Dans des études généralistes, on est en général plus proches de la représentation proportionnelle (avec de légères sur-représentations de certaines formules pédagogiques intéressantes mais relativement rares). Mais nous verrons au chapitre 10 le cas d’études centrées sur un thème précis d’où découlent des choix de population d’étude et de stratification assez éloignée de la logique de représentation proportionnelle.

D. L’échantillonnage en grappes

1. Le principe :

L’échantillonnage en grappes est celui pratiqué en général dans les enquêtes scolaires par échantillonnage. Pourquoi ?

- L’échantillon aléatoire simple est difficilement réalisable et en tout cas coûteux. Par ailleurs, il ne garantit pas à tout coup une bonne représentation de l’ensemble de la population scolaire (voir section précédente).

- L’échantillonnage stratifié, proportionnel ou non, permet de mieux contrôler la représentativité de l’échantillon tiré. Mais si on l’emploie seul, la difficulté de réalisation est encore accrue, puisqu’il s’agirait d’aller tester un par un des élèves dispersés sur l’ensemble du territoire.

- Par ailleurs, nos enquêtes sont rarement seulement centrées sur les élèves. Nous voulons en général aussi pouvoir comparer des classes entre elles (en particulier pour mesurer les effets des types de classes et des différentes caractéristiques des maîtres). On voit bien que de telles comparaisons seront très imprécises si on ne dispose que d’un élève dans chaque classe de notre échantillon[1] .

Que faire alors ? La technique de l’échantillonnage en grappes procède en plusieurs étapes, mêlant des ingrédients des techniques de stratification et de tirage aléatoire simple :

[1] Dans certains cas, l’intérêt est tellement centré au niveau des classes que les classes (plutôt que les élèves) constituent la véritable population d’intérêt. Mais on se trouve souvent dans le cas hybride où on veut tirer une information pertinente à la fois sur l’ensemble des classes et l’ensemble des élèves d’un pays.

- **1ère étape : la stratification :**
  Afin de s'assurer que les principales dimensions de la population scolaire qu'on veut étudier sont bien représentées, on procède d'abord à la définition de strates selon une ou plusieurs dimensions. La définition des strates ainsi que le choix de leur représentation (proportionnelle ou non) dépend des priorités de l'étude, comme on l'a vu dans la section précédente.

- **2ème étape : le tirage des classes :**
  A partir des listes de classes, on tire au sort au sein de chaque strate un certain nombre de classes. Si on connaît le nombre d'élèves dans chacune des classes, on tire au sort les classes en faisant en sorte que chaque classe ait une probabilité d'être tirée proportionnelle au nombre de ses élèves. C'est ainsi qu'on s'assure que chaque élève d'une strate a la même probabilité de figurer dans l'échantillon final.

- **3ème étape : le tirage des élèves :**
  Au sein de chaque classe, on tire au sort un certain nombre d'élèves.

On parle d'**échantillonnage en grappes** à cause de l'étape 2 : en effet, **on ne tire pas alors des individus, mais des groupes d'individus (des «grappes»)**.

- L'intérêt pratique est immédiat : l'administrateur qu'on envoie dans une école ne se déplacera pas pour un seul élève, mais testera un ensemble d'élèves de cette classe. Cela vaudra la peine de recueillir des informations sur la classe, son maître, l'environnement (par les questionnaires au maître et au directeur) puisque, à chaque fois, l'information pourra être rapportée à plusieurs enfants.
- Mais on peut déjà anticiper une perte de précision : en effet, les élèves d'une même classe ont toutes chances de se ressembler un peu. Vingt élèves d'une même classe ne nous donnent pas autant d'informations sur la diversité de la population étudiée que vingt élèves pris dans vingt classes différentes. Autrement dit, l'échantillon en grappes représente moins bien la diversité de la population ; il se traduit (par rapport à l'échantillon aléatoire simple et à taille d'échantillon constante) par une perte en précision pour l'estimation des caractéristiques de la population.

