

RÉSULTATS D'UNE ENQUÊTE SUR LA DISTRIBUTION DU SEXE DANS LES FAMILLES NOMBREUSES

par M. P. SCHÜTZENBERGER,

Attaché de Recherche à l'Institut National d'Hygiène

L'étude des processus qui déterminent le sexe de l'enfant à naître nécessite concurremment à toute recherche biologique une connaissance statistique précise qui lui indique des perspectives, contrôle ses déductions, fixe les limites *a priori* de sa validité.

Sous la direction de M. le Professeur R. Turpin, et grâce à l'appui du Fonds d'Etudes de la Société Médicale des Hôpitaux de Paris, nous avons effectué une semblable enquête. Elle nous a permis de mettre en évidence un fait nouveau que nous discuterons ici.

LES DOCUMENTS

Une attention particulière doit être apportée dans les enquêtes démographiques à ce que les facteurs économiques et sociaux aux sens le plus large, ne viennent masquer les phénomènes proprement biologiques dont on se propose l'étude. Ainsi, peut-on craindre qu'une restriction volontaire de la natalité n'intervienne dans une famille quand est né, enfin, un héritier mâle; ou qu'au contraire une longue succession de filles ne décourage certains parents... A ces influences, qui viennent le plus immédiatement à l'esprit (et dont nous ignorons complètement la portée effective), s'en ajoutent certainement d'autres plus subtiles, plus variables et plus importantes. La composition actuelle d'une famille du point de vue de sa répartition en garçons et en filles n'est sûrement pas sans influencer sur l'attitude des progéniteurs dans un très grand nombre de cas. Aussi, pour éviter au maximum les distorsions par rapport aux processus physiologiques qu'occasionnent ces influences, avons-nous limité notre enquête aux familles nombreuses et plus précisément aux familles ayant eu plus de 5 enfants. Peut-être eut-il mieux valu prendre une limite plus élevée; malheureusement, la rapidité de la décroissance du nombre des familles en fonction de leur importance, nous eut exposé à un autre danger: celui de ne disposer que d'un échantillon trop restreint. En effet, dans la population, on constate qu'avec une régularité assez remarquable il y a à peu près 1,6 fois moins de familles de $n + 1$ enfant que de familles de n , pour n plus grand que 5 tout au moins.

TABLEAU N° I
ORIGINE DES DOCUMENTS

Nombre d'enfants par famille	Ministère de la Santé	Région parisienne	Grandes familles du Nord	Fondation Cognac-Jay	Total
5	3.309	1.387	432		5.128
6	2.214	740	301		3.255
7	1.298	374	205		1.877
8	1.269	211	139		1.119
9	682	109	81		872
10	527	48	59		634
11	313	27	34	10	384
12	159	12	15	32	218
13	68	3	6	31	108
14	26	3	1	29	59
15	14		2	21	37
16	9		1	12	22
17	5			6	11
18				2	2
19				4	4
Total des familles.	9.893	2.914	1.270	147	14.230

larité assez remarquable il y a à peu près 1,6 fois moins de familles de $n + 1$ enfant que de familles de n , pour n plus grand que 5 tout au moins.

D'autre part, la considération des naissances gémellaires et de leur répartition introduisant un élément étranger au but que nous nous proposons (et qui n'eut pu être éliminé que par des artifices mathématiques compliqués), nous n'avons recueilli que des familles n'en comportant pas. Ainsi notre échantillon comprend au total 14.230 familles françaises contemporaines sans jumeaux, soit 95.725 naissances. Le détail des sources est indiqué au tableau I :

— Service de la Médaille des Familles au Ministère de la Santé Publique et de la Population (9.893 familles).

TABLEAU N° II
RÉPARTITION NUMÉRIQUE DES FAMILLES
SELON LE NOMBRE DE LEURS ENFANTS ET LE NOMBRE DE GARÇONS

Enfants	GARÇONS															
	0	1	2	3	4	5	6	7	8	9	10	11	12	13		14
5	5.144	810	1.544	1.585	841	24										8.128
6	54	325	717	979	785	350	65									3.255
7	21	111	287	480	529	304	125	19								1.877
8	8	61	180	327	412	406	170	49	6							1.619
9		20	72	155	204	184	140	74	22	1						872
10		5	23	64	120	165	128	78	39	11	1					634
11		3	15	27	51	85	92	61	37	9	3	1				384
12		4	3	18	19	32	50	49	23	16	2	1	1			218
13		2	1	3	9	12	24	19	14	9	13	2				108
14				2	2	6	9	11	11	8	6	1				59
15					2	2	4	6	8	5	5	1	4			37
16					2	3	3	4	3	4	1	2	2			22
17						1		1	1	3	2	3	1			11
18									1				1			2
19											1	1		1	1	4

— Fichiers des Associations de Familles Nombreuses de la Région Parisienne (2.914 familles).

