

L'univers étudié

L'univers observé est l'ensemble des femmes non célibataires (mariées, veuves, divorcées), nées en 1925 et après, c'est-à-dire ayant moins de 46 ans au 1^{er} janvier 1971.

Il y a environ 6 millions de femmes dans l'univers étudié (voir tableau 1). La plupart des femmes sont mariées.

Jusqu'à 45 ans, les veuves et les divorcées non remariées sont rares ; la plupart des ruptures d'union, par veuvage ou divorce, donnent lieu à des remariages. Le tableau 1, comme le recensement de la population, ne distingue pas les remariées des mariées pour la première fois.

L'entrée dans la population des femmes mariées se fait progressivement avec l'âge ; malgré les forts effectifs des générations nées depuis 1945, et la nuptialité précoce de ces générations, on compte moins de femmes mariées avant 25 ans, qu'au dessus : la pyramide des âges des femmes non célibataires va s'élargissant de 20 à 45 ans.

Une telle structure va directement ^{peser} opérer sur le plan de sondage adopté.

L'échantillon des femmes

On se propose d'observer environ 4000 femmes.

Le tirage de l'échantillon, en assurant aux femmes des chances égales d'être interrogée par l'enquête, pourrait se faire au taux constant de :

$$\frac{4.000}{6.000.000} = \frac{1}{1.500}$$

La structure par âge de l'univers étant respectée dans l'échantillon, on disposerait de jeunes femmes interrogées. Ce serait

là un inconvénient grave : ces jeunes femmes sont en pleine période de fécondité ; et l'évolution du nombre des naissances en France dans les dix prochaines années dépend du comportement qu'elles adopteront.

On a donc décidé d'adopter un taux de sondage plus fort dans les classes jeunes, de façon à disposer de deux sous-échantillons d'effectifs voisins, dans les deux groupes d'âge :

femmes jeunes, nées 1940 et après : taux 1/1000

femmes plus âgées, nées 1928-39 : taux 1/2000

Les effectifs à peu près égaux des deux sous-échantillons permettent une étude aussi précise des plus jeunes que des moins jeunes. L'égalité des effectifs rend maximale la précision des comparaisons entre les deux groupes d'âges.

Dans les exploitations d'ensemble, les femmes plus âgées, deux fois moins représentées, comptent deux fois plus que les jeunes. Cette pondération, proportionnelle à l'inverse des taux de sondages (1/1000, 1/2000) assure une description non biaisée de l'ensemble des femmes non célibataires de moins de 46 ans.

Difficultés d'observation

Dans la pratique le tirage d'un échantillon de femmes non célibataires se heurte à plusieurs difficultés. Au premier abord, la meilleure base de sondage semble être formée des bulletins individuels du recensement de 1968. Or, cette base n'est pas à jour ; elle diffère de l'univers étudié, fin 1970, sous deux aspects :

1 - Au point de vue matrimonial : les femmes dont le mariage se situe entre le 1^{er} mars 1968 et la fin de 1970, sont encore célibataires, (sauf quelques exceptions) au recensement de 1968. Il serait particulièrement gênant d'écarter les jeunes mariées de l'enquête ; on écarterait de l'observation directe la durée des 3 premières années de la vie conjugale. Seule

l'observation rétrospective serait possible.

2 - Au point de vue migratoire, les femmes qui déménagent entre le recensement et l'enquête seraient ^{difficile à observer.} presque toutes perdues pour l'enquête. L'ampleur de cette ^{la} mobilité est plus grande, particulièrement pour les jeunes adultes ; et la dimension de la famille est une des raisons de cette mobilité. Il faut donc craindre, en ^{ne} disposant que de l'adresse au recensement, vieille de près de 3 ans, que ~~ne~~ s'échappe à l'observation une partie de l'univers entraînant une déformation systématique des résultats.

Echantillon de logements

La solution adoptée est la suivante : constituer un échantillon de tous les logements, et observer dans les logements tirés, toute femme non célibataire née en 1925 et après. ^{au} L'échantillon des ^{tous les} logements inclut à l'échantillon des ^{entre les} femmes au moment de l'enquête qui est pris en compte : ainsi les femmes le plus récemment mariées peuvent être observées.