2. Un exemple d'échantillonnage en grappes : l'enquête du PASEC au Sénégal en 1995-96 :

L'évaluation du système éducatif sénégalais commanditée par la CONFEMEN a utilisé les techniques d'échantillonnage en grappes.
On peut citer ici le texte qui explique les choix réalisés dans cette expérience concrète[1] :

«Le choix d'un échantillon dépend de la question à laquelle on désire répondre par le moyen de l'enquête. Compte tenu des choix de la CONFEMEN, mais aussi des contraintes du terrain, pour construire l'échantillon, il fallait prendre en considération les contraintes suivantes :

- *On désire obtenir une évaluation du niveau des élèves de CP et de CM1. L'échantillon doit* ***être représentatif des élèves****.*

- *Il n'est pas possible d'envisager que certaines régions soient exclues de l'enquête. L'échantillon doit donc* ***couvrir tout le pays.***

- *On désire avoir des indications sur l'influence du mode d'organisation des écoles et des classes.* ***L'échantillon doit donc tenir compte de la diversité des situations****, même si certaines d'entre elles sont très minoritaires.*

  [conséquence : **échantillon stratifié** en fonction des différents modes d'organisation des classes]

- *Faire un échantillon direct d'élèves serait trop coûteux (on ne dispose pas de la liste exhaustive des élèves) et on n'obtiendrait alors aucune information sur l'influence de la classe et de l'école.* ***Il faut donc tirer d'abord un échantillon d'écoles.***

  [conséquence : **échantillon en grappes** et perte de précision par rapport à un échantillon aléatoire simple]

---

[1] *Evaluation du système éducatif sénégalais. Enseignement primaire. Rapport intermédiaire - juin 1995.*(Institut national d'étude et d'action pour le développement de l'éducation [INEADE, Dakar] et Centre international d'études pédagogiques [CIEP, Sèvres], sous la direction d'Emilie Barrier).
C'est nous qui ajoutons les remarques entre crochets.

- *Les écoles et les classes peuvent être de tailles très différentes. Pour que l'échantillon des élèves soit représentatif, le tirage des écoles doit se faire* ***proportionnellement à la taille de l'école*** *mesurée en nombre d'élèves concernés.*

  [conséquence : **chaque élève de la population a les mêmes chances de se retrouver dans l'échantillon final**]

- *Les données obtenues, en particulier celles concernant le niveau national, doivent être d'une précision raisonnable. En particulier, l'échantillon doit permettre d'obtenir des* ***indications fiables avec des sous-groupes****, comme par exemple, celui des classes à double flux. L'échantillon ne doit pas être trop petit.*

  [conséquence : une **sur-représentation de certaines strates peu nombreuses mais intéressantes** peut être souhaitable]

- *Les ressources financières sont limitées. L'échantillon ne doit donc pas être trop grand.*

- *Deux niveaux sont testés (CP et CM1). Pour des raisons d'économie, il convient de tirer des écoles contenant les deux niveaux, mais sans négliger celles qui n'en contiennent qu'un seul (écoles à classe unique souvent).*

- *Les élèves au sein d'un établissement se ressemblent plus entre eux qu'ils ne ressemblent à ceux d'un établissement voisin (du fait de l'influence de leur maître par exemple). Il faut tenir compte de ce phénomène (design effect).*

Toutes ces considérations et les données de la théorie des échantillons nous amènent à adopter les règles suivantes :

- *On tirera un échantillon d'au moins 100 écoles pour tenir compte de l'effet de grappe (design effet), avec une probabilité égale au nombre d'élèves de la classe concernée ou du nombre d'élèves total en CP et CM1.*

- *Certaines strates seront sur-représentées, pour avoir assez d'élèves dans ces strates.*

- *On tirera une classe de CP et une classe de CM1 dans chaque école (si elles existent)*

- *On tirera au hasard 20 élèves dans chaque classe.*

Les strates et l'échantillon retenus sont présentés dans le tableau ci-dessous :

Echantillonnage pour le PASEC Sénégal :

