LA DISTRIBUTION DES NOMS DE FAMILLE

COMME OUTIL POUR L'ANALYSE DES
DINAMIQUES MIGRATOIRES

Avant-propos

- L’étude de l’origine et de la diffusion des noms de famille est une thématique fortement interdisciplinaire. Pas seulement les sciences historiques et démographiques peuvent y contribuer, mais aussi la génétique, qui a étudié pour longtemps la distribution des noms, et la physique statistique, qui s'est approchée à cette matière dans le dernier décennie.

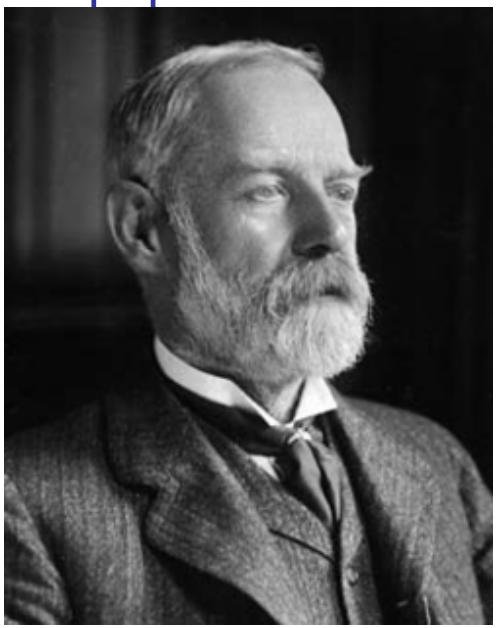
Introduction

- L’objectif de mon exposé est la présentation des principaux résultats obtenus jusqu'ici dans les études de la génétique, mais aussi des thèmes et des techniques mises récemment en champ par la physique statistique, en indiquant les informations les plus importantes qui peuvent être obtenues grâce à ces méthodologies et en illustrant les résultats avec des exemples concrets.

- L’emploi des noms de famille dans les études de biologie humaine a été proposé pour la première fois pendant la deuxième moitié du XIX siècle, mais l’étude systématique a été mis en train seulement dans les années Soixante du XX siècle et peut être reconduit à deux thèmes principaux, celui des études de consanguinité et celui de la distribution en fréquence des noms de famille et de sa possible origine.

Isonymie et consanguinité (XIX siècle)

- L’idée d'étudier le degré de consanguinité au moyen des noms de famille remonte à George Darwin (fils de Charles) qu'en 1875 calcula (pour l’Angleterre) la probabilité que dans un mariage les conjoints pouvaient avoir fortuitement le même nom de famille et la confronta avec les données empiriques, en concluant que la consanguinité dû aux mariages parmi des cousins rejoignait le 4,5% dans l’aristocratie, avec des valeurs plus basses pour la bourgeoisie et la population urbaine.

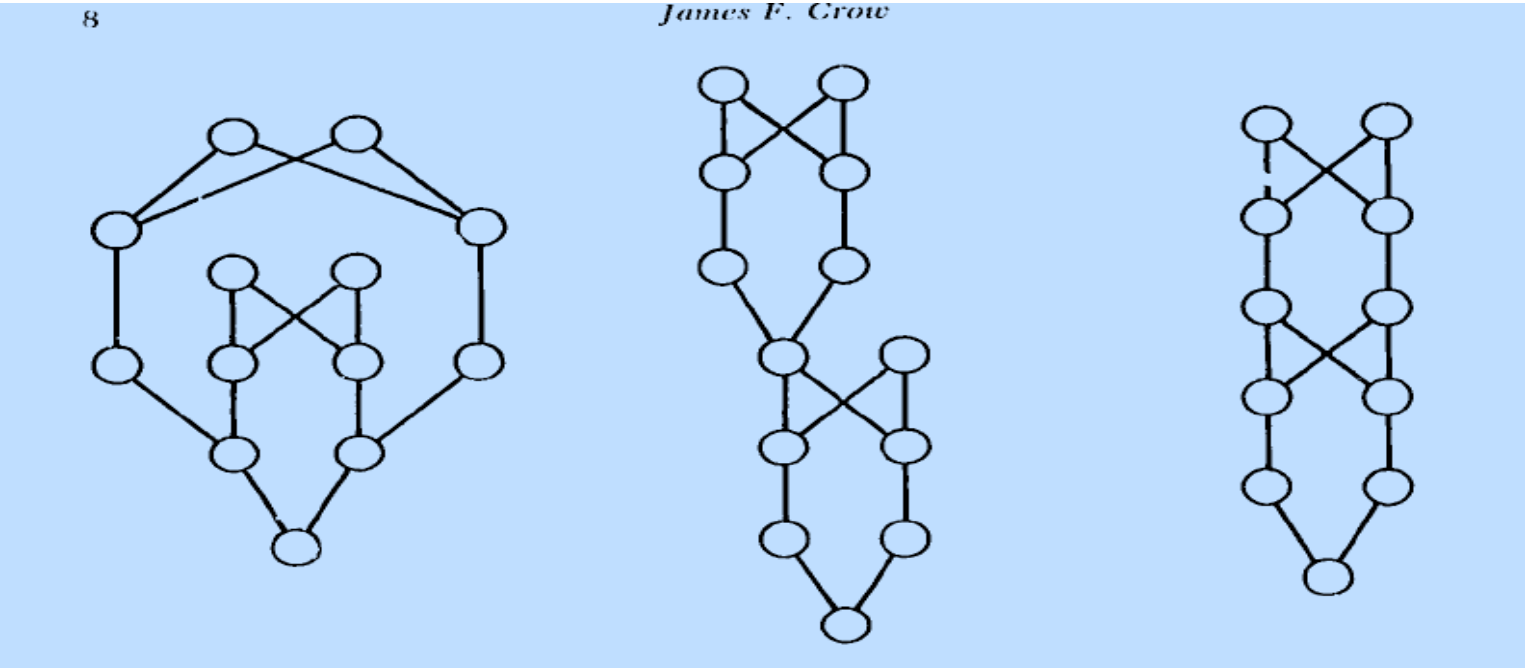


Isonymie dans le mariage

- Deux examples de mariage isonymique dans un petit cimitière de campagne

Isonymie et consanguinité (XX siécle)

- Ce genre d'étude fut repris seulement après 1965, lorsque Crow et Mange montrèrent que, pour la plupart des typologies de mariages entre parents, la probabilité que des individus apparentés aient le même nom est proportionnelle au degré de consanguinité. Par conséquent le pourcentage des mariages isonymiques constitue une mesure assez directe du niveau de consanguinité de la population.



Isonymie dans le mariage: les Xuetes (1)

	Fuster	Forteza	Aguilò	Cortés	Pinya	Pomar	Bonnìn	Martì	Mirò	Picò	Valls
Fuster	22	13	10	7	13	10	5	2	4	7	10
Forteza	11	19	15	17	5	7	2	10	7	1	3
Aguilò	15	18	12	15	7	6	8	11	4	2	5
Cortés	4	13	13	18	2	6	2	14	0	0	6
Pinya	17	7	11	0	6	13	10	2	3	10	7
Pomar	12	12	6	2	11	10	8	7	4	4	4
Bonnìn	14	5	7	4	9	11	4	0	8	8	5
Martì	5	16	6	20	0	5	2	10	0	0	0
Mirò	7	3	8	1	5	6	10	0	4	9	4
Picò	8	2	4	0	10	6	4	0	6	3	4
Valls	4	7	3	4	7	5	0	1	4	3	1

Isonymie dans le mariage: les Xuetes (2)

	Fuster	Forteza	Aguilò	Cortés	Pinya	Pomar	Bonnìn	Martì	Mirò	Picò	Valls
Fuster	15	15	13	12	10	11	7	7	6	6	7
Forteza	14	13	11	11	9	9	6	7	5	5	6
Aguilò	14	14	12	11	10	10	7	7	5	6	6
Cortés	11	11	9	9	8	8	5	6	4	4	5
Pinya	12	11	10	9	8	8	5	6	4	5	5
Pomar	11	10	9	8	7	7	5	5	4	4	5
Bonnìn	11	10	9	8	7	7	5	5	4	4	5
Martì	9	9	8	7	6	7	4	5	3	4	4
Mirò	8	8	7	6	5	6	4	4	3	3	4
Picò	7	6	5	5	4	5	3	3	2	3	3
Valls	5	5	4	4	4	4	3	3	2	2	2

Isonymie dans le mariage: les Xuetes (3)

	Fuster	Forteza	Aguilò	Cortés	Pinya	Pomar	Bonnìn	Martì	Mirò	Picò	Valls
Fuster	45%	-11%	-20%	-41%	28%	-6%	-30%	-73%	-29%	20%	50%
Forteza	-19%	46%	35%	60%	-45%	-26%	-68%	53%	40%	-81%	-49%
Aguilò	5%	31%	2%	34%	-27%	-40%	19%	59%	-24%	-64%	-20%
Cortés	-65%	18%	38%	101%	-74%	-25%	-63%	153%	-100%	-100%	20%
Pinya	45%	-38%	14%	-100%	-23%	59%	82%	-65%	-31%	122%	36%
Pomar	13%	17%	-32%	-76%	54%	34%	60%	36%	2%	-2%	-14%
Bonnìn	31%	-51%	-20%	-52%	26%	47%	-20%	-100%	103%	95%	7%
Martì	-47%	78%	-22%	172%	-100%	-24%	-54%	120%	-100%	-100%	-100%
Mirò	-13%	-61%	21%	-84%	-7%	6%	164%	-100%	34%	190%	13%
Picò	23%	-68%	-25%	-100%	130%	32%	31%	-100%	150%	20%	40%
Valls	-25%	37%	-32%	-4%	96%	34%	-100%	-61%	103%	46%	-57%

Parenté entre populations

- Plus généralement on établit le degré de parenté entre deux populations différentes à partir dès pourcentages de personnes ayant le même nom de famille. Il faut pourtant corriger les résultats pour tenir compte des deux faits très importants:
- le même nom peut avoir des origines différentes et n'être pas dû à un ancêtre commun
- les phénomènes migratoires peuvent influencer les distributions des noms

Populations iberiques et ibero-américaines

- Il vaut la peine de signaler que certains études empiriques, parmi les plus importants dans ce contexte, ont été réalisés sur des populations ibériques et ibéro-américaines, grâce au fait que le système espagnol des noms de famille fournit un immédiat et facile contrôle de l'isonymie des ancêtres.

