EMC/15/2/Ress_emc_introduction_465152.pdf
- **Le débat réglé ou argumenté**, accessible depuis la page "principes et démarches de l'EMC"
<http://eduscol.education.fr/cid92404/methodes-et-demarches.html>
http://cache.media.eduscol.education.fr/file/EMC/01/1/ress_emc_debat_464011.pdf
- **Le débat réglé ou argumenté - pour aller plus loin**, accessible depuis la même page
http://cache.media.eduscol.education.fr/file/EMC/02/3/Ress_emc_debat-plus-loin_464023.pdf
- **Les dilemmes moraux, une méthode du développement éthique**. Même page
http://cache.media.eduscol.education.fr/file/EMC/01/3/ress_emc_dilemmes_ethique_464013.pdf
- **La discussion à visée philosophique (DVP) ou oral réflexif**. Même page
http://cache.media.eduscol.education.fr/file/EMC/01/7/ress_emc_discussion_DVP_464017.pdf

L'ensemble des textes est réuni dans un seul et même dossier dropbox accessible à l'adresse suivante:
<https://www.dropbox.com/sh/tf0x2bj2p9v1ni0/AADceHLaARDqT-EsUWz2tHloa?dl=0>

C. PRÉSENTATION DE LA PARTIE PÉDAGOGIQUE ET PROFESSIONNELLE DU PROJET DE RECHERCHE (A. BERNARD)

On décrit ici le second volet du projet de recherche "les lois du hasard: enjeux historiques, mathématiques et citoyens", qui porte sur des questions pédagogiques, et de développement professionnel des enseignants qui s'intéressent aux thématiques du projet. Il s'intéresse aux conditions grâce auxquelles les documents et réflexions élaborées dans le volet historique du projet, peuvent devenir pour le collectif ouvert des enseignants participants, des ressources de développement professionnel. Cela suppose de définir au préalable les enjeux, les perspectives de travail et de discussion de ce collectif.

1. La base du projet: un collectif interdisciplinaire de réflexion professionnelle autour de questions de citoyenneté

L'enjeu de ce collectif vise à construire des supports de scénarios d'enseignement en mathématiques et en histoire, ou les deux (scénarios interdisciplinaires type EPI ou autres), selon un modèle voisin qui pourra servir ici de point de départ: celui des supports d'enseignement construites par l'ancien groupe "mathématiques et citoyenneté" de l'IREM de Paris Nord (Dutarte et al., 2007). Ces supports d'enseignement, qui ont été en partie scénarisés, avaient déjà un caractère interdisciplinaire puisqu'ils comportaient des volets tournés vers l'éducation morale et civique (ex-ECJS) ou vers des adaptations possibles en cours de lettres. Ils renvoient en outre à un modèle commun de scénario, inspiré du format général d'une démarche d'investigation:

- Une situation déterminée à partir de d'une documentation relative à des faits contemporains (encarts de journaux, sites internet, documentation spécialisée...) est présentée puis "scénarisée" pour donner lieu à une activité significative et accessible pour des élèves. Elle implique l'usage d'outils statistiques et probabilistes, et permet d'engager avec les élèves une réflexion ayant trait à la citoyenneté.
- Un dossier complémentaire est présenté permettant d'aller plus loin et notamment d'envisager des activités de type interdisciplinaire. Certains de ces dossiers évoquent des scénarios pour des contextes qui ne sont plus ceux d'un cours de mathématiques⁴.
- Dans les deux cas les objectifs pédagogiques sont explicités et brièvement présentés, et des copies d'élèves sont parfois présentées et analysées en rapport à des expérimentations effectives.

On propose ici d'adopter pour le collectif qui prendra en charge la construction de nouvelles ressources un mode de fonctionnement semblable à celui d'un groupe IREM habituel⁵, à l'image du collectif "statistiques et citoyenneté" qui a produit la brochure en question. Il s'agit donc, pour la première année, de convenir du type de projets dans lesquels chacun pourrait et voudrait s'engager, de définir la liste des collègues impliqués et d'un nombre limité de réunions pour la mise en commun et la

⁴ Voir par exemple le complément à l'activité "Inquiétude sur Woburn" (Dutarte et al., 2007, p. 47) qui propose une activité exploitable en ECJS, à partir du film *Prejudice* qui évoque d'une manière romancée l'événement à la base de l'activité.