Numéro de la strate	Identification de la strate	Nombre d'écoles	Nombre d'élèves	Nombre d'élèves de CP	Nombre d'élèves de CM1	% d'élèves	% d'élèves de CP	% d'élèves de CM1	Ecoles avec CP dans l'échantillon	Ecoles avec CM1 dans l'échantillon
1	Ecoles à CM1 seul, CM1 multigrade	19	1 304	0	300	0,18%	0%	0,27%	0	5
2	Ecoles à CM1 seul, CM1 simple flux	274	30 795	0	9 263	4,21%	0%	8,38%	0	5
3	Ecoles pilotes, avec CP et CM1	56	37 611	6 619	5 939	5,14%	4,75%	5,37%	10	10
4	Ecoles avec CP et CM1, CP et CM1 multigrades	30	6 583	1 185	1 238	0,90%	0,85%	1,12%	2	2
5	Ecoles avec CP et CM1, CP et CM1 double flux	68	81 586	14 690	12 429	11,15%	10,54%	11,24%	10	10
6	Ecoles avec CP et CM1, CM1 double flux, CP simple flux	11	7 104	1 203	1 154	0,97%	0,86%	1,04%	2	2
7	Ecoles avec CP et CM1, CP double flux, CM1 simple flux	51	48 511	8 703	7 057	6,63%	6,25%	6,38%	6	6
8	Ecoles avec CP et CM1, CP et CM1 simple flux, Dakar	268	174 561	30 191	26 583	23,87%	21,67%	24,05%	22	22
9	Ecoles avec CP et CM1 simple flux, Djourbel	37	21 038	3 973	3 169	2,88%	2,85%	2,87%	3	3
10	Ecoles avec CP et CM1 simple flux, Fatick	88	28 582	5 716	5 144	3,91%	4,10%	4,65%	4	4
11	Ecoles avec CP et CM1 simple flux, Kaolack	76	34 239	6 208	5 223	4,68%	4,46%	4,72%	4	4
12	Ecoles avec CP et CM1 simple flux, Kolda	76	24 263	4 895	4 234	3,32%	3,51%	3,83%	3	3
13	Ecoles avec CP et CM1 simple flux, Louga	47	20 357	3 765	3 227	2,78%	2,70%	2,92%	3	3
14	Ecoles avec CP et CM1 simple flux, St Louis	107	44 326	8 584	6 725	6,06%	6,16%	6,08%	5	5
15	Ecoles avec CP et CM1 simple flux, Tambacounda	39	15 720	3 215	2 295	2,15%	2,31%	2,08%	2	2
16	Ecoles avec CP et CM1 simple flux, Thiès	123	58 793	10 719	9 474	8,04%	7,69%	8,57%	7	7
17	Ecoles avec CP et CM1 simple flux, Ziguinchor	102	41 089	7 154	7 096	5,62%	5,13%	6,42%	5	5
18	Ecoles avec CP seulement, CP multigrade	11	997	323	0	0,14%	0,23%	0%	5	0
19	Ecoles avec CP simplement, CP simple flux	452	53 948	22 187	0	7,38%	15,92%	0%	6	0
Total		1935	731 407	139 330	110 550	100%	100%	100%	99	98

On peut faire quelques remarques, en comparant les colonnes d'effectifs de la population (intitulées «% d'élèves de CP» et «% d'élèves de CM1") avec les colonnes de nombre de classes retenues[1] (intitulées «Ecoles avec CP dans l'échantillon» et «Ecoles avec CM1 dans l'échantillon») :

- Les strates 1 à 7 et 18 à 19, qui correspondent à des types d'organisation spécifiques et relativement rares au Sénégal, ont été définies sur l'ensemble du territoire sénégalais. En effet, morceler ces strates davantage (par région) aurait conduit à définir des strates beaucoup trop petites, représentant chacune un trop faible pourcentage de la population.

- La strate 1 a été sur-représentée, afin de pouvoir mieux étudier les effets du multigrade au CM1 ;

- La strate 2 a été, elle, sous-représentée, sans doute parce qu'on n'attendait rien de spécifique de ces écoles qui ne comportent pas de CP mais sont à simple flux, et qu'il y avait un coût à envoyer un administrateur de tests pour un seul niveau ;

- Les écoles pilotes ont été sur-représentées, pour mieux étudier leurs effets mais aussi parce que toutes les classes d'une école pilote ne sont pas des classes pilotes.

- Les différentes régions ont été représentées à peu près proportionnellement.

3. Les propriétés d'un échantillon en grappes :

Pour un échantillon en grappes, le calcul de la précision des estimations se complique du fait de la complexité du plan expérimental. C'est pourquoi il n'est pas possible de donner ici des formules de calcul ; il s'agit plutôt de comprendre l'intuition (pourquoi l'existence de grappes fait-elle perdre en précision alors que la stratification permet au contraire d'en gagner ?), et d'apprendre à lire les tables qui indiquent la précision qu'on peut attendre en fonction des différentes caractéristiques de la population et de l'échantillonnage.

[1] On peut considérer qu'il s'agit approximativement de % d'élèves retenus dans l'échantillon, car le nombre total de classe est très proche de 100 et que, sauf classes de moins de 20 élèves, le même nombre d'élèves est retenu dans chaque classe.