- Premièrs études:
- J. Pinto-Cisternas *et al., Estimation of Inbreeding from Isonymy in Iberoamerican Populations,* Am.J. Hum. Gen. 37 (1985) 373-385
- G.W.Lasker, *Datos sobre los apellidos hispanoamericanos en los estudios de biologia humana,* Ann. Antrop. (Mex) 28 (1993) 107-128

Populations ibero-américaines

Etudès régionaux (États iberoaméricaines):

- ARGENTINA: Alfaro et Dipierri (1994-1997) [Province de Jujuy]
  Colantonio *et al* (1998-2002) [Province de Còrdoba]
- BOLIVIA: Chakraborty *et al* (1989) [Aymara]
- COSTA RICA: Madrigal *et al* (1997-1999) [Escazu]
- MEXICO: Allen (1987-1988) [Mennonites]
  Christensen (2002) [Maya XVI siècle, Mixtec]
- PÉROU: Pettener (1997) [Santa Ana]
- VENEZUELA: Castro de Guerra *et al* (1997-1999)
  Pinto-Cisternas *et al* (1990-1997)
  Rodriguez-Larralde (1989-2000) [Lara, Sucre, Yachira, Anzoategula, Trujillo, Guarico, Zulia, Aragua, …]

Populations iberiques et françaises

Études régionaux (Éspagne, Portugal):

- Eizaguirre (1997): Vallée du Minho
- Fuster *et al* (1996): Vallées de Gredos
- Rodriguez-Larralde *et al* (2003): Noms de famille en Éspagne
- Branco *et al* (2005): Azores
- Esparza *et al* (2006): Règion du delta de l'Ebro
- Lermo *et al* (2006): Population gitane d'Éspagne

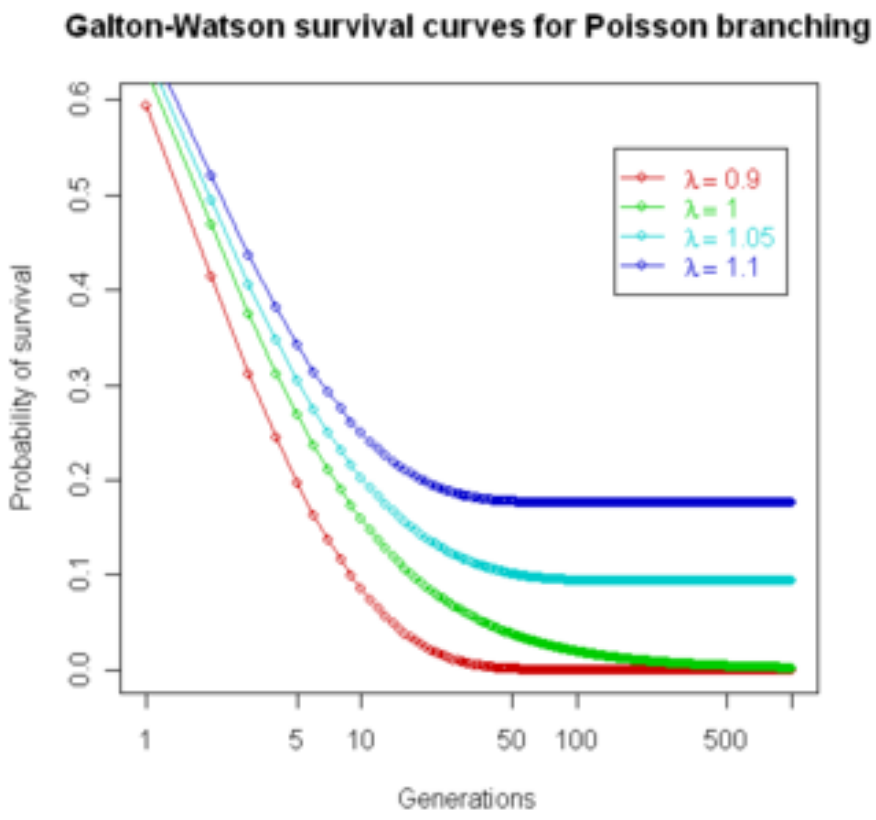
Études régionaux (France):

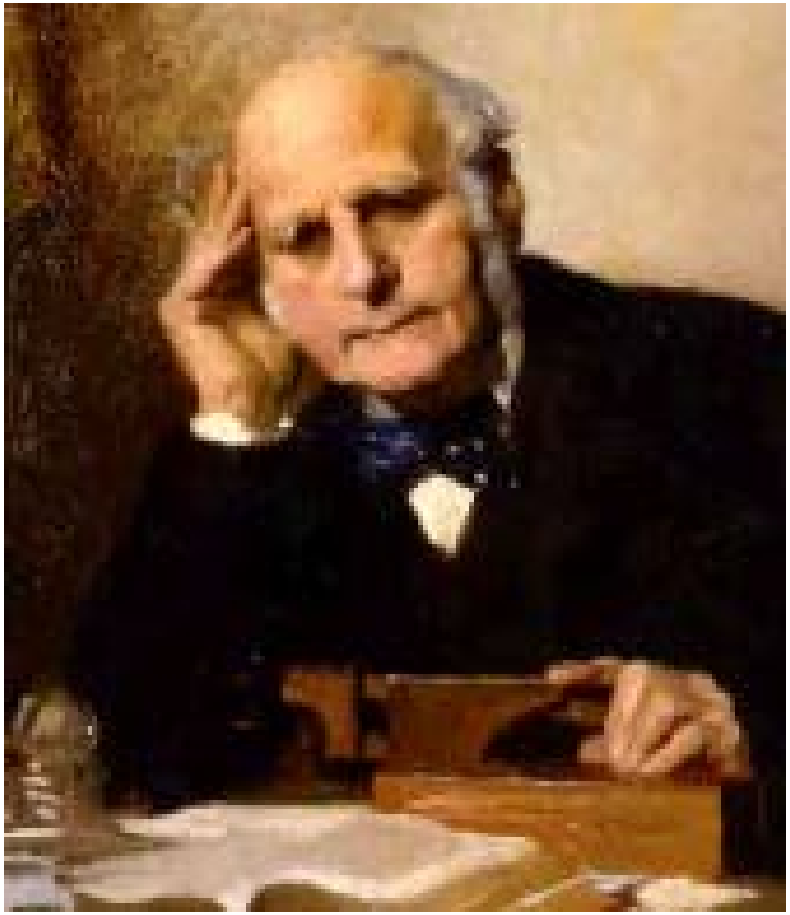
- Darlu, Ruffie (1992-1997): Patronymiques de France
- Degioanni *et al* (1996-2001): Immigration Italiènne en France
- Vernay (2000-2001): Distribution des noms de famille en France,
  Noms de famille dans le Jura (1763-1972)
- Scapoli *et al* (2005): Noms de famille et dialectes en France
- Prost (2005): Isonymie matrimoniale dans le sud-est (1546-1899)

Distribution des noms de famille (XIX siécle)

- La deuxième ligne d'enquête fait son début en 1874 avec le travail de F.Galton (cousin de Charles Darwin) et H. W. Watson sur la probabilité d'extinction des noms de famille. Ce travail fut motivé en ce temps-là par l'observation soucieuse que nombreuses familles de l'aristocratie britannique étaient éteintes ou risquaient de s'éteindre. Cet étude entama une entière ligne de recherches mathématiques sur les procès de ramification (*branching processes*). Le premier résultat prévoyait erronément que chaque nom avait probabilité 1 de s'éteindre après un temps suffisamment long; en suite on éclaircit que telle probabilité n'est pas 1 dans une population croissante, cependant elle a une valeur pas négligeable et prochaine à 0.5.