⁵ Un groupe IREM rassemble généralement des enseignants de tous niveaux (primaire, secondaire, universitaire), généralement de mathématiques (mais il existe des groupes interdisciplinaires) pour conduire des recherches sur l'enseignement des mathématiques et produire des ressources (stages de formation continue, brochures, publications...). Pour une image actualisée des activités des IREM, voir le portail en ligne des IREM: <http://www.univ-irem.fr> et plus précisément la page décrivant les missions et moyens: <http://www.univ-irem.fr/spip.php?article6>.

présentation des travaux en cours. Le groupe étant ouvert, on peut encore adjoindre des collègues qui ne sont pas venus le 20 janvier mais pourraient être intéressés à collaborer au projet. Un coordinateur ou une coordinatrice s'occupe du calendrier et de l'ordre du jour des réunions.

Pour la première année, le délai sera généralement trop court pour aboutir à des premières ressources complètement abouties et expérimentées: il suffira de pouvoir définir fin 2017 un premier rapport d'étape pour l'année 2 (2018), faisant état de l'avancement des projets envisagés, et de produire les données exploitables sur la genèse de ces scénarios, dont il sera question plus bas.

L'enjeu ultime de ce travail préparatoire est la constitution possible d'un nouveau collectif "mathématiques, histoire et citoyenneté", dans un cadre institutionnel à définir, et qui pourrait fonctionner dans les années suivantes (2018 et suivantes) le temps d'expérimenter les ressources et démarches envisagées dans la phase préparatoire.

En résumé, la base du projet est donc de créer les conditions collectives pour constituer des scénarios d'enseignement en mathématiques et en histoire, ou les deux (scénarios interdisciplinaires type EPI, TPE, ou autres), selon un modèle voisin des supports d'enseignement construits par l'ancien collectif "statistiques et citoyenneté" de l'IREM de Paris Nord.

2. Se donner les moyens d'étudier les genèses documentaires à l'œuvre dans le collectif.

Par rapport à ce modèle de départ, le premier aspect crucial de ce volet du projet vise à interroger la manière de construire de telles situations intéressantes pour les apprentissages mathématiques, historiques et citoyens, *en explicitant les démarches et réflexions qui y ont conduit*. Cet aspect génétique est en effet peu explicité dans la brochure qui nous sert de point de départ. Pour cette dernière on identifie bien le collectif des enseignants impliqués, et elle comporte une analyse explicite des objectifs de chaque activité. Mais les ressources utilisées ne sont que partiellement répertoriées, et les débats et les solutions alternatives qui ont été discutées en amont des expérimentations ou pendant ces dernières, ne sont pas explicités. En s'inspirant des recherches contemporaines sur les genèses documentaires des collectifs d'enseignants, à partir d'un ensemble varié de ressources (Gueudet, Pepin, & Trouche, 2011), il est donc possible d'envisager un travail de réactualisation de l'esprit de ces scénarisations, qui explicitent les aspects génétiques en amont ou dans le cœur de la construction de ces ressources.

Or ce travail paraît difficilement envisageable dans un collectif restreint qui doit en priorité se consacrer à l'élaboration efficace de ressources. Il devient par contre possible dans un collectif élargi qui comprend des chercheurs, des (ex-) étudiants MEEF, et dans lequel on peut partager les rôles. Certains membres du collectif seront en effet prioritairement intéressés à créer / élaborer des ressources et scénarios nouveaux, voire les expérimenter en fonction du temps. D'autres pourront se consacrer à des tâches d'observation raisonnée de ce travail d'élaboration. D'autres encore parmi nous qui ont une expérience de formation, pourront se consacrer à une tâche intermédiaire, qui est d'accompagner la démarche d'explicitation réflexive de la genèse des scénarios.

Pour cela il s'agit (a) de se mettre d'accord, au sein du groupe qui travaillera en cours d'année, sur qui fait quoi et sur ce qui est rendu visible: nous proposons une première nomenclature de ces rôles, à titre indicatif, dans la dernière partie (§5); (b) d'identifier au mieux l'ensemble des ressources vives mises en jeu dans le travail d'élaboration, et notamment le produit des recherches historiques.

3. Le rôle des produits de la recherche historiographique dans les genèses documentaires

Parmi ces ressources impliquées dans la construction pédagogique, il y a potentiellement les sources historiques issues du 1^{er} volet du projet de recherche. Cela nous conduit au second aspect crucial de ce dernier, dans lequel on s'intéresse à la question suivante: quel est le rôle précis des productions historiographiques (issues du 1er volet, de recherche historique) sur ce travail de genèse documentaire?