De même, les femmes le plus récemment arrivées dans un nouveau logement sont interrogées ; la mobilité des ménages et des personnes n'est donc plus une cause de déformation systématique de l'échantillon observé.

L'échantillon de logements permet donc d'atteindre un échantillon sans biais, de toutes les femmes nées en 1925 et après, qui ne sont pas célibataires à la date de l'enquête.

Dans de nombreux logements, il n'y a ^{aura pas de} à l'enquête aucune femme non célibataire née en 1925 et après ; il n'est nullement nécessaire en théorie de visiter ces logements. Mais dans la pratique, cette visite est ^{indispensable} nécessaire pour connaître les occupants à la date de l'enquête, ^{pour} et savoir s'il y a, ou non, une femme non célibataire. On n'échappe pas à l'obligation de visiter un grand nombre de logements beaucoup plus grand que le nombre de femmes enquêtées.

Environ 2 logements sur 3 n'abritent pas de femmes à étudier.

On a cherché à éviter une dépense excessive dans l'étude de cette catégorie de logements, pour consacrer une plus grande part des crédits à l'étude des femmes. Bien que l'on ignore, avant l'enquête, la composition du ménage qui habitera chaque logement au moment de l'enquête, il est possible de s'en faire quelque idée à partir des résultats du recensement de 1968 : malgré les mouvements démographiques (mortalité, nuptialité, divorcialité, natalité), et les mouvements migratoires, il y a une corrélation assez forte entre la composition des ménages de l'échantillon de logements au recensement, et à la date de l'enquête.

On classe dans une strate particulière les logements qui n'abritent pas de femmes membres de l'univers étudié, à la date du recensement ; cette strate est étudiée avec un taux de sondage plus faible : 1/4000.

A partir du recensement, on a finalement constitué trois strates de logements, étudiées respectivement avec les taux 1/1000, 1/2000 et 1/4000. (tableau 2)

Une dernière strate est celle des logements neufs, dont la construction n'était pas achevée au recensement. Les enquêtes sur les logements non agricoles, ou dans les grands ensembles, laissent prévoir qu'il y aura presque toujours une femme mariée dans ces logements neufs ; on étudie cette strate au taux de 1/1000.

Le tableau 2 laisse prévoir les effectifs des logements à visiter ; De façon très conjecturale, on donne une prévision du nombre de femmes de moins de 46 ans, non célibataires lors de l'enquête.

Tableau 2 - Stratégie des logements (évolution sommaire)

	Univers : milliers logements	Echantillon provisoire taux unique 1/1000	Echantillon final		
			taux par catégorie	taux final	logement à observer
Logements où il y avait au 1/3/68					
<u>Au moins 1 femme non célibataire d'âge (au 1/1/1969) :</u>	milliers	unités			
- moins de 29 ans A	1.760 ^(a)	1.760	1/1	1/1000	1.760
- 29-44 ans B	4.314 ^(a)	4.314	1/2	1/2000	2.157
<u>Aucune femme non célibataire de moins de 45 ans au 1^{er}/1/1969 :</u>					
- Résidences principales ..C1	9.704 ^(b)	9.704	1/4	1/4000	2.425
- Résidences secondaires et logements vacantsC2	2.478	2.478	1/4	1/4000	620
	18.256	18.256			6.953
<u>Logements achevés depuis 1968 (ordre de grandeur) :</u>					
D	1.200	1.200	1/1	1/1000	1.200
Tous logements à visiter par les enquêteurs :					8.163
Evolution du nombre de logements habités (c)					7.443
Observations utiles (d). Logements avec une femme non célibataire née en 1925 et après					4.868
Observations perdues pour l'enquête					3.295

(a) On suppose dans ce tableau, qu'il n'y a jamais, dans un même logement, deux femmes non célibataires de moins de 44 ans. Ce qui permet de poser l'égalité : nombre de femmes = nombre de logements, dans les catégories A et B.