Le modèle de Galton et Watson



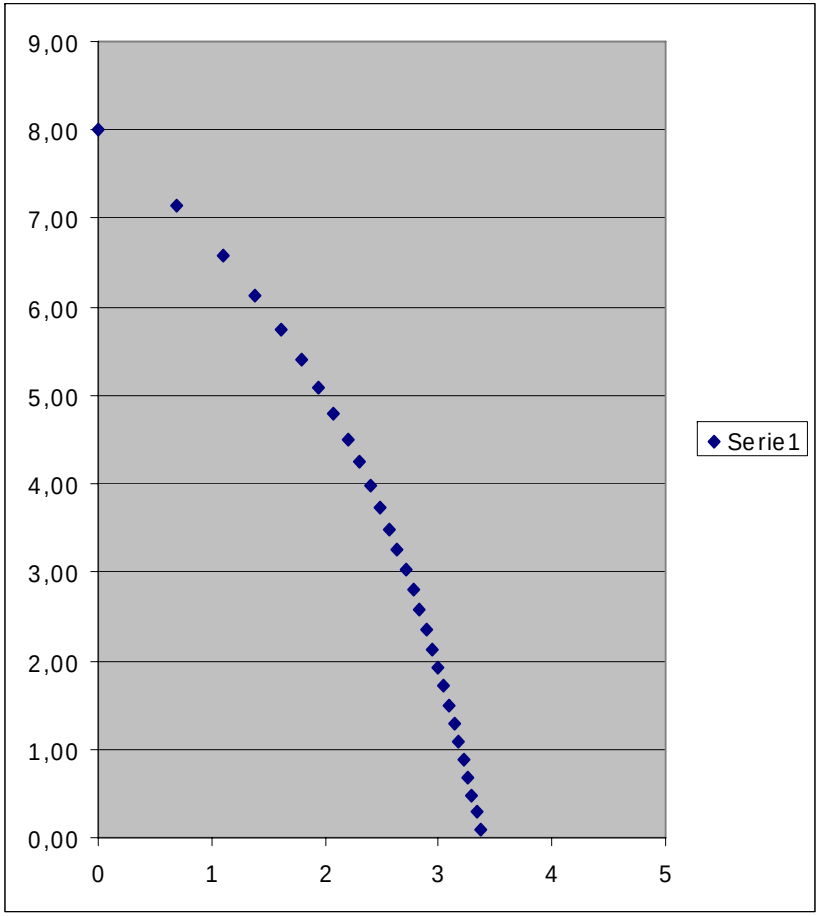


Les noms de famille en génétique

- La génétique ne s'intéressa pas trop aux noms de famille jusqu'aux premières années Soixante-dix, lorsque Yasuda, Cavalli-Sforza et leurs collaborateurs reconnurent que la transmission des noms obéit aux mêmes lois biologiques et statistiques de la transmission des allèles neutres (la large partie du patrimoine génétique qui ne conditionne pas le phénotype, et donc n'est pas assujettie à la pression évolutive). En particulier les noms de famille européens se transmettent exactement comme le chromosome Y, en passant de père en fils dans la ligne masculine, et, comme pour les gènes, leur distribution se développe dans le temps pour effet des deux phénomènes concomitants: la mutation et la migration.

Le modèle de Karlin et McGregor

- Par conséquent on appliqua aux noms de famille les mêmes conclusions et les analyses déjà effectués pour les allèles neutres, et en particulier la théorie mathématique formulée en 1967 par Karlin et McGregor, qui tenait explicitement compte de la possibilité que, avec le temps, le patrimoine génétique était sujet à des mutations.
- La version simplifiée de la théorie, applicable dans le cas de populations très nombreuses (distribution de Fisher), devint très rapidement le moyen ordinaire de paramétrisation des distributions de noms.



Le modèle de Fox et Lasker

- Un grand intérêt fut aussi soulevé par l'observation de Fox et Lasker (1983), qui relevèrent, sur une base purement empirique, que la fréquence des noms était décrite avec remarquable précision par la courbe de Pareto, originairement introduite pour décrire la distribution des revenus, mais plus tard réutilisée dans l'analyse quantitative de nombreux différents phénomènes sociaux.

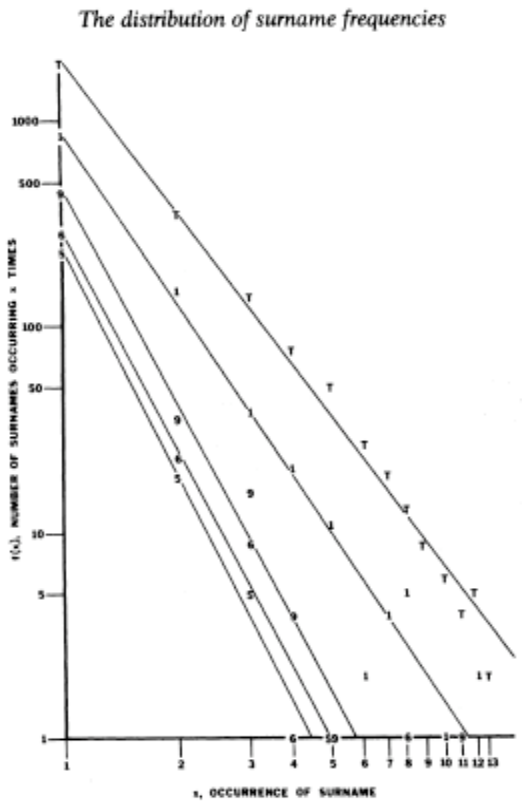
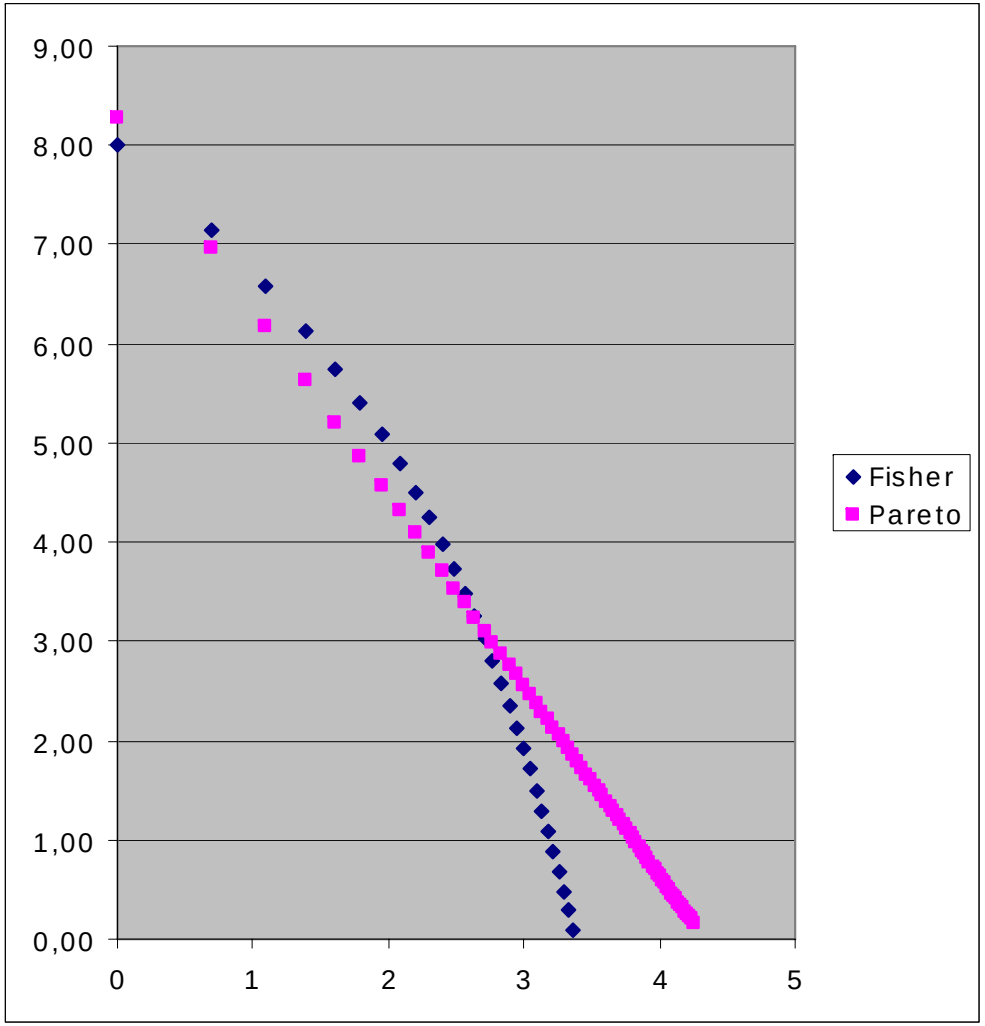


Figure 1. *Number of surnames occurring x times, f(x) plotted against x, on logarithmic scales, with fitted lines of the discrete Pareto distributions (districts, 1, 5, 6, 9 and all districts combined, T).*

Courbe de Fisher et courbe de Pareto



L’importance des phénomènes migratoires

- Un des aspects plus intéressants des analyses sur la distribution des noms de famille est la possibilité de calculer, sur la base des données empiriques relatifs à la distribution actuelle, une estime statistique pour la fraction de noms qui se sont modifiés ou ajoutés par rapport à une hypothétique population initiale. Ainsi on peut estimer l'importance des phénomènes migratoires si le nombre des mutations peut être considère relativement négligeable.

La distribution des noms de famille en Europe

- L’étude des phénomènes migratoires a attrait beaucoup d’attention par les génétistes, qui ont étudié les dynamiques migratoires en Sardaigne (1984), ensuite en Italie (1987) et en France (1992), pendant que le groupe dirigé par J. Barrai, en exploitant, à partir de 1995, les possibilités offertes par les versions digitalisées des listes téléphoniques, a étendu l’analyse à tous les principaux Pays européens, jusqu'à couvrir moyennement le 9% de l’entière population, et a fourni des estimes intéressantes sur l’entité des migrations à niveau continentale et sur les effets de séparation et d'isolation dus à la distance spatiale et aux barrières géographiques et linguistiques.