Pour comprendre ce point, il suffit de partir d'un constat simple concernant la brochure qui nous sert de premier modèle. Aucune des activités proposées n'utilise effectivement de documentation à portée culturelle et historique à titre de matériel *directement* exploité. On trouve, au mieux, des articles de portée épistémologique rédigés dans un contexte contemporains et sous une forme suffisamment accessibles pour pouvoir être exploités dans le contexte d'un enseignement type ECJS⁶. Or il est parfaitement clair, dans le cas de l'ancien groupe "mathématiques et citoyenneté", que ce n'est faute ni de connaissances ni d'intérêt pour l'histoire: les publications de certains membres (Dutarte, 2005) indiquent suffisamment que l'histoire ancienne et récente des statistiques et probabilités jouait un rôle vivant dans la construction des démarches. En outre la dimension historique est de manière générale inévitable sur les questions de citoyenneté. Mais le rapport entre ces réflexions et connaissances, et les scénarios élaborés, n'était pas un rapport de réinvestissement direct.

Au regard de la littérature sur ce type d'activité (Barbin et al., 2002), il n'est pas surprenant que ce soit pour l'essentiel la réflexion épistémologique et didactique des enseignants qui joue un rôle crucial et soit au cœur du dispositif d'élaboration des ressources. Les données historiques qui permettent de reconsidérer en profondeur les enjeux et la nature des savoirs ou des méthodes mis en jeu (en l'occurrence probabilistes et statistiques, ou extra-mathématiques), font naturellement l'objet d'une digestion, d'une "macération réflexive" dont les chemins sont multiples. Un de ces chemins est la discussion collective, d'autant plus importante quand il s'agit d'activités en partie interdisciplinaires qui supposent de s'entendre, de bâtir des éléments d'intercompréhension d'une discipline à l'autre. L'autre chemin est le partage de lectures spécifiques et la co-élaboration de ressources visant à partager le savoir.

4. Étudier concrètement les genèses documentaires: outils méthodologiques envisageables.

La question qui doit donc nous intéresser ici est justement d'étudier plus précisément certains de ces chemins, en postulant que ce travail de macération prend trois voies principales, dont chacun demande à être étudiée et documentée à l'aide d'outils spécifiques:

- i. les échanges "à vif", notamment les controverses d'interprétation, à l'occasion de réunions plénières ou d'exposés; c'est ici qu'on peut suivre le travail complexe du dialogisme mis en évidence (entre autres) dans les travaux de psychologie du travail (Clot, 2008, pp. 203–230)⁷;
- ii. la reprise, au prix éventuel d'un détournement, de certains outils d'exploration historiographique: documentation primaire (corpus de textes publiés, extraits de correspondances), répertoires raisonnés de textes; schémas ou tables prosopographiques, notices biographiques et littérature secondaire (articles existants);
- iii. les documents de synthèse rédigés ou préparés par les membres, pour faire le point sur tel ou tel aspect de la recherche et notamment ceux qui indiquent, à l'image de l'introduction ou de la conclusion du texte de C. Ehrhardt (1er volet, p.21-24), les enjeux épistémologiques ou en

⁶ Voir, par exemple, l'article de 2003 sur la notion de "risque d'inondation" cité en note complémentaire à l'activité proposée sur la crue décennale (Dutarte et al., 2007, p. 63).

⁷ Le postulat fondamental est que les documents mis en circulation feront l'objet le plus souvent de divergences d'interprétation qui sont dues en partie aux différences de perspective sur la documentation.

général philosophiques du projet. Cette documentation peut prendre l'aspect d'interventions enregistrées, à l'image des vidéos mises à disposition après la journée du 20 janvier. Elle peut prendre aussi la forme de documents de synthèse circulant au sein d'une partie du groupe.

Chacun de ces grands aspects deviennent accessibles par des outils méthodologiques spécifiques: (i) les échanges dialogiques sont en partie analysables à partir de transcriptions des échanges, ou bien à partir des traces que laissent ou peuvent laisser *après coup*, dans les entretiens avec des participants ou dans leurs écrits, ces discussions; (ii) le devenir des outils de la recherche historique est analysable à partir du moment où on peut identifier ces réutilisations. De là vient la nécessité de répertorier le mieux possibles les ressources effectives du travail d'élaboration, même quand il s'agit de ressources non publiées mais en circulation effective au sein du groupe (des entretiens ciblés peuvent avoir cette fonction d'identification); (iii) enfin le devenir des documents devient plus aisément accessibles si les membres du groupe, du moins les membres volontaires, livrent leurs réflexions sous différentes formes: notes de lecture, entretiens, ou encore des textes de réflexions sur la pratique enseignante, plus construits. Un tel exemple de texte de ce genre est accessible sur le billet rédigé par une des membres de l'actuel groupe "histoire et épistémologie" de l'IREM Paris Nord (Szmigel, 2016). On peut aussi penser aux écrits thématiques élaborés dans l'ancienne instance INNOVALO⁸ du rectorat, et qui supposait un accompagnement spécifique en écriture.