— Annuaire dit « Les grandes familles du Nord de la France » (édition 1948), (1.276 familles).

— Dossiers de la Fondation Cognacq-Jay (années 1920-1924 : 145 familles). Outre les personnes appartenant à ces divers services et qui ont bien voulu faciliter notre accès à leurs documents, nous tenons à remercier Mmes Godin, Deroche et Pouzol qui nous ont aidé à recueillir près de la moitié des familles au Ministère de la Santé Publique et de la Population. En plus de son aspect matériel, leur aide nous a été précieuse puisqu'elle nous a permis de contrôler l'objectivité de notre échantillonnage : les calculs limités aux seuls documents que nous n'avons pas recueillis personnellement conduisent dans l'ensemble aux mêmes conclusions que ceux que nous avons collectés nous-même. Ce point est plus important qu'il ne paraît puisqu'aussi bien une certaine part d'arbitraire était inévitable dans l'enregistrement d'un aussi grand nombre de familles et qu'on eut pu suspecter l'influence d'une tendance inconsciente en faveur de nos thèses dans les cas douteux (prénoms ambigus tels que Claude, Dominique, etc., élimination de documents de base suspects, ou même erreur pure et simple dans l'enregistrement).

LA VARIATION DE LA SEX RATIO
SELON LE RANG DE NAISSANCE

Préalablement à toute autre recherche, il était nécessaire de contrôler l'hypothèse fréquemment admise d'une variation systématique de la fréquence des garçons selon le rang de naissance (les premiers-nés seraient plus fréquemment du sexe masculin) et en fonction de l'âge de la mère (les femmes âgées ayant proportionnellement moins de filles). En ce qui concerne la première de ces théories, nos documents donnent une réponse très nette : si cette déviation existe, elle ne saurait être que très faible, puisque la proportion des garçons parmi les premiers-nés n'est que de 0,5097 contre 0,5062 aux autres rangs. Le léger écart constaté pourrait très légitimement être imputé à une simple fluctuation d'échantillonnage car l'écart type de la première fréquence est de 0,0042 et celui de la seconde de 0,0018. En ce qui concerne la deuxième hypothèse, nos documents se prêtent moins bien à une analyse rigoureuse. Cependant, il est clair que si elle était vraie on devrait constater une augmentation du nombre des garçons parmi les naissances de rang le plus élevé. On en déduit un test asymptotiquement équivalent à celui de J. B. S. Haldane et Cedric B. Smith (pour tester l'influence du rang de naissance sur l'incidence d'une maladie héréditaire) en calculant le rang moyen de naissance des garçons. Aucune déviation de la valeur observée à la valeur théorique que prévoit l'hypothèse de la constance de la probabilité n'a pu être établie de façon statistiquement significative. En présence de ce résultat négatif, qui retrouve une recherche antérieure de Wicksell sur les naissances en Suède (corrélation entre sex-ratio et âge maternel), nous supposons dorénavant que la probabilité de naissance d'un garçon ne varie pas systématiquement avec le rang de la naissance.

LA VARIANCE DE LA SEX RATIO
D'UNE FAMILLE A L'AUTRE

L'affirmation de la constance de la sex ratio dans l'espèce humaine, même si on tient compte de possibilités de variations systématiques telle que l'augmentation si souvent invoquée du nombre des garçons après les guerres, re-

couvre en réalité toute une série d'hypothèses entre les deux théories limites suivantes :

— à chaque naissance, la probabilité *a priori* de naissance d'un garçon a une certaine valeur, constante d'une famille à une autre ;

— au contraire, d'une famille à une autre, cette probabilité varie de façon importante et c'est seulement sa valeur moyenne qui est constante dans la population.