Noms de famille en Éspagne

TABLE 3. Fifty most frequent Spanish paternal and maternal surnames in a sample of 3,625,688 telephone users

Paternal	Frequency	Maternal	Frequency
Nunez	6,537	Herrero	6,357
Cano	6,626	Nunez	6,376
Garrido	6,697	Pascual	6,471
Santos	6,744	Iglesias	6,532
Prieto	6,798	Morales	6,728
Herrero	6,822	Lozano	6,954
Ramirez	7,371	Prieto	6,971
Castillo	7,534	Castillo	6,980
Marin	7,545	Garrido	7,064
Calvo	7,707	Calvo	7,718
Diez	7,850	Marin	7,765
Morales	7,886	Diez	7,918
Iglesias	7,909	Ramirez	8,263
Delgado	8,212	Delgado	8,310
Molina	8,246	Sdarez	8,676
Suarez	8,367	Ramos	8,802
Ortega	9,122	Molina	8,992
Ortiz	9,170	Ortiz	9,321
Ramos	9,183	Ortega	9,571
Rubio	9,518	Torres	9,703
Castro	10,019	Castro	9,736
Torres	10,208	Rubio	9,979
Serrano	10,997	Dominguez	11,134
Vazouez	11,110	Sanz	11,262
Dominguez	11,623	Vazquez	11,831
Sanz	11,854	Blanco	12,085
Blanco	12,022	Gil	12,137
Gil	12,041	Serrano	12,184
Navarro	14,434	Navarro	14,489
Romero	14,983	Romero	15,326
Gutierrez	15,704	Gutierrez	15,659
Munoz	20,248	Alonso	19,942
Alonso	20,967	Munoz	20,228
Moreno	23,289	Hernandez	23,227
Hernandez	23,628	Moreno	23,530
Jimenez	25,432	Diaz	25,922
Diaz	27,697	Jimenez	25,955
Alvarez	27,924	Alvarez	27,174
Ruiz	30,885	Ruiz	31,103
Gomez	37,727	Gomez	38,556
Martin	37,755	Martin	39,672
Perez	60,726	Perez	64,475
Sanchez	63,596	Sanchez	64,508
Rodriguez	64,155	Rodriguez	65,478
Gonzalez	66,546	Gonzalez	68,915
Martinez	71,052	Martinez	72,752
Lopez	72,432	Lopez	75,736
Fernandez	79,248	Fernandez	79,960
Garcia	118,940	Garcia	121,614

TABLE 2. Hundred most frequent Spanish double surnames in a sample of 3,625,688 telephone users[1]

Gomez Lopez	929	Garcia Gomez	1,431
Hernandez Hernandez	929	Perez Lopez	1,435
Gonzalez Martin	931	Perez Martinez	1,474
Rodriguez Martin	938	Perez Fernandez	1,503
Aevarez Rodriguez	944	Fernandez Alvarez	1,517
Rodrigoez	949	Sanchez Lopez	1,526
Garcia Diaz	953	Rodriguez Lopez	1,532
Garcia Moreno	966	Fernandez Martinez	1,546
Martinez	969	Alvarez Fernandez	1,554
Gonzalez Alvarez	988	Rodriguez Gonzalez	1,571
Alvarez Gonzalez	993	Lopez Rodriguez	1,579
Diaz Fernandez	993	Martinez Fernandez	1,601
Hernandez Garcia	1,018	Gonzalez Lopez	1,609
Ruiz Garcia	1,024	Fernandez Perez	1,619
Garcia Hernandez	1,028	Martin Garcia	1,636
Garcia Jimenez	1,031	Garcia Martin	1,684
Gomez Gonzalez	1,046	Alvarez Alvarez	1,706
Lopez	1,047	Garcia	1,715
Perez Gonzalez	1,050	Gonzalez Rodriguez	1,801
Perez Sanchez	1,070	Lopez Martinez	1,873
Martinez Rodriguez	1,082	Martinez Lopez	1,884
Sanchez Martin	1,092	Lopez Fernandez	1,965
Fernandez	1,099	Martin Martin	2,146
Martin Sanchez	1,115	Gonzalez Fernandez	2,167
Garcia Ruiz	1,125	Rodriguez Fernandez	2,204
Rodriguez Sanchez	1,130	Fernandez Lopez	2,217
Jimenez Jimenez	1,139	Fernandez Rodriguez	2,303
Sanchez Perez	1,165	Perez Garcia	2,315
Rodriguez Martinez	1,182	Garcia Perez	2,317
Lopez Perez	1,195	Garcia Rodriguez	2,423
Diaz Garcia	1,200	Sanchez Garcia	2,445
Gomez Gomez	1,205	Rodriguez Garcia	2,519
Martinez Gonzalez	1,218	Fernandez Gonzalez	2,588
Sanchez Rodriguez	1,221	Martinez Garcia	2,596
Gonzalez Perez	1,249	Perez Perez	2,641
Garcia Alvarez	1,266	Garcia Sanchez	2,650
Gonzalez Sanchez	1,293	Gonzalez Garcia	2,679
Fernandez Sanchez	1,323	Lopez Garcia	2,778
Sanchez Martinez	1,331	Garcia Lopez	2,855
Alvarez Garcia	1,365	Garcia Gonzalez	2,940
Gonzalez Martinez	1,366	Garcia Martinez	2,952
Martinez Sanchez	1,373	Rodriguez Rodriguez	3,135
Lopez Gonzalez	1,381	Gonzalez Gonzalez	3,138
Martinez Perez	1,386	Garcia Fernandez	3,142
Perez Rodriguez	1,388	Sanchez Sanchez	3,348
Rodriguez Perez	1,388	Fernandez Garcia	3,679
Sanchez Gonzalez	1,396	Lopez Lopez	3,889
Gomez Garcia	1,409	Martinez Martinez	3,890
Sanchez Fernandez	1,413	Fernandez Fernandez	5,891
Lopez Sanchez	1,424	Garcia Garcia	7,722

[1] Five of these are single surnames.

Noms de famille en France

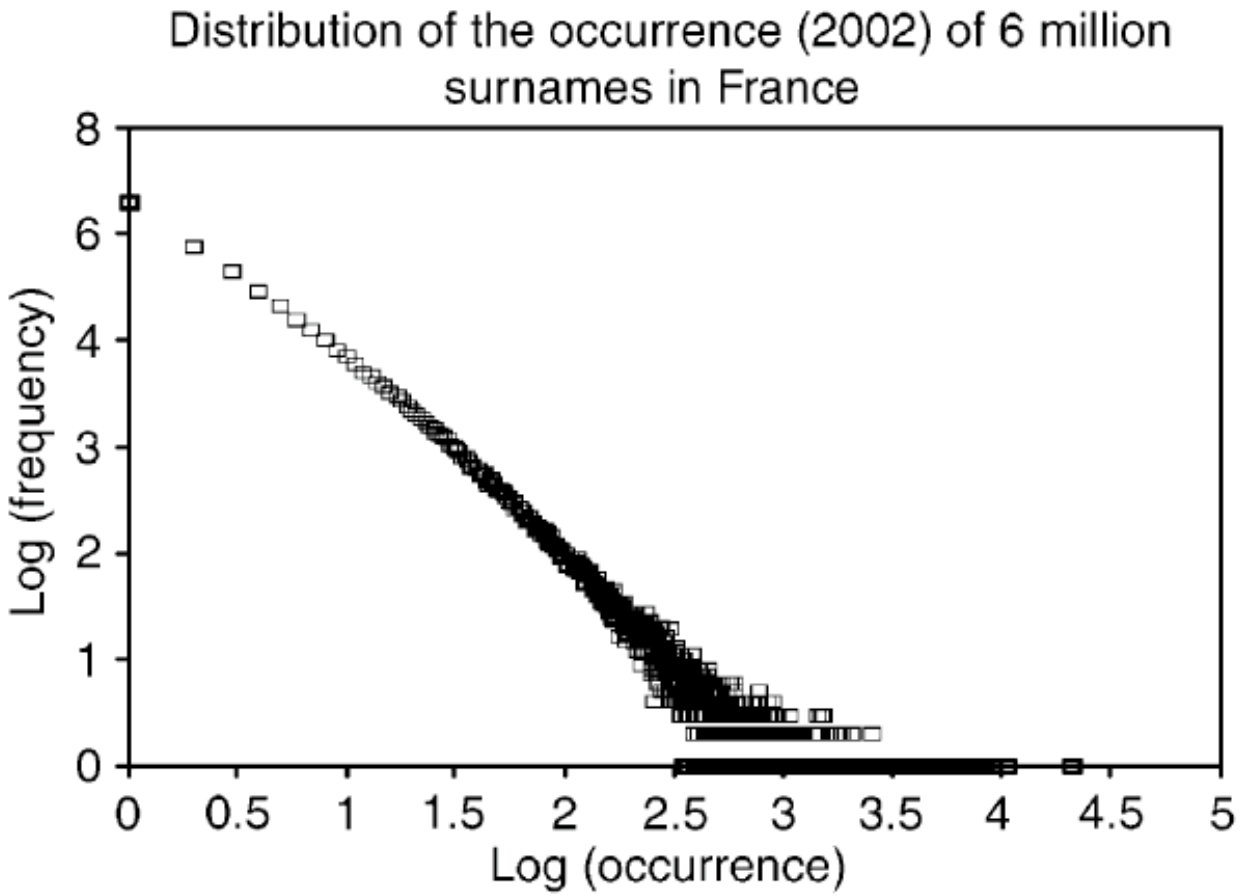
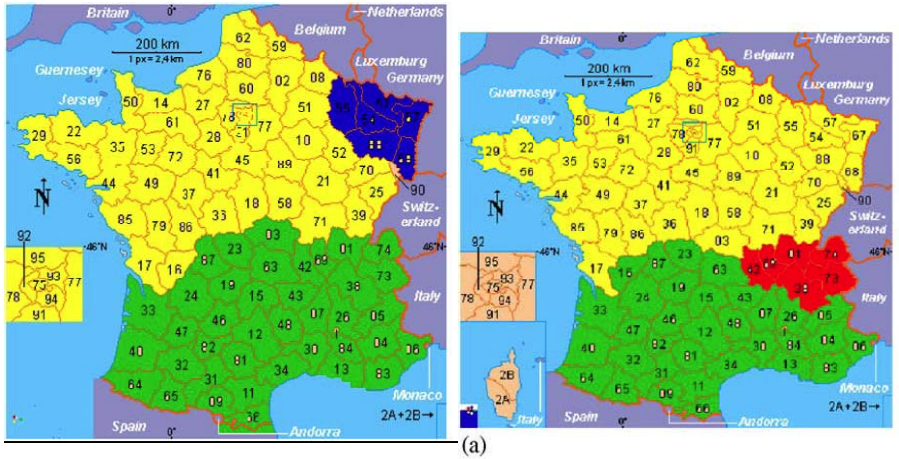


Fig. 2. The log–log distribution of the frequency of occurrence of surnames in France.