5. Rôle possibles au sein du groupe de travail

Le projet, pour être viable, doit donc conjuguer de manière équilibrée des rôles de conception et d'observation. Bien sûr les deux ne sont pas mutuellement exclusifs, *l'observation participante* étant une catégorie reconnue en sciences sociales. Mais il n'est généralement pas viable de prétendre vouloir tout faire à la fois, chaque collègue intéressé-e à participer doit plutôt proposer une contribution raisonnable et proportionnée au temps qui lui est réellement disponible. Cela revient à se définir un rôle raisonnablement identifié dans le projet collectif, quitte à ce que dernier évolue. Cette question est importante pour les étudiants MEEF ou néotitulaires, qui doivent conjuguer leur participation à beaucoup d'autres contraintes. Voici une liste indicative de possibilités:

- Rôle de contributeur / -trice à l'élaboration de nouvelles ressources pédagogiques;
→ *potentiellement membre ultérieurement d'un nouveau collectif "mathématiques-histoire-citoyenneté", dans un cadre IREM ou autre, si la phrase probatoire de 2017 donne lieu à un projet à plus long terme*
- Rôle d'observateur / -trice du travail de genèse documentaire: travail d'observation des échanges, de l'élaboration d'une réflexion commune; de repérage des questionnements et "objets frontières" qui circulent dans ce travail
→ *peut bien convenir à des étudiants MEEF ou néotitulaires, qui pourront ainsi contribuer au projet de recherche et enrichir leur "répertoire professionnel". Leurs observations et questions pourront aussi stimuler des questionnements chez les collègues*
- Rôle d'accompagnateur / -trice du type de travail d'élaboration réflexive évoqué plus haut, et permettant de retracer le cheminement heuristique conduisant à de nouvelles ressources
→ *peut bien convenir à des collègues ayant une expérience / expertise de formateur*
- Rôle de coordinateur / chercheur → *pour la coordination des recherches, la définition des problématiques, les enquêtes en archives ou sur le terrain, et la rédaction + valorisation des résultats (articles, conférences, rédaction de nouveaux projets en phase 2..)*

⁸ ex-Mission innovation / valorisation du rectorat de l'académie de Créteil

D. BIBLIOGRAPHIE

NOTA : cette bibliographie a été établie sous le logiciel zotero, pour les personnes qui en ont l'usage il est possible de demander à rejoindre le groupe **Lois du hasard** pour accéder à une version électronique de la bibliographie, et participer à son enrichissement.

Bibliographie de la journée d'étude (AM - livret 2).

- [AR] ARBUTHNOT, JOHN (1710) « An argument for Divine Providence, taken from the Constant Regularity observ'd in the Births of both sexes », *Philosophical Transactions*, n° 328, 1710, vol. 27 pour les années 1710-1712.
- BARBIN, E., BAGNI, G. T., GRUGNETTI, L., KRONFELLNER, M., LAKOMA, E., & MENGhini, M. (2002). Integrating history: research perspectives. In J. Fauvel & J. V. Maanen (Eds.), *History in Mathematics Education* (pp. 63–90). Springer Netherlands. Retrieved from http://link.springer.com/chapter/10.1007/0-306-47220-1_3-[DM] DE MOIVRE (1756) « The doctrine of chances », 1756.
- CLOT, Y. (2008). *Travail et pouvoir d'agir*. Paris: Presses universitaires de France.
- [DP] POISSON (1830) « Mémoire sur la proportion des naissances des filles et des garçons.... », *Mémoires de l'Académie royale des sciences de l'Institut de France* », t. IX, Paris, Firmin-Didot, 1830, p 239-308.
- DUTARTE, P. (2005). *L'induction statistique au lycée: illustrée par le tableur*. Paris, France: Didier, 2005.
- DUTARTE, P., DELZONGLE, F., MAATI, H., CARDINAL, J.-P., COUPRY, A., & DHERISSARD, S. (2007). *Statistiques et citoyenneté, le citoyen face au chiffre. IREM de Paris Nord*. <http://irem.statistiques.free.fr/telechargement/brochure135.pdf>
- [FH] MAURICE FRÉCHET ET MAURICE HALBWACHS, 1924, *Le calcul des probabilités à la portée de tous*, Dunod Paris.
- GUEUDET, G., PEPIN, B., & TROUCHE, L. (2011). *From text to “lived” resources: mathematics curriculum materials and teacher development*. Dordrecht; New York: Springer.
- [H] MAURICE HALBWACHS, « Recherches statistiques sur la détermination du sexe à la naissance », *Journal de la Société de statistique de Paris*, 74^e année n°5, mai 1933.
- [HS] MAURICE HALBWACHS ET ALFRED SAUVY, *Le point de vue du nombre*, 1936, édition critique sous la direction de Marie Jaisson et Éric Brian, INED, 2005.
- [L] LAPLACE, *Essai philosophique sur les probabilités*, Paris, 1814.
- SZMIGEL, C. (2016, September 26). *L'évolution d'un enseignement du théorème de Pythagore vers une démarche par problèmes* [Carnet de recherche]. <https://problemata.hypotheses.org/462>, accessible sous le mot de passe pythagore
- ZEISEL, H., KAYE, D. H., & WEINSTEIN, J. B. *Prove it with Figures: empirical methods in law and litigation*. Etats-Unis d'Amérique: Springer-Verlag New York. 1997