<u>Propriété 1 :</u> l'échantillon en grappe permet de donner une estimation sans biais des caractéristiques de la population (moyenne, proportion) :

· Soit toutes les strates ont été représentées proportionnellement dans l'échantillon, si bien que tous les individus ont la même probabilité de figurer dans l'échantillon final, et alors une moyenne (ou proportion) sur l'ensemble de l'échantillon est une estimation sans biais de la moyenne (ou proportion) sur l'ensemble de la population.

· Soit certaines strates ont été sur-représentées, et alors il convient de pondérer les moyennes obtenues dans les différents sous-échantillons par le poids de la strate correspondante dans la population[1].

<u>Propriété 2 :</u> la précision d'une estimation par échantillon en grappes est moins grande que pour un échantillon aléatoire simple de même taille. La perte de précision («l'effet de plan expérimental»[2]) est d'autant grande que les groupes sont différenciés, c'est-à-dire que les individus se ressemblent entre eux au sein d'un groupe (dans le cas d'enquête scolaire, les groupes sont les classes) et diffèrent d'un groupe à l'autre. Cette ressemblance des individus au sein des groupes est mesurée à l'aide du **coefficient de corrélation intraclasse (noté ρ)**. ρ varie entre 0 et 1. Lorsque les groupes sont très hétérogènes entre eux, mais que les individus sont très homogènes à l'intérieur d'un groupe, ρ prend des valeurs proches de 1 ; il se rapproche de 0 lorsque les groupes se ressemblent et que la diversité est plutôt au sein de chaque groupe.

La formule pour estimer ρ est la suivante, et peut-être utilisée une fois qu'on a tiré un échantillon :

$$\hat{\rho} = \frac{n_g s_g^2 - s^2}{(n_g - 1) \times s^2}$$

où n_g est le nombre d'individus dans un sous-échantillon[3];

[1] Voir ci-dessus le calcul du score moyen lorsque les classes multigrades sont sur-représentées.
[2] Appelé «design effect» en anglais.
[3] Nous appelons échantillon l'ensemble des individus représentant la population, et sous-échantillon l'ensemble des individus représentant un groupe (ici, les groupes sont les classes, et un sous-échantillon est un ensemble de n_g élèves représentant la classe g). n_g vaut 20, par exemple, si on tire 20 élèves par classe.

s_g^2 est la variance des moyennes de groupe (on calcule les moyennes dans chaque groupe -- par exemple chaque classe -- et on en calcule la variance);

s^2 est la variance de l’échantillon.

Au total, la précision d’un échantillon en grappes (pour une population suffisamment grande) dépend de trois paramètres principaux :
1. Le coefficient de corrélation intraclasse, ρ ;
2. Le nombre de groupes (par exemple, de classes) tirés ;
3. Le nombre d’individus (par exemple, d’élèves) tirés dans chaque groupe.
Des tableaux sont utilisés pour présenter les différents échantillons qui permettent d’obtenir telle ou telle précision en fonction de ρ.

Illustration :

L’intuition de la perte de précision liée à l’échantillon par grappe peut être perçue avec l’image suivante :

Supposons que notre problème soit de décrire la composition d’une salade de fruits. On a le droit, pour cela, de manger un nombre donné de morceaux de fruits (50). Deux méthodes sont possibles : dans le cas de l’échantillon aléatoire simple, on pique avec les doigts, successivement, 50 morceaux de fruits un peu partout dans le saladier. Dans le cas d’un tirage en grappe, on utilise une cuillère pour prendre dix cuillerées de cinq morceaux de fruits chacune. La deuxième méthode, plus propre, est beaucoup moins précise que la première, pour peu que la salade de fruit n’ait pas été parfaitement mélangée : chaque cuillerées va comporter plus ou moins les mêmes fruits, et il sera difficile, en dix cuillerées, de rendre compte des différents fruits qui constituent notre salade.
Le rôle du coefficient de corrélation intra-classe est décisif : il correspond à la façon dont la salade de fruits a été mélangée. Il est faible lorsque la salade est bien mélangée, ce qui a pour conséquence que chaque cuillerée est assez variée, au point que l’ensemble des cuillerées parvient à rendre bien compte de la salade dans son ensemble. La perte de précision est alors minime. Mais plus il est élevé, moins la salade a été mélangée et plus les quelques cuillerées absorbées risquent de nous tromper sur la composition de la salade.