Noms de famille et dialectes en France

Table 2
The 100 most frequent surnames in 6 million users in France. Increasing order of frequency

BENOIT	2910	FABRE	4285
MARTY	2912	MULLER	4287
ROLLAND	2925	NICOLAS	4314
LEMOINE	2934	MATHIEU	4375
PIERRE	2957	CLEMENT	4385
COHEN	2970	VIDAL	4393
RENARD	3003	GAUTIER	4397
ROGER	3032	ROBIN	4450
LEMAIRE	3052	BOYER	4559
RENAUD	3083	PEREZ	4611
PICARD	3097	SANCHEZ	4646
LACROIX	3128	LEFEVRE	4718
NGUYEN	3148	CHEVALIER	4778
COLIN	3186	GAUTHIER	4793
AUBERT	3210	DUPONT	4797
REY	3262	MERCIER	4810
BARBIER	3324	PERRIN	4844
LUCAS	3335	LEROUX	4897
GERARD	3344	GUERIN	4963
NOEL	3358	GARNIER	5008
BEN	3363	LOPEZ	5073
DUMONT	3376	LAMBERT	5166
DUFOUR	3387	MOREL	5169
MARIE	3417	ANDRE	5256
JOLY	3436	VINCENT	5309
DUMAS	3443	LEFEBVRE	5356
ROY	3509	ROUSSEAU	5571
JEAN	3514	FOURNIER	5775
RODRIGUEZ	3566	FAURE	5802
GAILLARD	3626	GIRARD	5882
MARCHAND	3629	BLANC	5990
DUVAL	3635	BONNET	6123
MEUNIER	3657	DAVID	6221
MEYER	3662	BERTRAND	6256
ROCHE	3685	MARTINEZ	6283
DENIS	3710	ROUX	6784
BRUNET	3792	LEROY	6809
ARNAUD	3796	LAURENT	7249
LEGRAND	3821	SIMON	7594
BLANCHARD	3875	MICHEL	7707
FONTAINE	3889	DUBOIS	7765
GIRAUD	3894	PETIT	8113
DASILVA	3944	GARCIA	8169
ROUSSEL	3965	ROBERT	8278
FERNANDEZ	3980	MOREAU	8286
FRANCOIS	4043	RICHARD	8458
BRUN	4056	THOMAS	9171
HENRY	4066	DURAND	9193
MASSON	4124	BERNARD	10663
MORIN	4218	MARTIN	21043

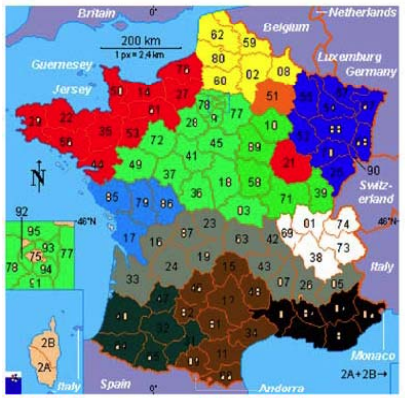


(a)

Map of France derived from the matrix of dialect distances between departments.

The departments of most of the Franco-Provençal area cluster together.

The North-South boundary mostly coincide with that identified by surname distances.



(b)

The subclusters formed by the matrix of dialect distances between departments.

Some of the subclusters correspond to well defined linguistic areas.

The Limousin area is not separated from the Auvergnat and from part of the Gascon at west and Provençal at East.

Noms de famille en Europe

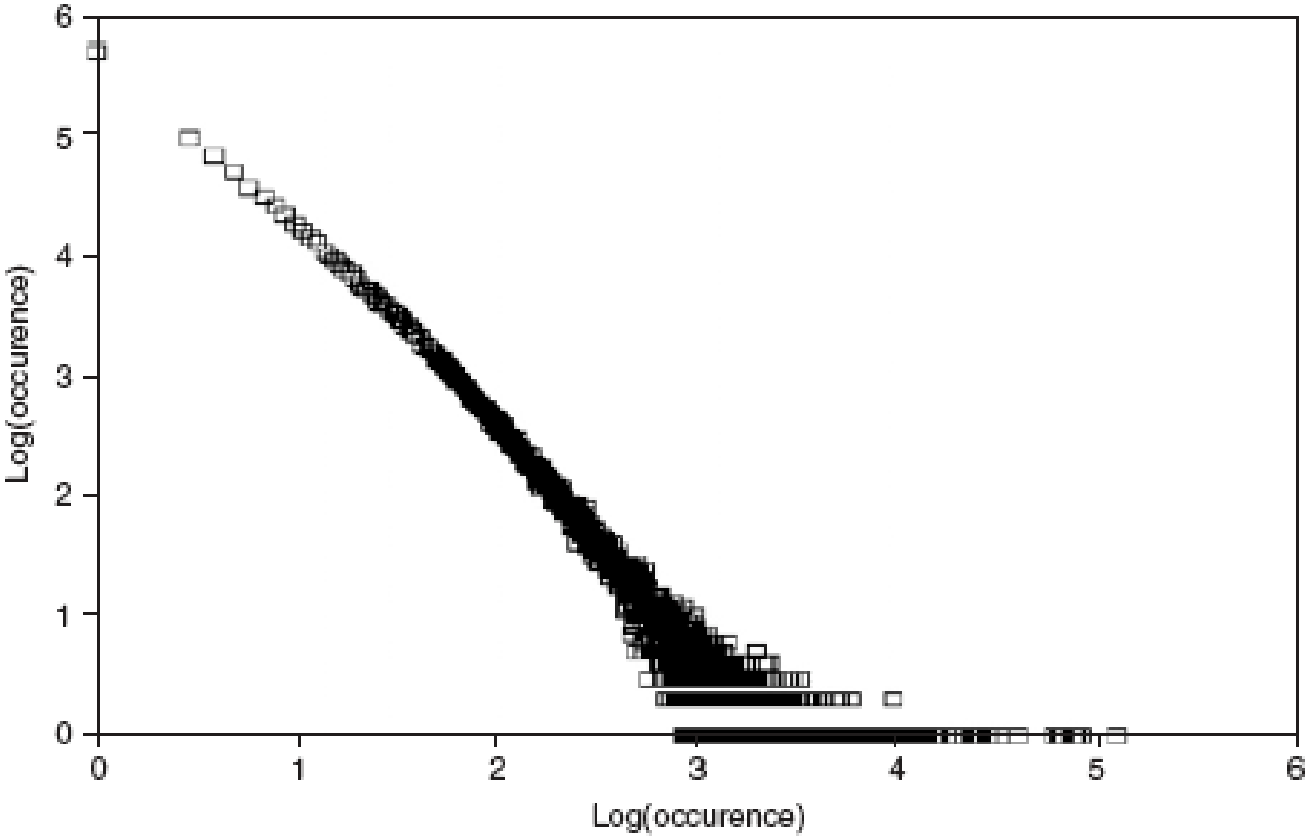
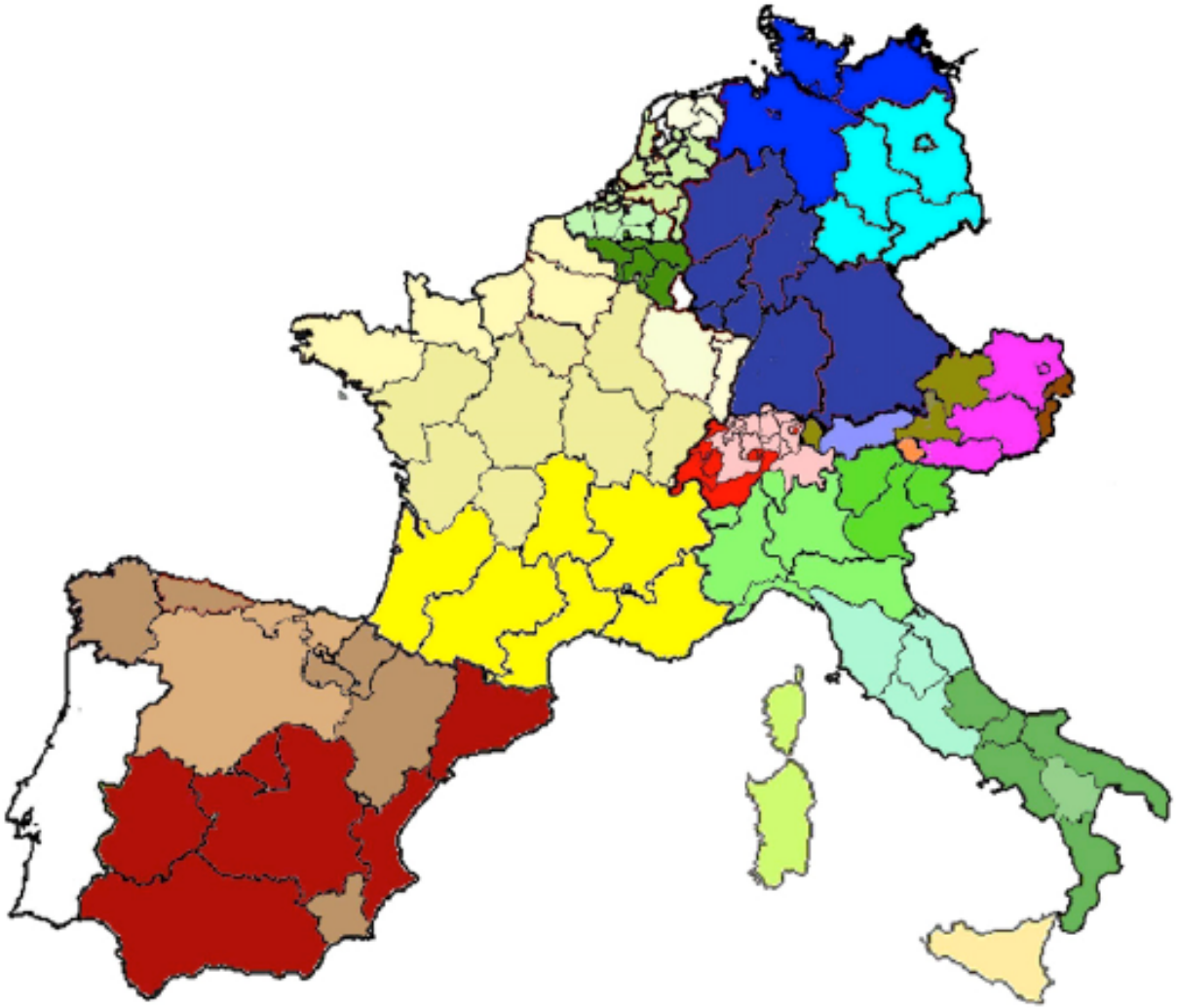