E. ANNUAIRE DE LA JOURNÉE

Ce tableau vous donne la possibilité de contacter l'une ou l'autre personne parmi les organisateurs, intervenants ou participants de la journée du 20 janvier .

Nom, prénom	Affiliation, établissement	Coordonnées
BERNARD, Alain	Centre A. Koyré, UMR 8560, UPEC-ESPE, IREM Paris Nord	alain.bernard@u-pec.fr ou alainguy.bernard@gmail.com tél 06.29.37.60.93
BERTHOUET Joël	Prof. mathématiques. Lycée Pablo Picasso, Fontenay-ss-Bois	joel.berhouet@ac-creteil.fr
CHABOT DÉJÀ Flavie	Prof stagiaire / étudiante MEEF - Lycée G. Bachelard, Chelles	flavie.chabotdeja@gmail.com
DELZONGLE Françoise	Prof et formatrice mathématiques retraîtée	Delzonglemaths@gmail.com
DENOMME Élise	Prof. Histoire-Géographie, Collège Monthety, Pontault-Combault	elisedenomm@gmail.com
DURAND Loanne	Prof. mathématiques, Lycée Paul Doumer, Le Perreux sur Marne	loanne.durand@gmail.com
DUTARTE Philippe	IPR mathématiques, rectorat de Créteil	dutarte@gmail.com
EHRHARDT, Caroline	Centre de Recherches Historiques, Univ. Paris 8	caroline.ehrhardt@univ- paris8.fr
GAUDRON, Isabelle	LAGA, UMR 7539, CNRS & SPC- Univ. Paris 13	gaudron@math.univ- paris13.fr
GLEBA Fabienne	Prof. mathématiques, Collège Delattre - Le Perreux sur Marne	Fabienne.Gleba@ac-creteil.fr
GOSZTONYI, Katalin	Univ. Eötvös Loránd, Budapest	katalin.gosztonyi@gmail.com
LE COENT Sébastien	Prof. mathématiques - Collège Gustave Monod Vitry-sur-Seine	sebastienlecoent@gmail.com
L'HOSPITALIER Franck	Prof. histoire-géographie, Lycée Sonia Delaunay Cesson	flhospitalier@gmail.com
MAZLIAK, Laurent	Laboratoire de Probabilités et Modèles Aléatoires (UMR 7599), Univ. Pierre et Marie Curie	laurent.mazliak@upmc.fr
MONNET CHALOIN Pascale	Formatrice Histoire Géographie ESPE de Créteil	pascale.monnet-chaloin@u- pec.fr
NORMAND DISPAGNE Caroline	Prof. histoire-géographie, collège Louis Issaurat, Créteil	caroline.dispagne@gmail.com
SCHWER, Sylviane	LIPN, UMR 7030, CNRS & SPC- Univ. Paris 13, IREM Paris Nord	schwer@lipn.univ-paris13.fr
HAZARD-TOU- RILLON Anne-Marie	IPR Histoire-Géographie, rectorat de Créteil	Anne-Marie.Tourillon@ac- creteil.fr

These goals require a number of administrative steps.¹⁰ The first one is to compile a list of all eligible persons. In each state and in the federal system, statutes define the qualifications of persons eligible to serve as jurors. For example, persons who have been convicted of a felony and who are unable to speak or read English usually are not eligible. A court clerk, judge, or specially appointed jury commissioners prepare a list of the eligible jurors within the jurisdiction.¹¹ The names on this master list may come from voter registration lists and other sources.¹² The resulting jury “pool” or “wheel” is the population from which smaller “venires” or “panels” are summoned to appear in court.