Utilisation des tables d’échantillonnage :

Le choix d’un échantillon (méthode d’échantillonnage et taille de l’échantillon) résulte toujours d’une optimisation sous contrainte : sachant les ressources dont on dispose, quel est le dispositif optimal pour estimer les grandeurs d’intérêt ou, symétriquement, sachant les exigences de précision imposées, quel est le dispositif qui me permet de minimiser le coût d’enquête ?

En particulier, dans le cas des échantillons en grappes (utilisés dans les enquêtes de rendement scolaire que nous menons), quel est le nombre de classes qu’il faut choisir ? Combien d’élèves faut-il retenir dans chaque classe ? Il est en effet possible de choisir soit beaucoup de classes avec un petit nombre d’élèves, soit beaucoup d’élèves dans chaque classe mais un petit nombre de classes.

Les tables d’échantillonnage permettent de faire ces choix, en proposant différentes configurations équivalentes : par exemple, examinons la première table donnée ci-dessous :

- La première ligne indique le degré de certitude exigé pour les intervalles de confiance : 95%. Autrement dit, cette table établit des fourchettes dans lesquelles on a 95% de chances que se trouvent les vraies valeurs.
- La deuxième ligne donne différentes tailles de fourchettes (c’est-à-dire d’intervalles de confiance). Les plus resserrées se trouvent dans les colonnes de gauche :
  1. «0,05 s» signifie qu’on exige que les moyennes (par exemple, le score moyen des élèves) soient estimés à une marge de 0,05 écart type près. Autrement dit, si l’écart type du score des élèves est de 20 points, le score moyen des élèves doit être estimé à ±1 point près.
  2. «2,5%» signifie que les proportions doivent être estimées à 2,5% près. Par exemple, si on évalue le pourcentage de filles, on exige que celui-ci soit connu à ±2,5% près.

  Selon la même logique, les colonnes de droite donnent des exigences moindres : 0,20 s signifie que, si l’écart type est de 20, le score moyen des élèves doit être estimé à 4 points près (en plus ou en moins).

- Le tableau est ensuite divisé en plusieurs tables selon la valeur du paramètre ρ. Il faut en effet connaître le coefficient de corrélation intraclasse

Tables d'échantillonnage :

Taille des échantillons

Limites de confiance : 95% pour les moyennes et pour les pourcentages	0,05 s 2,5%		0,10 s 5,0%		0,15 s 7,5%		0,20 s 10,0%	
Taille des sous-échantillons	Nombre de groupes	Nombre d'élèves	Nombre de groupes	Nombre d'élèves	Nombre de groupes	Nombre d'élèves	Nombre de groupes	Nombre d'élèves
ρ=0,1								
1	1600	1600	400	400	178	178	100	100
2	880	1760	220	440	98	196	55	110
5	448	2240	112	560	50	250	28	140
10	304	3040	76	760	34	340	19	190
15	256	3840	64	960	29	435	16	240
20	232	4640	58	1160	26	520	15	300
30	208	6240	52	1560	24	720	13	390
40	196	7840	49	1960	22	880	13	520
50	189	9450	48	2400	21	1050	12	600
ρ=0,2								
1	1600	1600	400	400	178	178	100	100
2	960	1920	240	480	107	214	60	120
5	576	2880	144	720	65	325	36	180
10	448	4480	112	1120	50	500	28	280
15	406	6090	102	1530	46	690	26	390
20	384	7680	96	1920	43	860	24	480
30	363	10890	91	2730	41	1230	23	690
40	352	14080	88	3520	40	1600	22	880
50	346	17300	87	4350	39	1950	22	1100
ρ=0,3								
1	1600	1600	400	400	178	178	100	100
2	1050	2100	260	520	116	232	65	130
5	704	3520	176	880	79	395	44	220
10	592	5920	148	1480	66	660	37	370
15	555	8325	139	2085	62	930	35	525
20	536	10720	134	2680	60	1200	34	680
30	518	15540	130	3900	58	1740	33	990
40	508	20320	127	5080	57	2280	32	1280
50	503	25150	126	6300	56	2800	32	1600

*Source : **Kenneth N. Ross, T Neville Postlethwaite.*** **Sample Design Procedures for the IEA International Study of Reading Literacy.** ***IEA, 1988.***

Tables d'échantillonnage (suite) :