(b) sur 18.256.000 logements, 15.778.000 sont des résidences principales, les autres sont des résidences secondaires ou des logements vacants, à la date du recensement (2.478.000)

Le calcul de C1 est fait à partir des résidences principales.

$$15.778 - (1.760 + 4.314) = 15.778 - 6.074 = 9.704$$

(c) Ordre de grandeur en supposant que les logements C2 sont inhabités (620) ainsi que 100 autres de l'échantillon.

(d) On suppose que l'on trouvera une femme non célibataire née en 1925 et après :

- dans 90% des logements A et B
- dans 10% des logements C1
- dans aucun des logements C2
- dans 90% des logements D

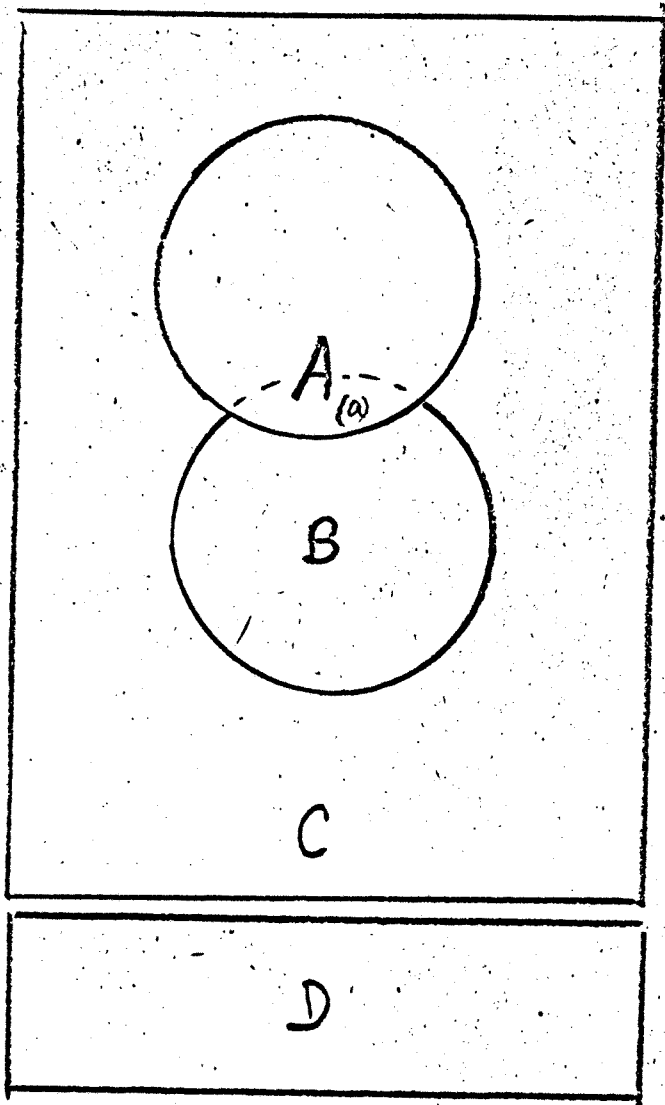


Schéma des
4 strates de
logement
A, B, C, D.

(a), dans certains
logements A, il y a
au moins une femme jeune
et une femme plus âgée

Figure 1.

Les taux (1/1000, 1/2000 et 1/4000) déterminant les poids (1,2,4) par lesquels on multipliera toute observation faite dans ces logements.

Il faut noter que les logements de la strate A, où il y a eu recensement au moins une femme née en 1940 et après, non célibataire en 1968, peuvent abriter, à la même date, d'autres femmes non célibataires plus âgées. La strate B, constituée après A, ne comprend aucun logement où il ~~avait eu recensement, une ou plusieurs femmes non célibataires, nées en~~ 40 et après. (Voir figure .1)

L'observation

Puisque l'échantillon des femmes est inclus dans un échantillon de logements, plus vaste, le travail de l'enquêteur comprend donc deux parties :

1 - rechercher le logement tiré au sort et établir la composition des ménages qui y vit. S'il n'y a pas de femmes nées en 1925 et après, non célibataires à la date de l'enquête, l'observation est terminée.