Pour étudier cette question, la statistique théorique montre qu'il suffit de posséder pour un échantillon assez nombreux non pas seulement la sex ratio totale, mais la composition de chacune des familles en garçons et en filles. En effet, si la première hypothèse était vraie, l'onde x devrait observer à peu près $N \binom{g+f}{p} \left[\frac{g}{p} \right]$ familles de g garçons et f filles si le nombre total des familles de $g+f$ enfants est N , et si la sex ratio générale est p . Au contraire, si la deuxième hypothèse se trouvait vraie, les valeurs probables correspondant aux familles où l'un des sexes prédomine nettement sur l'autre devraient être plus fortes que ne l'indiquent les formules précédentes. Sans entrer dans le détail ni des calculs ni des méthodes employées, nous nous bornerons à affirmer que, dans notre échantillon, les chiffres observés obligent à rejeter sans appel la première hypothèse : pour qu'une déviation aussi forte entre les valeurs qu'elle prévoit et les faits existant il eut fallut que se produisit un événement de probabilité inférieure à 1/100.000. Confirmant un résultat déjà obtenu par R. A. Fisher sur les matériaux réunis par Geissler, nous avons donc montré que *lorsque sont en causes des collections très importantes de familles, la théorie selon laquelle la probabilité a priori de naissance d'un garçon est une constante absolue ne peut être observée*. Pour donner une idée de l'importance de ce phénomène, indiquons que 65 familles de 6 enfants sans une seule fille ont été recueillies, alors que 55 seulement auraient dû être observées ; que nous avons trouvé 21 familles de 7 filles sans un seul garçon au lieu des 13 prévisibles dans l'hypothèse d'une probabilité identique pour toutes les familles, etc. Il n'est pas inutile d'insister sur le caractère, somme toute minime, de ces divergences et qui seraient passées inaperçues dans une enquête moins étendue (quelques milliers de familles seulement).

TABLEAU N° III
VALEURS CARACTÉRISTIQUES DES HISTOGRAMMES

Nombre d'enfants	Nombre de familles	Fréquence des naissances masculines	Variance observée	Ecart de la variance (t)
5	5.128	0,5085	1.300	+ 4,6
6	3.255	0,5057	1.595	+ 6,3
7	1.877	0,5061	1.892	+ 6,5
8	1.619	0,5011	2.069	+ 2,7
9	872	0,4980	2.535	+ 10,7
10	634	0,5178	2.583	+ 1,9
11	384	0,5066	2.976	+ 3,8
12	218	0,5022	3.667	+ 8,0
13	108	0,5207	4.090	+ 2,9
14	59	0,5170	3.322	— 1,0
15	37	0,5441	4.095	+ 4,3
16	22	0,4682	5.110	+ 3,7
17	11	0,6112	3.072	— 0,8
18	2	0,5566	8.000	+ 3,5
19	4	0,6316	3.333	— 0,1

(t) Ecart entre valeur observée et valeur théorique exprimé en écart-type.

TABLEAU N° VI

COMPARAISON ENTRE LES HISTOGRAMMES OBSERVÉS ET LES SCHEMAS PROBABILISTES

Nombre de garçons	Nombre observés	Nombres théoriques dans l'hypothèse binominale	Nombres théoriques dans l'hypothèse de Markoff.	
FAMILLES DE CINQ ENFANTS				
6	204	171	190	chi carré = 9,73
4	841	833	852	
3	1.585	1.624	1.386	chi carré = 5,88
2	1.544	1.581	1.543	
1	810	769	788	
0	144	150	169	
FAMILLES DE SIX ENFANTS				
6	65	55	63	chi carré = 9,75
5	330	321	336	
4	785	783	776	chi carré = 2,24
3	979	1.017	987	
2	717	743	733	
1	325	289	304	
0	54	47	54	
FAMILLES DE SEPT ENFANTS				
7	19	16	19	chi carré = 15,00
6	126	109	117	
5	304	320	323	chi carré = 8,16
4	529	520	507	
3	460	507	493	
2	327	296	208	
1	111	96	104	
0	21	13	10	
FAMILLES DE HUIT ENFANTS				
8	6	7	5	chi carré = 11,34
7	49	55	60	
6	170	185	190	chi carré = 14,44
5	406	356	351	
4	412	441	432	
3	327	348	343	
2	180	171	176	
1	61	47	52	
0	8	6	9	