Taille des échantillons

Limites de confiance : 95% pour les moyennes et pour les pourcentages	0,05 s 2,5%		0,10 s 5,0%		0,15 s 7,5%		0,20 s 10,0%	
Taille des sous-échantillons	**Nombre de groupes**	**Nombre d'élèves**	**Nombre de groupes**	**Nombre d'élèves**	**Nombre de groupes**	**Nombre d'élèves**	**Nombre de groupes**	**Nombre d'élèves**
				ρ=0,4				
1	1600	1600	400	400	178	178	100	100
2	1120	2240	280	560	125	250	70	140
5	832	4160	208	1040	93	465	52	260
10	726	7260	184	1840	82	820	46	460
15	704	10560	176	2640	79	1185	44	660
20	688	13760	172	3440	77	1540	43	860
30	672	20160	168	5040	75	2250	42	1260
40	664	26560	166	6640	74	2960	42	1680
50	660	33000	165	8250	74	3700	42	2100
				ρ=0,5				
1	1600	1600	400	400	178	178	100	100
2	1200	2400	300	600	134	268	75	150
5	960	4800	240	1200	107	535	60	300
10	880	8800	220	2200	98	980	55	550
15	854	12810	214	3210	95	1425	54	810
20	840	16800	210	4200	94	1880	53	1060
30	827	24810	207	6210	92	2760	52	1560
40	820	32800	205	8200	92	3680	52	2080
50	816	40800	204	10200	91	4550	51	2550
				ρ=0,6				
1	1600	1600	400	400	178	178	100	100
2	1280	2560	320	640	143	286	80	160
5	1088	5440	272	1360	122	610	68	340
10	1024	10240	256	2560	114	1140	64	640
15	1003	15045	251	3765	112	1680	63	945
20	992	19840	248	4960	111	2220	62	1240
30	982	29460	246	7380	110	3300	62	1860
40	976	39040	244	9760	109	4360	61	2440
50	973	48650	244	12200	109	5450	61	3050

Source : ***Kenneth N. Ross, T Neville Postlethwaite.*** **Sample Design Procedures for the IEA International Study of Reading Literacy.** ***IEA, 1988.***

pour estimer l'effet du plan expérimental. Ce coefficient peut être connu à la suite d'autres études, ou estimé à partir d'études dans des pays comparables. A défaut, on peut prendre la valeur de 0,3. Un calcul ex post permettra d'ajuster l'estimation (et de réévaluer la précision effective de l'échantillonnage).
Restons pour notre part dans la première table (ρ=0.1), qui correspond à des classes peu différenciées.

➢ Le tableau se lit ensuite par colonnes. On requiert souvent que les scores des élèves soient connus à 0,10 écart type près avec une probabilité de 95%. C'est donc la deuxième configuration (en partant de la gauche) qui nous intéresse, soit les colonnes 4 et 5 du tableau. Elles se lisent de la façon suivante : pour obtenir des fourchettes de ±0,10 écart type avec une probabilité de 95%, alors que le coefficient intraclasse est de 0,10 (classes peu différenciées), on a le choix entre :
- 400 groupes de 1 individu chacun, c'est-à-dire un échantillon aléatoire simple de 400 élèves ;
- 220 groupes de 2 individus chacun, c'est-à-dire un échantillon en grappes de 220 classes avec deux élèves testés par classe ;
- 112 groupes de 5 individus chacun, c'est-à-dire un échantillon en grappes de 112 classes où on teste à chaque fois 20 élèves ;

(…)

- 48 groupes de 50 individus chacun, soit 48 classes dans lesquelles on testerait à chaque fois 50 élèves.

La table nous indique donc que tous ces échantillons sont aussi précis les uns que les autres. Si un autre degré de précision est souhaité, il est possible de se référer à d'autres colonnes.

Comment choisir entre différents échantillons de précision équivalente ?

- Si le seul objectif de l'étude est l'évaluation d'un niveau moyen, ces différents échantillons sont strictement équivalents. Il ne reste qu'à choisir le moins coûteux. C'est vraisemblablement celui qui comprend le moins de classe, pour peu que le déplacement des enquêteurs coûte plus cher que l'impression, la correction et la saisie des instruments.
- Mais souvent, l'étude a d'autres objectifs, en particulier celui d'étudier les effets de différentes configurations scolaires sur les résultats des élèves. Dans ce cas, on souhaite avoir un nombre de classes différentes élevé,

de façon à pouvoir mener des comparaisons entre classes. Le mieux est alors de choisir un maximum de classes possibles. Pourtant, des considérations de coût conduisent sans doute à ne pas retomber dans l'échantillon aléatoire simple (un élève par classe), dans la mesure où, une fois que l'enquêteur s'est déplacé, interroger plusieurs élèves vaut largement la peine.