2 - Sinon : interroger toute femme non célibataire, née en 1925 et après, et vivant dans le logement :

Deux formulaires sont utilisés :

- un questionnaire "logement" rempli pour tout logement de l'échantillon
- un questionnaire " femme", qui peut exister à plusieurs exemplaires, pour un même logement, ou ne pas exister, selon le nombre de ~~femmes non célibataires, nées en 1925 et après.~~ ^{à étudier}

Des observations complémentaires peuvent ^{aussi} actuellement être faites : questionnaire rétrospectif aux femmes de 46 à 65 ans, question/d'opinion posées à d'autres catégories de la population.

Sondage à deux degrés

Un échantillon de logement, tiré dans chacune des quatre strates (A,B,C,D) conduirait à une très grande dispersion de la collecte sur le territoire.

Pour éviter cet inconvénient, le plan de sondage est articulé en deux degrés (2):

- 1 - tirage d'unités primaires (lieux de collecte)
- 2 - tirage des logements (unités secondaires) dans les bases de sondages des unités primaires tirées.

Dans le cas présent, on a décidé de ne pas disperser l'observation en plus de 100 unités primaires.

Une unité primaire est, selon les cas :

- soit une unité urbaine définie par le recensement de 1963 : agglomération urbaine multicommunale, ou ville isolée.
- soit un canton totalement formé de communes rurales,
- soit l'ensemble des communes rurales d'un canton partiellement urbain (c'est-à-dire où il y a au moins une commune appartenant à une unité urbaine.)

L'agglomération parisienne est désignée d'office, ainsi que les six plus grandes agglomérations de province : LYON, MARSEILLE, LILLE-ROUBAIX-TOURCOING, BORDEAUX, NICE, TOULOUSE. Dans chaque agglomération, l'échantillon de logement est tiré avec un taux constant dans chaque strate : (1/1000, dans les strates A et D ; 1/2000 dans la strate B; 1/4000 dans la strate C).

On trouvera le détail des opérations en annexe.

Pour les autres unités primaires, l'ordre des opérations a été le suivant :

- 1 - Stratification des unités primaires
- 2 - Tirage des unités primaires
- 3 - Tirage d'un échantillon provisoire dans l'unité primaire.
- 4 - Stratification des échantillons provisoires
- 5 - Tirage de l'échantillon à observer.

(2) ^{et} Elaboration de ce plan, le tirage des unités primaires et secondaires, a été effectué sous la direction de M. F. Chartier, chef de la division " méthodes statistiques " de l'I.N.S.E.E.

Tableau 3 - Strates des unités primaires

N° de strate		Nbre unités Pri-maires	Echantillon des logements tirés				TOTAL à visé
			Recensement 1968			Logements neufs	
			Echantillon provisoire		Echantillon final		
			1/1000				
			Par Unité (a)	Par Strate (b)	par strate (c)	(d)	(e)
6	Sept agglomérations retenues d'office :						
	PARIS		3.181		1.214		
	LYON		376		143		
	MARSEILLE		343		131		
	LILLE-R.T		294		112		
	BORDEAUX		193		74		
	NICE		172		66		
	TOULOUSE		151		58		
	Total :	7	4.710	4.710	1.798	310	2.100
5	Unités urbaines de 200.000 habitants et plus	9	154	1.386	528	91	619
4	- - de 100 à 200.000	9	149	1.341	511	88	599
3	- - de 20 à 100.000	16	150	2.400	915	158	1.073
2	- - de moins de 20.000	18	146	2.628	1.002	174	1.176
1	Parties rurales de cantons partiellement urbains	19	149	2.831	1.080	185	1.265
0	Cantons totalement ruraux	20	148	2.960	1.122	194	1.323
		98		18.256	6.963	1.200	8.163

(a) Dans les strates n° 6, 5 et 4, des tirages de communes ou de quartiers ont été faits dans certaines unités urbaines.

(b) effectifs exactement tirés dans chaque strate, au 1/1000.

(c) Evaluation en prenant 38/3 % de (b). Les effectifs exacts seront connus après tirage.

(d) Evaluation très grossière, prévue pour le total des strates.

(e) Evaluation grossière, cumulant les imprécisions de (c) et (d).