Once the proper list is compiled, the second task begins — equitable random selection from that list. This is usually done through so-called systematic sampling: every n th name on the list is picked. The number n is determined by the foreseeable number of jurors the system may need. Normally these persons receive a communication or summons through the mail. At this point, the second difficulty of equitable selection begins because mail questionnaires have a notoriously low response rate. Some citizens respond immediately to the invitation to serve; the majority of recipients needs coaxing.¹³ In this respect, court practices differ. Some courts eventually bring over 90% of the originally notified jurors into court; a decreasing number of circuits stop their efforts after the first, voluntary response wave.

There is another, seldom discussed, requirement of the random selection principle. In earlier times, jury commissioners loved to reinvoke (after the completion of the service-free interval guaranteed by most jurisdictions after jury service) the jurors who had served before. These veteran jurors required little processing time when they returned. But that procedure has the effect of concentrating jury service to a docile minority of the population instead of spreading it widely. Occasional polls inquiring about jury experience reveal with uniformity that only 10% to 20% of our adult population have ever served on a jury. Many more could serve.

Systematic Discrimination. There is a considerable body of law on the way to select the jury venires from the jury pools. Various constitutional constraints imply that distinctive segments of the community must not be excluded, in whole or part, as venires are drawn.¹⁴ Such exclusion might be proved directly, by inquiring into the mechanism of the selection procedure itself. For example, in *Bush v. Kentucky*,¹⁵ a statute expressly excluded African-Americans from jury service; in *Norris v. Alabama*,¹⁶ the jury commissioners themselves testified that they “never discussed” the

names of any African-Americans for inclusion on the jury roll; and in *Cassell v. Texas*,¹⁷ jury commissioners virtually admitted that they limited the number of African-Americans to one per grand jury panel. In more modern cases, however, the alleged exclusion is less blatant, and the proof of exclusion may be less direct. While there may be proof that the officials responsible for selecting venires had the opportunity to discriminate against a racial or other cognizable group,¹⁸ evidence of the outcome of the selection procedure can help support an inference of bias in the operation of the procedure — even when this bias is not observed directly. Consistent underrepresentation of a racial or similar group on the venires implies that members of this group have less of a chance of selection — that they have been excluded, at least in part, from jury service.¹⁹

How, then, does one prove systematic underrepresentation? In this regard, the use of probability theory to see whether the system for selecting jurors is working properly has become well established. To illustrate this type of analysis, we turn to a leading case, *Castaneda v. Partida*.²⁰ In a southern Texas border county Rodrigo Partida was indicted and convicted for burglary with intent to rape. He then attacked this conviction,²¹ alleging that the Texas “key man” system for impaneling grand jurors discriminated against Mexican-Americans. Under this system, the jury commissioners compiled lists, and the county judge interviewed these potential jurors under oath to ensure that they met such statutory criteria as literacy and “good moral character.” Partida relied on census data and court records to show that over an 11-year period 39% of the persons summoned to grand jury duty had Spanish surnames, compared to 79% of the general adult population. Furthermore, over the 2.5 years during which the judge in Partida’s case was involved in the selection of grand jurors, the proportion of Spanish-surnamed grand jurors was 56%. Referring to the discrepancy between the population proportion and the proportion over 11 years of venires, the Supreme Court wrote:

The mathematical disparities that have been accepted by the Court as adequate for a prima facie case have all been within the range presented here. For example, in *Whitus v. Georgia* . . . , the number of Negroes listed on the tax digest amounted to 27.1% of the taxpayers, but only 9.1% of those on the grand jury venire. This disparity was held to be sufficient to make out a prima facie case of discrimination. See *Sims v. Georgia* . . . (24.4% of tax lists, 4.7%

of grand jury lists); *Jones v. Georgia* . . . (19.7% of tax lists, 5% of jury list). [T]he proof in this case was enough to establish a prima facie case of discrimination against Mexican-Americans in the Hidalgo County jury selection.²²

As the Court's citation of prior cases indicates, the prevailing approach to identifying systematic underrepresentation was to appraise, according to some unspecified criteria, the difference in the proportion of the protected class in the total population and the proportion called for jury service over a longish period of time.