Le nombre d'élèves testés dans chaque classe s'étage entre 5 et 20 selon les études. Cela dépend vraiment de la priorité qui est donnée à l'évaluation du niveau moyen ou à la comparaison des différents types de classes, ainsi que de la structure des coûts de l'enquête.

Attention! Quelle que soit le nombre d'élèves par classe retenu, il ne s'agit surtout pas d'éliminer de l'échantillon les classes de moins de 20 élèves. Cela aurait pour conséquence de biaiser l'échantillon en omettant toute une catégorie de petites classes (par exemple, des classes multigrades). Au contraire, quand une telle classe est tirée, on la maintient dans l'échantillon. Simplement, on testera tous ses élèves.

En principe, le lecteur dispose à ce stade de tous les éléments nécessaires pour choisir son échantillon. En raison des nécessités de l'exposition, ces éléments sont néanmoins éparpillés au fil du chapitre. La section suivante va préciser la procédure à suivre en détaillant les différentes étapes.

E. Récapitulatif des étapes à suivre pour un échantillonnage en grappes[1]

1. **Définir les contraintes** qui vont nous guider : il s'agit des limites administratives et financières, ainsi que des exigences scientifiques de précision, qui vont limiter l'ensemble des échantillonnages faisables et acceptables. C'est dans ce sous-ensemble qu'on choisira le meilleur d'entre eux.

[1] Ce récapitulatif est largement inspiré d'un module de formation intitulé *Les échantillons* d'Emilie Barrier (CIEP, 1995) et d'un article de Kenneth N. Ross : «Comment établir le plan d'échantillonnage dans les études internationales sur le rendement scolaire ?» in *Perspectives*, vol XXII, n°3, 1992 (83), pp. 343 à 355.

2. **Définir la population d'étude** : parfois, on exclut une partie des individus qui figurent pourtant dans la population cible «désirée» (par exemple, parce que les listes sont incomplètes, ou parce que telle région est inaccessible pour cause de troubles,…). C'est à éviter au maximum, mais lorsque c'est inévitable, il faut le préciser clairement pour définir les limites de validité de l'échantillon. Cela se traduit par la construction de la base de sondage (par exemple, les listes d'écoles et de classes avec leurs effectifs).

3. Au moyen des tables d'échantillonnage, **déterminer le nombre d'écoles et d'élèves adéquat** par rapport aux moyens et aux objectifs de l'étude. Un des paramètres déterminants est le coefficient de corrélation intra-classe qui, s'il n'est pas connu par d'autres études dans le pays ou un pays voisin, peut être posé égal à 0,3 (on le calculera plus précisément une fois les données collectées).

4. **Définir les différentes strates** et choisir leur représentation (proportionnelle ou non). Le choix des strates et de l'importance de leur représentation dépend des priorités de l'étude[1].

5. **Tirer au sort les écoles** avec une probabilité proportionnelle à leur taille.

   Justification : on veut que le sous-échantillon correspondant à la strate indexée k soit représentatif de cette strate. Si on tire les écoles avec la même probabilité, puis qu'on tire un nombre fixe d'élèves dans chaque école, les élèves des écoles peu nombreuses auront plus de chances de figurer dans l'échantillon.

   Par exemple, supposons qu'on tire dix écoles parmi 100, puis 20 élèves de CM1 dans chaque école. Un élève qui se situe dans une école A avec 20 élèves en CM1 (ou moins) a une chance sur 10 d'être dans une école du sous-échantillon, auquel cas il est automatiquement présent dans le sous-échantillon final. Avec un tirage équiprobable des écoles, un élève d'une classe de 20 élèves a donc 1 chance sur 10 d'être dans le sous-échantillon final. Mais un élève qui se situe dans une école B avec 100 élèves de CM1 a une chance sur 10 d'être dans une école retenue, puis seulement 1 chance

[1] Voir les discussions de la section sur la stratification.

sur 5 (soit 20/100) d'être parmi les élèves retenus de cette école. Il a donc au total 1 chance sur 50 d'être dans le sous-échantillon final. La conséquence, c'est qu'un tirage équiprobable des écoles va sur-représenter les élèves des écoles peu nombreuses et sous-représenter les élèves des écoles nombreuses.