Fig. 2. The log-log distribution of the frequency of occurrence of surnames in Western Europe.

Noms de famille en Europe: les régions



Nom de famille et chromosome Y

- Une autre ligne de recherche a été engendrée par l'observation de l'analogie entre le nom de famille et le chromosome Y, et concerne le problème de l'origine (monophylétique ou polyphylétique) des noms de famille. L'étude que Sykes et Irven (2000) ont conduit sur le DNA d'un groupe de porteurs mâles du nom Sykes, pour lequel les sources documentaires fouillées par les historiens semblaient indiquer une origine multiple, a montré que les données génétiques sont par contre cohérentes avec l'hypothèse que, compte tenu des inévitables cas de non-paternité, tous les Sykes descendent d'un seul ancêtre mâle commun.

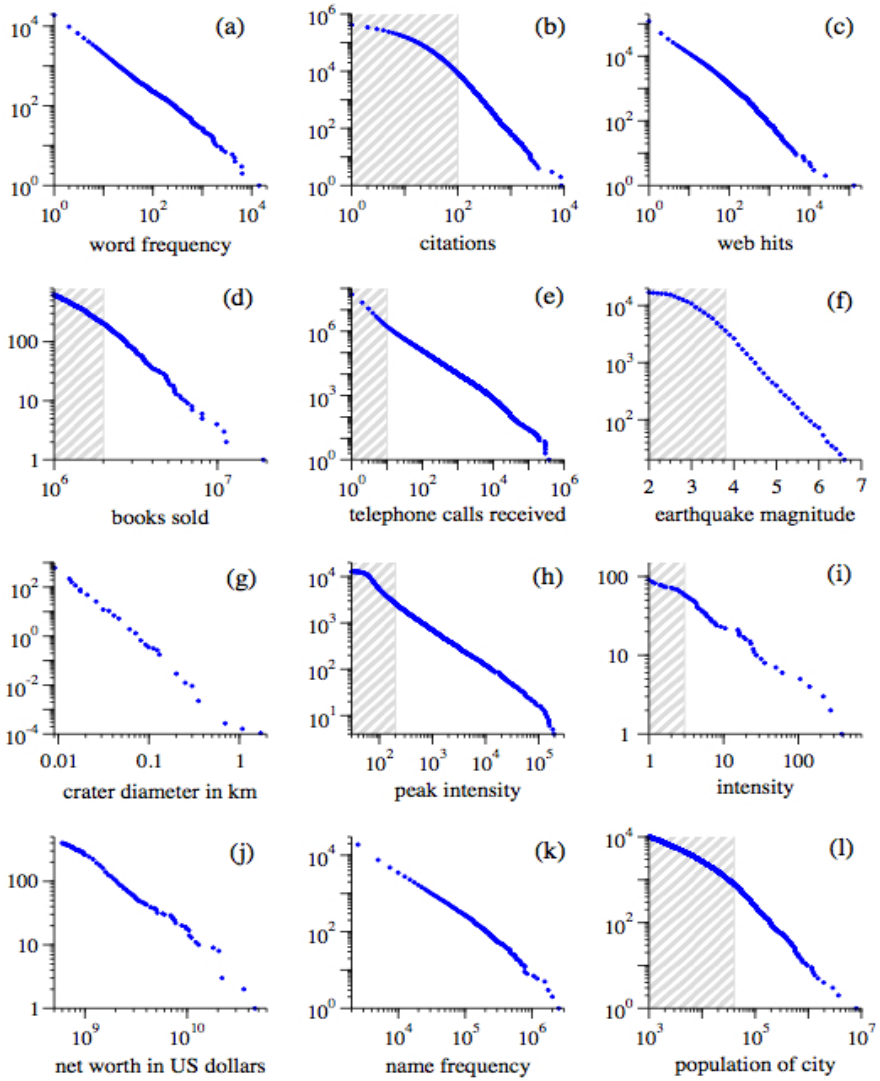
Byran Sykes

La génétique comparé à l'histoire

- Autres intéressantes études sont ceux relatives au nom hébreu Cohen, pour lequel l'hypothèse biblique d'une descendance patrilinéaire directe à partir d'Aaron frère de Moïse résulte génétiquement plausible, ceux sur l'ascendance séparée des irlandais ayant un nom de famille d'origine gaélique et de ceux ayant un nom anglo-saxon, et tout récemment (2007) les recherches sur l'influence génétique des Vikings dans le nord-ouest de l'Angleterre.

Noms de famille et physique statistique

- Prenons enfin en considération les aspects du phénomène qui ont attrait particulièrement l'attention des mathématiciens et des physiciens statisticiens, qui sont beaucoup attirés par l'existence de structures universelles en phénomènes très différents et cherchent à expliquer leur origine dynamique par la construction de modèles mathématiques (dans ce cas modèles d'évolution des populations).



La statistique des arbres généalogiques

- L'analyse de la consanguinité a reçu une importante contribution de la publication du travail de Derrida, Manrubia et Zanette (1999), dans lequel on a comparé un modèle théorique avec les valeurs empiriques de la répétition des ancêtres dans l'arbre généalogique d'Edouard III, roi d'Angleterre.