There are a number of difficulties with this approach. For one, this difference in the proportions is somewhat removed from the quantity of legal interest — the degree to which the chance or odds of being called for jury duty are impaired. Some statistics that better convey this idea have been proposed²³ but are not yet in general use.²⁴ Second, while the intuition that “[i]f a disparity is sufficiently large, then it is unlikely that it is due solely to chance or accident”²⁵ is appealing, the approach leaves judges who must decide when a disparity is “sufficiently large” entirely at sea.

The majority opinion in *Castaneda*, however, offers a solution to the second problem. In a footnote to the text quoted above, the Court wrote:

If the jurors were drawn randomly from the general population, then the number of Mexican-Americans in the sample could be modeled by a binomial distribution. . . . Given that 79.1% of the population is Mexican-American, the expected number of Mexican-Americans among the 870 persons summoned to serve as grand jurors over the 11-year period is approximately 688. The observed number is 339. Of course, in any given drawing some fluctuation from the expected number is predicted. The important point, however, is that the statistical model shows that the results of a random drawing are likely to fall within the vicinity of the expected value. . . . The measure of the predicted fluctuations from the expected value is the standard deviation, defined for the binomial distribution as the square root of the total number in the sample (here 870) times the probability of selecting a Mexican-American (0.791) times the probability of

selecting a non-Mexican-American (0.209). . . . Thus, in this case the standard deviation is approximately 12. As a general rule for such large samples, if the difference between the expected value and the observed number is greater than two or three standard deviations, then the hypothesis that the jury drawing was random would be suspect to a social scientist. The 11-year data here reflect a difference of approximately 29 standard deviations. A detailed calculation reveals that such a substantial departure from the expected value would occur by chance is less than 1 in 10^{140} .²⁶

This footnote may do more to impress the reader with its erudition than to explicate how probability theory can assist in ascertaining whether a discrepancy is “sufficiently large.” The underlying idea is to compute the probability of selecting so few Mexican-Americans on the assumption that every Mexican-American has a known chance (0.791) of appearing on a grand jury venire every time a person is selected for a venire. In other words, the Court models the selection process as a series of coin flips in which each selection of a grand juror is like an independent flip of a weighted coin that has an unchanging probability, 0.791, of heads. There have been 870 tosses of this coin, and only 339 heads. How often would such a coin produce so few heads?

In principle, the probability for 339 or fewer heads can be calculated directly. We want to know the probability of no heads, of one head, of two heads, and so on, through 339 heads. The probability for no heads is easily found. This event is composed of a series of identical, independent outcomes — no heads 870 times in a row: $(1-.791)^{870} = (.209)^{870}$. The probability for one head is only slightly more complicated. We can obtain exactly one head any one of 870 ways — a head on the first toss then 869 tails, a tail on the first toss followed by a head on the second followed by 868 tails, etc. In other words, there are 870 disjoint possibilities, and each has the probability $.791(1-.791)^{869}$ associated with it. So the probability for one head is $870(.791)(.209)^{869}$, which is somewhat larger than the previous probability. A pattern is beginning to emerge. The probabilities for each possible number of heads form the “binomial distribution” of which the Court speaks. In the end, we could compute the 340 binomial probabilities and add them up.

The Court does not talk in precisely these terms because it uses a shortcut to figure the probability of 339 or fewer heads. Instead of computing the binomial distribution to arrive at an exact answer, it pretends that the number of heads follows what is called a "normal distribution" with a mean of $870(.791) = 688.2$ and a standard deviation of $[870(.791)(.209)]^{1/2} = \sqrt{143.8} = 12$. It then relies on the properties of the normal curve to conclude that the probability in the trailing end²⁷ of the normal distribution (corresponding to as few as 339 heads) is $1/10^{40}$.

The details of the calculation are not of great interest here. For the numbers in question, the binomial and normal distributions are practically indistinguishable, so the Court's approximation is quite accurate. But one should not lose sight of the fact that the use of the normal distribution with a specified mean and standard deviation is merely a convenient device for calculating some numbers. The important value is not the number of standard deviations, but the probability, often called a P-value, of such extreme results given the probability model of the selection process.²⁸

In *Castaneda*, that probability is $1/10^{40}$. Now 10^{40} is an incomprehensibly huge number. Even if we could flip the weighted coin 870 times in a mere second, the odds are that we would have to wait many times the current age of the universe to see as few as 339 heads. The nearly irresistible inference is that grand jurors are not selected in the manner posited: the simple coin-flipping model may be inapposite, selection may not be random, or the individual probability in the binomial model may have a value lower than 0.791.