Les autres unités urbaines et les cantons ruraux, sont regroupés en strates, d'après l'effectif de la population (tableau. 3)

Dans chaque strate, les unités primaires sont classées par département, et ceux-ci par région ; ce qui assure un balayage systématique du territoire lors du tirage au sort.

Les unités primaires sont de tailles très inégales : on désire cependant observer à peu près un même nombre de logements dans chacune : environ une soixantaine. La précision des estimations serait très faible, au dessous et l'organisation de la collecte ^{plus} difficile.

En réalité on a ^{constitué} estimé des échantillons provisoires de 150 logements environ, dans lesquels on tire après coup un échantillon de 60 (voir annexe 2).

Taux de sondage

Une fois les strates ^(s) constituées, on connaît le nombre total de logements qu'elles contiennent (N_s) ; on ~~va~~ constitue un échantillon provisoire de logements (n_s)

$$n_s = \frac{1}{1000} N_s$$

Cet échantillon provisoire doit être réparti en unités primaires de 150 logements environ. Ce qui donne le nombre (m_s) ^{d'}unités primaires à tirer dans la strate (s) :

$$m_s = \frac{n_s}{150} = \frac{N_s}{150 \times 1000}$$

Dans la strate (s), où le nombre total des logements est N_s , chaque unité primaire (i) a un nombre de logements N_{si} , connu par le recensement.

On tire les unités primaires avec une probabilité proportionnelle à la part des logements qu'elles détiennent parmi tous ceux de la strate : N_{si} / N_s .

La probabilité qu'un logement de la strate s soit tiré est constante et égal à $1/1000$; c'est le produit des probabilités :

- que l'unité primaire soit tiré (~~dans le groupe de m_s~~)
- que le logement soit tiré

$$f = m_s \times \frac{N_{s1}}{N_s} \times \frac{150}{N_{s1}} = \frac{1}{1000}$$

Ainsi le taux unique de l'échantillon des logements permet-il d'exploiter sans pondération l'ensemble des strates d'unités primaires (catégorie de localités). *(feuille 3)*

Mais, dans chaque unité primaire, demeure la stratification en 4 lots A, B, C, et D ; dont les taux de sondage sont différents, comme on l'a vu. *(feuille 2)*

Les figures 2 et 3 montrent la répartition des unités primaires ^{tirées} sur le territoire, et au sein de l'affiliation par commune (annexe 1).

Enquête "Fécondité"

1^{er} juin 1970
Unités Primaires



Figure 2



Annexe 1

Détail du sondage dans la strate 6

Ville de Paris

Au taux de 1/1000, l'échantillon provisoire de logement dans le recensement donnerait 1.222 logements.

Pour concentrer la collecte sur le terrain, on a tiré un échantillon de 20 quartiers sur 80, au taux de 1/4. L'échantillon des logements dans chaque quartier tiré, est constitué avec le taux de 1/250 ; on a encore :

$$1/4 \times 1/250 = 1/1000$$

Le tirage des quartiers se fait après constitution de 20 strates de 4 ; et tirage d'un quartier par strate . Les variables de stratification ont été :

- x : un indicateur de structure par âge, lié à la fécondité : rapport des effectifs féminins de 0-4 ans à l'effectif féminin de 25-34 ans (au recensement de 1969).
- y : un indicateur socio-professionnel : proportion des chefs de ménages "ouvriers spécialisés et manoeuvres" parmi tous les chefs de ménages ordinaires du quartier.

On regroupe les quartiers dont les points représentatifs sont proches sur ^{un} le graphique en x et y.

Un premier tirage de 20 quartiers rassemblait, dans l'échantillon au 1/500 : 1.038 logements, Résultat toujours inférieur à 1222.

San multistation elections, le résultat est et
En procédant à un nouveau tirage de 2 quartiers, dans 2 strates convenablement choisies, on obtient un échantillon numériquement acceptable : 1218 logements.

Banlieue proche de Paris

26 communes proches de Paris sont traitées ensemble ; ce sont les plus anciennement urbanisées, et les plus peuplées de la banlieue.

Un échantillon simple au 1/1000 donnerait 679,logements.