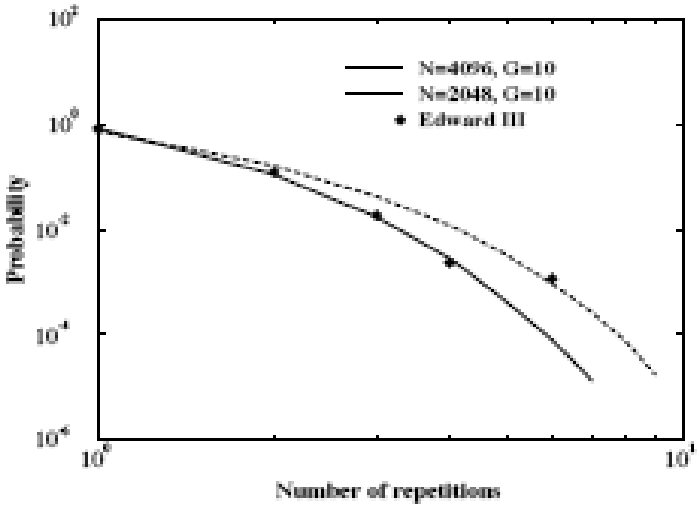
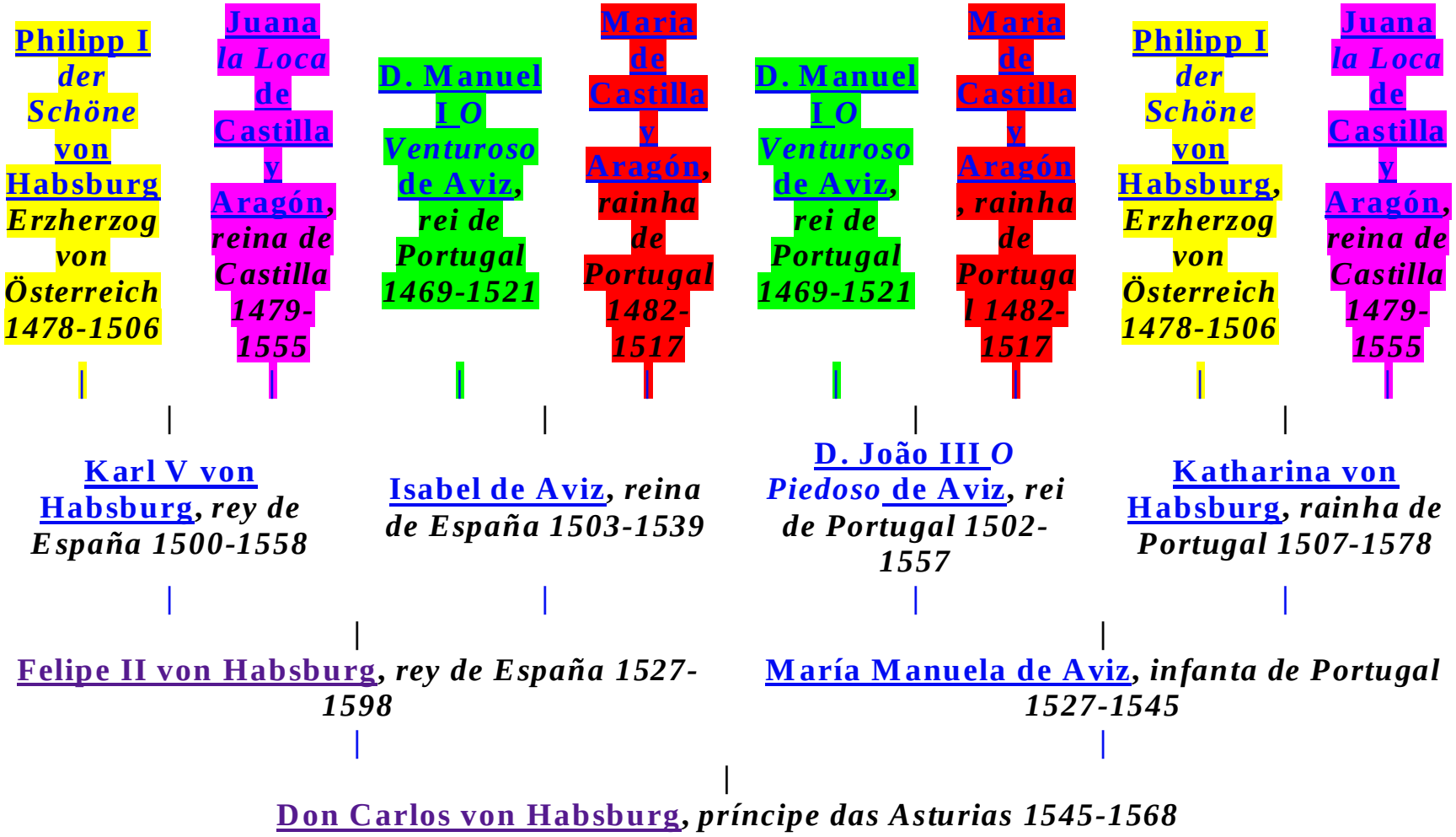


FIG. 1. Probability of ancestor repetitions in the genealogical tree of the king Edward III [5]. The continuous and dashed lines represent the results of simulations of $F(r)$ in a closed population with 2^{11} and 2^{12} individuals for our model. Averages have been performed over the ten first generations of 10^3 independent trees.

Les ancêtres de Don Carlos (1)



Les ancêtres de Don Carlos (2)

- La sixième génération: 12 ancêtres (en principe 32):
- Juan II rey de Aragón x 4
- Juana Enriquez (de Castilla) x 4
- Juan II rey de Castilla x 4
- Isabel de Aviz x 4
- Duarte I de Aviz x 2
- Leonor de Aragón x 2
- Joao de Aviz x 2
- Isabel de Bragança x 2
- Friedrich III d'Habsburg x 2
- Leonor de Aviz x 2
- Charles de Bourgogne x 2
- Isabelle de Bourbon x 2

- La dixième génération: 83 ancêtres (en principe 512):

Un programme de recherche

- Malheureusement, ne sont pas encore suivis autres études plus systématiques, mais aujourd'hui, avec l'énorme diffusion de matériels généalogiques dans Internet et avec la création de bases de données contenant des centaines de milliers de noms et leurs connexions familiales, les temps sont mûrs pour un étude beaucoup plus étendu, qui pourrait parvenir à une validation définitive du modèle.

- Quelques examples, extraits de la base de données «roglo»:

  - 33.677 ancêtres du roi d'Éspagne

  - 30.028 ancêtres de la reine d'Angleterre

  mais aussi:

  - 6.497 ancêtres de M. Jean Dupont (†1981) (nom pri au hasard)

Le modèle general de Baek, Tiet et Kim (1)

- Pour ce qui concerne la distribution des noms, le résultat le plus avancé est dû au groupe coréen de Baek, de Tiet et de Kim, qui ont formulé en 2007 un modèle très général de dynamique des populations, qui englobe les modèles antécédents, en introduisant les probabilités de naissance, de mort, de mutation et d'immigration et en tirant des lois statistiques pour la croissance du nombre des noms de famille par rapport à la croissance globale de la population, et pour la probabilité qu'un nom paraisse dans la population en fonction du nombre des individus porteurs du nom.

Le modèle general de Baek, Tiet et Kim (2)

- Dans tous les cas on a montré que la distribution tend vers une courbe de Pareto, mais le paramètre caractéristique de la courbe dépend dès modalités avec lesquelles les nouveaux noms sont ajoutés, et prend des valeurs significativement différents dans le cas d'absence de mutation et dans celui d'absence de migration, avec des valeurs intermédiaires dans les situations mixtes.
- L'analyse des données empiriques, en particulier pour Pays, comme la Corée et la Chine, où les mutations des noms ont été toujours négligeables, montre un remarquable accord entre le modèle et la réalité historique.

La physique des systemes complexes

- Ces résultats ont été tout récemment reproduits dans la thèse d'A. De Luca, qui a utilisé le groupe de renormalisation, un outil tout propre de la physique théorique, ce qui a permis de reconduire la dynamique des noms de famille au contexte plus général des phénomènes de criticité auto-organisée. Ces phénomènes semblent caractériser beaucoup de systèmes complexes, soit physiques que biologiques, économiques et sociaux, dans lesquels ne soit pas présente une échelle caractéristique et les conditions d'auto similarité soient vérifiées.

Distribution des noms de famille en Italie

954 / BARRAI ET AL.