Notice that the conclusion is not that discrimination necessarily is present. The state could agree that Mexican-Americans have a selection probability less than the population proportion 0.791, yet deny any discrimination against this group. Thus, three dissenting justices in *Castaneda* argued that:

Respondent offered no evidence whatever with respect to other basic qualifications for grand jury service. The statistics relied on in the Court's opinion suggest that 22.9% of Spanish-surnamed persons over age 25 in Hidalgo County have had no schooling at all. . . . Since one requirement of grand jurors in Texas is literacy in the English language, approximately 20% of adult-age Mexican-Americans are very likely disqualified on that ground alone.²⁹

The point is quite general. The population of eligible jurors will almost never be identical to the adult community as a whole. Valid statutory disqualifications for jury duty, such as conviction of a felony, usually fall more heavily on some identifiable groups than others. Consequently, an analysis that rests on general population figures could be misleading. Since the law requires something akin to random selection only from the eligible population, ideally the degree of underrepresentation should be determined with regard to this population. Indeed, some courts have insisted that where the venires are drawn from registered voters, the population proportion must be taken from the voting rolls rather than the total adult population as counted in the last census.³⁰

In cases like *Castaneda*, probability models have proved informative and are now widely accepted. Once it is agreed that underrepresentation can signal discrimination, it becomes all but impossible to avoid some numerical measure of the degree of underrepresentation. From there, it is a short step to P-values (or related quantities)³¹ to help the judge evaluate the import of the measured underrepresentation. Small P-values mark events that are improbable if the assumptions that underlie their computation are correct. Although the improbable does happen from time to time, observing such improbable events should prompt us to question the assumptions that make them seem so improbable. Proof of very small P-values for substantial discrepancies in the chances of being called for jury service in a system where the opportunity to discriminate is present may amount to a prima facie case, as in *Castaneda*.

Sometimes a different approach than that employed in *Castaneda* is appropriate. The analysis in *Castaneda* pools all venires over some long time period and asks how frequently the observed degree of underrepresentation as seen in these pooled venires would arise if the selection process were independent of race, as the law requires. If there are, say, 50 venires with a total of 3,000 persons summoned over the time period selected for study, then the observed proportion is a weighted average of the proportions in each of the 50 panels. An analysis of the distribution of these proportions in the separate venires, however, may reveal a more subtle pattern of discrimination not apparent in the pool as a whole. For instance, the jury that in 1968 convicted the well-known pediatrician Benjamin Spock of advocating the destruction of draft cards was devoid of women, largely because the venire from which this jury was formed contained only nine women. The lack of women was of special concern, since public opinion polls suggested that antiwar sentiment was more

common among women than men at that time. Allegedly, the jurors were drawn more or less at random from the list of residents over age 21. Thus, the nine women posed a puzzle: How could such a selection reduce the proportion of women from over 50% to 9% in the jury venire?

On this ground the defense challenged the array of jurors, and I (H.Z.) undertook an inquiry into this perplexing sex distribution.³² Lists of 46 jury venires summoned for trials before the seven judges of the district court in the two-and-a-half year period before the trial showed that one judge — the *Spock* judge — had remarkably fewer women in the venires for his trials than any other judge. The probability that this judge would have had venires at least this different if his venires had been selected in the same way as those of his colleagues was approximately 10^{-18} . Although Dr. Spock was acquitted on appeal, the court of appeals did not rely on this probability, but grounded its decision in the grand design of the First Amendment.³³

Cases like *Castaneda* and *Spock* show that a simple probability model of juror selection can lead to powerful demonstrations that the actual system of choosing jurors departs from random selection.³⁴ Because the model of random selection is usually an apt description of how the process of drawing potential jurors *should* operate, comparison of the observed outcomes to those predicted by this probability model is a standard method of proof in jury discrimination cases. The calculated probabilities can shed light on how jurors actually were selected and, in some cases, place the burden of proving nondiscriminatory selection on the government.

12.3 *Selecting from the venire*

A distinguished Swedish judge once remarked at an international conference on trial by jury that Americans take more trouble than any other nation to make the jury representative of the community, and at the last moment before trial we allow this representativeness to be destroyed.¹ There is truth in this reproach. Three types of removals take place at the beginning of the trial: (1) supplementary excuses for hardship threatened by the circumstances of the particular trial, its unusual length, and perhaps the need for sequestration;² (2) exclusions for cause, if a prospective juror is judged to be incapable of being impartial in that particular trial; and (3) exclusions of jurors who, for reasons that need not be articulated, are not acceptable to a party.