Pour résoudre cette difficulté, on va tirer les écoles plus nombreuses avec une plus forte probabilité. On veut 200 élèves dans le sous-échantillon. Supposons qu'il y en ait en tout 10 000 dans la strate. Pour que le sous-échantillon soit représentatif, il faut que chaque élève ait 1 chance sur 50 d'être retenu. Pour tenir compte du fait que l'école A a 20 élèves et l'école B 100, on va s'arranger pour que l'école A ait 1 chance sur 50 d'être retenue. Ses élèves ont donc 1 chance sur 50 d'être retenus. On va donner par contre 1 chance sur 10 à l'école B d'être retenue. Ses élèves auront alors 1 chance sur 50 d'être retenus (1 chance sur 10 que leur école soit retenue, puis une chance sur 5 d'être parmi les élèves retenus dans cette école). On constate que pour assurer une représentation équiprobable des élèves de la strate, on a tiré l'école B avec une probabilité cinq fois plus élevée que l'école A, ce qui correspond au rapport de leurs nombres d'élèves (100/20).

<u>Procédure recommandée :</u> La **«méthode de la loterie»** est souvent utilisée. On fait une tombola dans laquelle on attribue autant de billets gagnants à chaque école qu'il y a d'élèves dans cette école. Ainsi, par exemple, l'école A aura 20 tickets (les numéros 1 à 20) et l'école B 100 (les numéros 21 à 120). Il y a en tout autant de billets que d'élèves (10 000). Il faut tirer au sort 10 école parmi 100. On tire pour cela 10 billets sur 10 000, soit 1 billet tous les 1 000 (ce nombre 1 000 est appelé «l'intervalle de tirage»). Il ne reste plus qu'à tirer un numéro de départ parmi les 1000 premiers. Par exemple, si on tire le numéro 87, on retient les billets 87, 1087, 2087, 3087,…, 9087. Les écoles qui ont ces billets sont retenues. Par exemple, l'école B, qui a le billet 87, fait partie de l'échantillon.

6. **Tirer un échantillon de remplacement** :

Il se peut que, pour une raison ou une autre (maladie des responsables de l'école, changement de certaines caractéristiques de l'école dû à l'utilisation de listes trop anciennes…), certaines écoles ne puissent participer à l'étude. Ce n'est jamais bon pour la fiabilité de l'échantillon. La première règle est donc de faire le possible pour aller dans toutes les écoles tirées (rappelons qu'une école difficile d'accès présente peut-être des caracté-

ristiques intéressantes qu'il ne faut pas occulter dans l'échantillon). Mais il est prudent d'avoir prévu une règle pour remplacer les écoles vraiment inutilisables.
Une façon de faire est de tirer un deuxième échantillon d'écoles en suivant exactement les mêmes règles de stratification et de tirage que dans le premier. Au sein de chaque strate, on attribue à chaque école du premier échantillon un remplaçant dans le second échantillon. Il est important que cela soit fait à l'avance, de sorte que des solutions de dernière minute improvisées sur le terrain ne viennent pas fausser l'échantillonnage.

7. **Vérifier la qualité de l'échantillon** avec des variables de contrôle :
On s'est certes entouré de nombreuses précautions qui nous donnent un maximum de chances d'avoir un échantillon représentatif de la population. Mais il existe toujours un risque de tomber sur un cas particulier. La stratification visait justement à éliminer certains cas particuliers où des strates sont complètement absentes. Mais au sein de chaque strate, les tirages au sort peuvent avoir conduit à un sous-échantillon très spécifique. On peut essayer de déceler de tels cas en comparant, pour les informations disponibles, la composition du sous-échantillon et celle de la strate : par exemple, taux de réussite des écoles aux examens, pourcentage de filles, effectif des écoles,… Si jamais les discordances sont vraiment très grandes, il est possible de tirer un nouvel échantillon. Mais ce devrait être rarissime!

En conclusion, l'échantillonnage constitue une technique que peut s'approprier toute équipe d'évaluation. Les connaissances techniques nécessaires sont relativement modérées ; il nous a semblé plus opportun d'insister sur l'intuition des conséquences des différents choix de stratification et de types de tirages, car c'est cette compréhension qui permettra de déjouer au mieux les difficultés du terrain en distinguant l'essentiel de l'accessoire.