Pour concentrer la collecte, on décide de tirer 7 communes sur les

26. Chacune des 7 sera observée avec un échantillon de $679 : 7 = 97$ logements.

On fait le tirage des 7 communes en donnant à chaque commune des chances proportionnelles à son nombre de logement (L_c); ce qui assume le taux final de $1/1000$:

$$7 \times \frac{L_c}{679.000} \times \frac{97}{L_c} = \frac{7 \times 97}{679.000} = \frac{1}{1000}$$

Reste de la Banlieue de Paris

En dehors des 26 communes traitées ci-dessus, les communes de la banlieue parisienne rassemblent 1.280.000 logements en 1968. On a ^{utilisé} ~~considéré~~ les deux procédés exposés ci-dessus :

1) échantillon de 63 communes (en réalité 64) 2 petites sont regroupées en une ^{seule} unité, avec probabilité proportionnelle au nombre de logements, et nombre presque constant à observer ≈ 20 ^{log} (en fait en moyenne 20,85). 2) Pour augmenter la liste à observer dans chaque commune tirée, on a constitué un échantillon de 16 communes sur les 64 $\times (1/4)$, après avoir stratifié les 64 communes en 16 groupes de 4 ^(taux 1/4).

Pour rapprocher les communes par ressemblance, on a utilisé notamment les résultats d'un classement factoriel des communes fait par l'I.A.U.R.P. ⁽¹⁾.

On tire dans chaque commune ^{un} échantillon ^{de} : $20,25 \times 4 = 81$ logements.

$$63 \times \frac{L_c}{1.280.000} \times \frac{1}{4} \times \frac{81}{L_c} = \frac{1}{1000}$$

Banlieues des six plus grandes agglomérations de province

Dans ^{les} agglomérations, désignées d'office, la ville centre est observée à partir d'un échantillon simple, à taux unique :
Echantillon provisoire au $1/1000$.

Certaines communes sont traitées avec la ville centre ; Villeurbanne avec Lyon, Mérignac avec Bordeaux, Roubaix et Tourcoing avec Lille etc...

(1) "Comparaison et classification des communes de l'agglomération parisienne" : 106 p. Cahier n°3 de l'Institut d'Aménagement et d'Urbanisme de la Région parisienne.

Le reste de la banlieue est traitée comme la banlieue de Paris dans le paragraphe précédent :

- Tirage de communes, proportionnellement au nombre de logement, dans l'ensemble de la banlieue;
- Tirage d'une commune sur 4, parmi le premier échantillon.

On a constitué des strates de 4, en mélangeant les agglomérations, en fonction des critères x et g, utilisés dans la ville de Paris.

Tous les échantillons au 1/1000 sur provisoires.
On applique ensuite la procédure de tirage de
l'échantillon définitif (annexe 2).

Tirage dans l'échantillon provisoire au 1/1000

Pour chaque unité primaire, les logements appartiennent aux trois strates A, B et C.

Comme il serait trop onéreux de stratifier chaque unité primaire tirée, puis faire un échantillon dans chacune.

On a procédé ainsi :

- tirage non stratifié d'un échantillon simple de tous les logements de l'unité primaire, au taux de 1/1000.

- stratification de cet échantillon (et non de la totalité de l'unité primaire) en trois catégories A, B, C. La répartition obtenue en A, B, C, n'est donc qu'une estimation de la répartition réelle de l'univers de l'unité primaire.

- La strate A de l'échantillon est conservée en entier (échantillon au 1/1000); ^{dans} la strate B, on retient un logement sur deux, ce qui donne un échantillon au 1/2.000 ; de la strate C, on retient un logement sur quatre (échantillon 1/4.000).

On ignore la distribution de chaque unité primaire en 3 strates (A, B, C). Pour l'ensemble de la France, le tableau 2 montre que l'échantillon des logements à observer sera 38,13 % de l'échantillon provisoire au 1/1000; 40% en arrondissant.

Pour observer 60 logements, il faut que l'échantillon provisoire ~~en~~ contienne, en moyenne :

$$60 \times \frac{100}{40} = 150 \text{ logements}$$

C'est la taille retenue pour l'échantillon provisoire, constant dans chaque unité primaire des strates 5 à 0.

Effectifs des logements attendus (a)

(b) Sans compter les logements construits depuis le Recensement