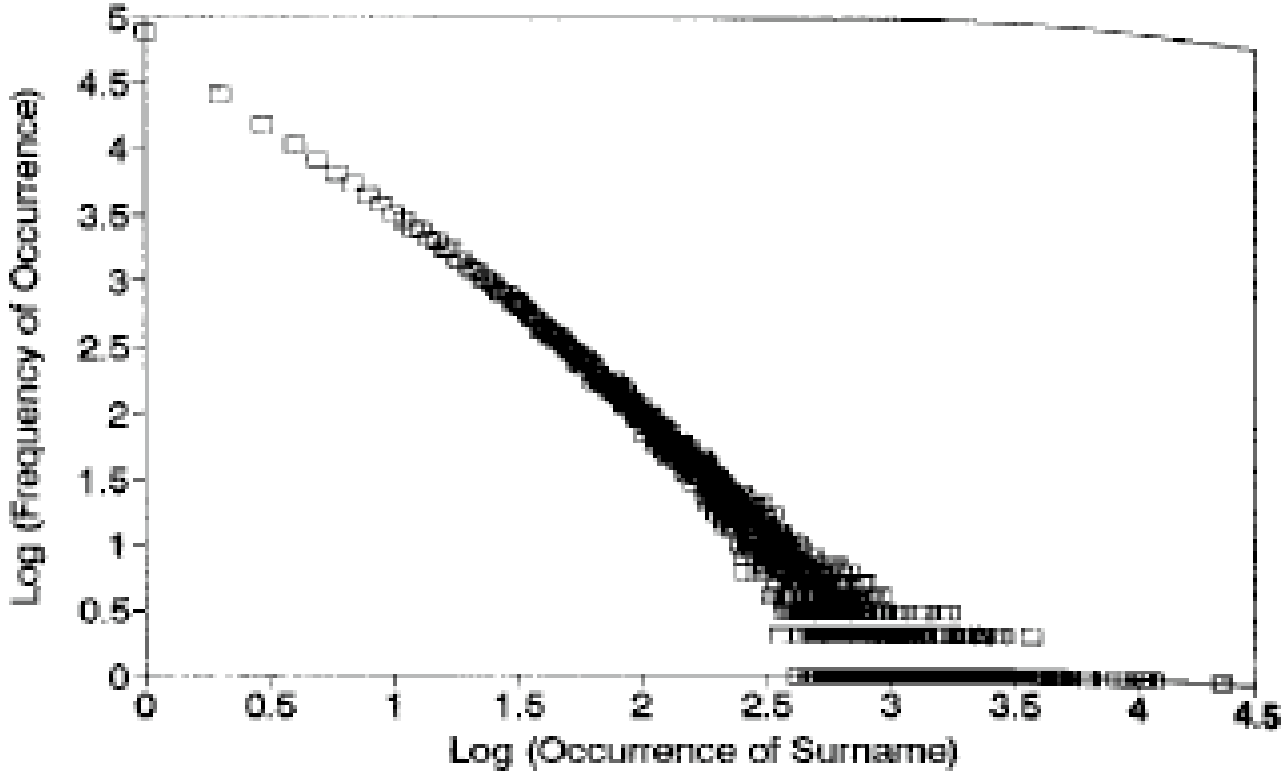
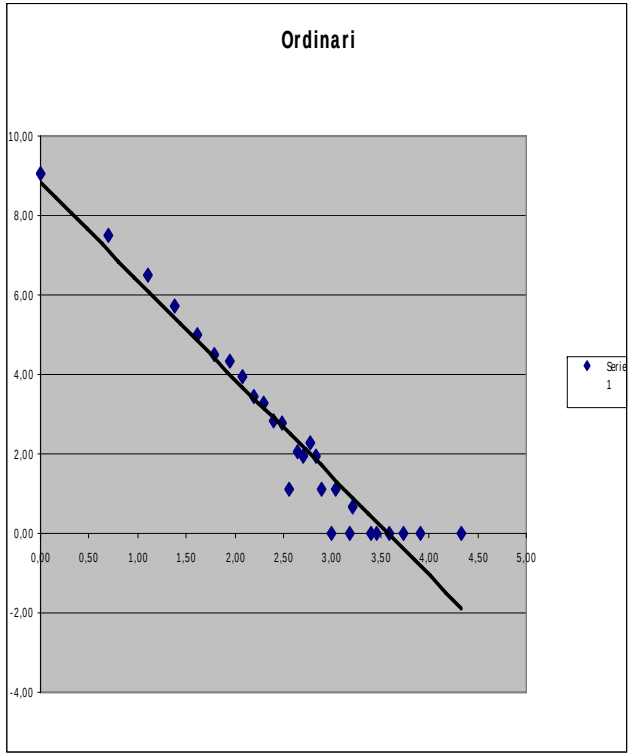
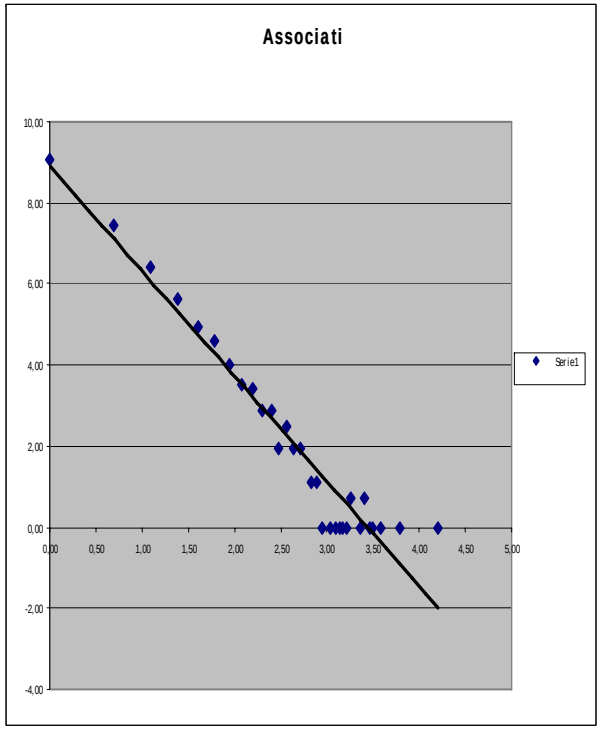
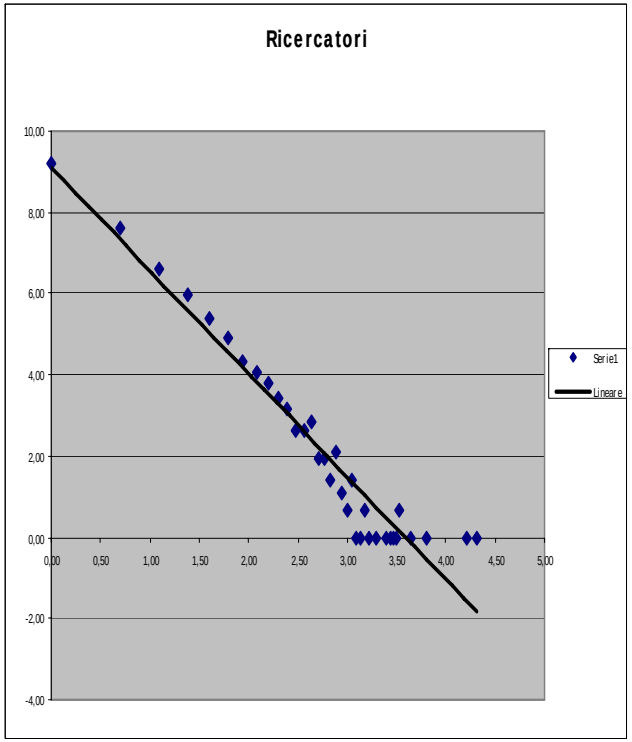


Figure 2. Log-log distribution (natural logarithms) of the occurrence of surnames in Italy in a sample of 5 million individuals in 1996.

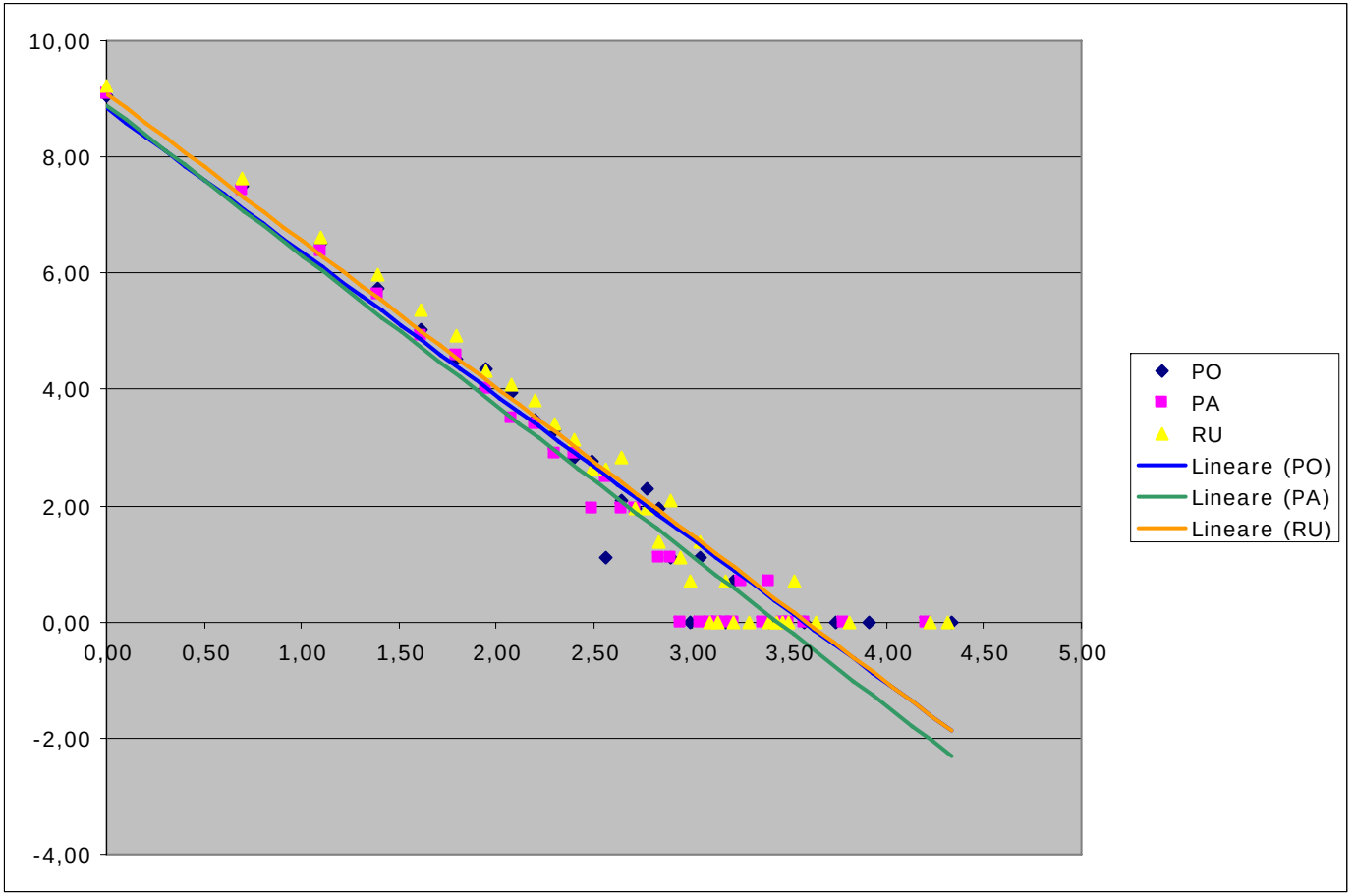
Les noms de famille dans les Universités italiennes







Noms de famille dans l'Université



Conclusions: le modèle et la réalité

	ITALIE (1997 annuaire téléphonique)	UNIVERSITÉ (modèle)	UNIVERSITÉ (2007 annuaire du Ministère)
INDIVIDUS	5.031.000	62.697	62.697
NOMS DE FAMILLE	215.623	26.408	26.562
FRÉQUENCE = 1	72.450 (33,6%)	16.953 (64,2%)	16.304 (61,4%)
FRÉQUENCE = 2	24.000 (11,1%)	4.122 (15,6%)	4.613 (17,3%)
FRÉQUENCE = 3	14.450 (6,7%)	1.743 (6,6%)	1.992 (7,5%)
ISONYMIE	0,00017	0,00017	0,00016

Questa presentazione è coperta, per le sue parti originali, da licenza Creative Commons:

attribuzione, non commerciale, condividi allo stesso modo

http://creativecommons.org/licenses/by-nc-sa/2.5/it/legalcode

rossi@df.unipi.it