DISCUSSION DU SCHEMA DE MARKOFF

La première interprétation de ces résultats, celle qui correspond le plus étroitement au formalisme analytique, est l'hypothèse d'une influence réelle du sexe d'un enfant sur le sexe de l'enfant à la grossesse suivante : pour étayer biologiquement cette théorie, on pourrait faire appel à l'exemple du *free martinisme* où, en effet, une semblable action (mais contemporaine !!!) est indiscutable. Il faudrait alors admettre une sorte d'imprégnation de l'organisme maternel, persistant quelques mois ou plutôt quelques années et pouvant influencer le sexe de l'embryon suivant. Plus vraisemblablement, pourrait-on penser que cette imprégnation augmente électivement le nombre de ces avortements précoces qu'a étudiés Parker quand le sexe du nouveau fœtus diffère de celui de l'enfant précédent. Mais il est une autre interprétation de ces phénomènes qui nous semble plus aisée à admettre dans l'état actuel de nos connaissances. Supposons en effet que, sous l'influence de causes extérieures variées, l'organisme maternel présente des périodes de durée et d'intensité aléatoires pendant lesquelles est favorisée la conception d'enfants de tel ou tel sexe, ou bien est augmentée la létalité des embryons mâles ou femelles (on pourrait songer par exemple à des modifications du pH utérin ?). Si ces périodes sont assez longues pour chevaucher de manière relativement fréquente sur

plusieurs grossesses, leur action se manifesterait en première approximation par des naissances réparties selon un schéma de Markoff. A vrai dire, si ces influences étaient très importantes, ou bien si leur fréquence ou leur durée compensaient leur faible action, il serait possible d'en tester rigoureusement l'existence. De fait, l'ordre de grandeur de l'incrément de probabilité auquel nous avons abouti ne permet guère de pousser plus à fond la discussion, à moins de pouvoir doubler ou tripler le volume des documents de base, ce qui semble pratiquement irréalisable en France actuellement. Par conséquent, il nous faut faire des hypothèses qui peuvent sembler parfaitement arbitraires pour évaluer la portée réelle de l'action de ces périodes gynophiles et androphiles. Supposons que les périodes aient une durée moyenne de deux gestations avec une probabilité de $p+x$ et de $p-x$. Pour obtenir un incrément de liaison de 0,015, il suffit que $x=0,146$. Autrement dit, si l'on veut bien admettre que la probabilité de naissance d'un garçon fluctue lentement et de manière aléatoire entre 0,35 et 0,65 environ, mais de telle sorte que les deux tiers des naissances en moyenne se trouvent dans la même phase, cela suffit pour rendre compte de la corrélation que nous avons observée. Si ces périodes étaient en moyenne deux fois plus longues, il suffirait que la probabilité oscillât entre 0,40 et 0,60. Pour discuter l'hypothèse qui nous semble la plus vraisemblable : seules certaines femmes présenteraient de semblables fluctuations de la probabilité de naissance d'un garçon, il faudrait faire encore de nouvelles hypothèses qui perdraient alors tout sens. Cette dernière théorie a cependant pour elle et d'expliquer mieux la constance relative de la sex ratio générale et le fait que la liaison résiduelle entre naissance aux rangs n et $n+2$ est très faible, sinon nulle, résultat peu compatible avec l'existence relativement fréquente de périodes de longue durée.

CONCLUSIONS

Arrivé au terme de cet exposé, il convient de faire le point des résultats obtenus et des perspectives qu'ils peuvent ouvrir à de nouvelles recherches.

1° Il ne semble pas que la fréquence des garçons varie de façon importante ni systématique selon le rang de naissance ;

2° L'hypothèse d'une distribution des sexes selon un schéma binomial (c'est-à-dire spécifiant la constance de la probabilité et l'indépendance des naissances) ne peut être conservée ;

3° L'hypothèse d'une liaison simple entre les sexes dans les naissances consécutives rend compte convenablement des faits observés et s'interprète de façon satisfaisante par l'existence de périodes aléatoires gynophiles et androphiles ;

4° Il ne semble pas possible en se limitant à une formation de même nature que celle utilisée ici (succession des sexes dans les familles), d'obtenir autre chose qu'une confirmation ou une infirmation de nos thèses.

Par contre, en faisant usage d'autres éléments, et en particulier de la durée de l'intervalle qui sépare les naissances, il serait sans nul doute possible de préciser les résultats déjà obtenus et de serrer de plus près les hypothèses entre lesquelles il est encore difficile de choisir. Sous la direction de M. le Professeur R. Turpin, ces travaux sont actuellement en cours au Centre de Génétique de l'hôpital Saint-Louis et laissent espérer de rapides progrès dans la connaissance des mécanismes qui déterminent le sexe dans l'espèce humaine.

REFERENCES

R. TURPIN et M. P. SCHÜTZENBERGER. — Recherche statistique sur la distribution du sexe à la naissance. *Comptes rendus à l'Académie des Sciences*, 226. 31-5-1948.