

# CHAPITRE 4 : L'AGRÉGATION DES DONNÉES ÉLÉMENTAIRES ET LE CALCUL DES QUOTIENTS DE RÉFÉRENCE

## 4.1 L'agrégation des données élémentaires

Une fois le zonage constitué, l'étape suivante consiste à rassembler l'information de base nécessaire à la projection. Il convient à ce titre de sélectionner son année de départ, le recensement annuel permettant de faire démarrer une projection de n'importe quelle année.

L'**agrégation** consiste à sommer les données de base exprimées au niveau communal pour obtenir des données se rapportant au zonage considéré. Trois processus d'agrégation vont être réalisés :

- ❖ l'agrégation des données du recensement fournissant une population de départ détaillée par sexe, âge, zone de résidence actuelle et commune de résidence antérieure. Cette première agrégation permettra de constituer les pyramides des âges de départ ainsi que les flux moyens entre les différentes zones ;

$$pop_{z,s,i} = \sum_{c \in z} pop_{c,s,i,r}$$

avec  $z$  désignant la zone considérée,  $c$  désignant ses communes constitutives,  $s$  désignant le sexe,  $i$  désignant l'âge et  $r$  désignant la commune de résidence antérieure

- ❖ l'agrégation des décès de l'état civil cumulés sur le même cycle quinquennal que celui du recensement et détaillés par sexe, âge et zone de décès, puis ramenés ensuite à une moyenne annuelle. Cette phase permettra ensuite de calculer les quotients de décès quinquennaux, par chaînage des quotients de décès annuels sur cinq âges successifs ;

$$décès_{z,s,i} = \frac{1}{5} \sum_{a=2005}^{2009} \sum_{c \in z} décès_{c,s,i,a}$$

$a$  désigne ici l'année

- ❖ l'agrégation des naissances de l'état civil sur le même cycle quinquennal que celui du recensement et détaillées par âge de la mère et zone de naissance, puis ramenées à une moyenne annuelle. Cette étape permettra de simuler des naissances par pas quinquennal, par chaînage sur cinq âges successifs.

$$naissances_{z,s,i} = \frac{1}{5} \sum_{a=2005}^{2009} \sum_{c \in z} naissances_{c,s,i,a}$$

$i$  désignant ici l'âge de la mère et  $a$  l'année

## 4.2 La notion d'âge dans les pyramides des âges

À la différence des précédents recensements et des données de l'état civil, où les concepts d'âge et de génération se confondent, les effectifs par âge figurant dans les pyramides des âges découlent du cumul de cinq années d'un cycle de recensement et prennent en compte des individus issus de cinq générations différentes. Par ailleurs, la

méthode d'agrégation retenue ignore les éventuels déséquilibres de la zone en termes d'années de collecte. L'importance des différences de représentativité entre années est néanmoins à relativiser :

- ❖ des travaux méthodologiques menés sur l'utilisation des données du recensement ont démontré que, compte tenu de la constitution des groupes de rotation des communes recensées, les niveaux supra-communaux bénéficient d'un bon équilibre entre les différentes années de collecte. D'autre part, étant donnée l'inertie caractérisant les phénomènes démographiques, ces questions d'équilibre restent du second ordre par rapport à la taille des populations ;
- ❖ exceptés les éventuels gradients d'effectifs par génération, le raisonnement en « moyenne par âge » permet de bien prendre en compte que, sur une période de cinq ans, compte tenu de l'inertie des phénomènes migratoires, les comportements démographiques (donner naissance, décéder, migrer) dépendent davantage de l'âge que de la génération.

### 4.3 Le calcul des quotients de référence

La réalisation d'une projection suppose la spécification d'un certain nombre de paramètres (horizon de la projection, scénario démographique avec hypothèses sur l'évolution de la fécondité, de la mortalité et des flux migratoires), qui seront détaillés dans les chapitres suivants. Une fois ces paramètres déterminés, le processus de projection manipule des données issues de l'agrégation, de la modélisation des échanges avec l'étranger et de l'application d'évolutions aux différents quotients de référence (projection des quotients). Plus précisément, il mobilise :

- ❖ la population pour l'année de référence par zone, sexe et âge (issue de l'étape d'agrégation) ;
- ❖ les quotients de fécondité par zone et âge de la mère ;
- ❖ des ratios naissances-enfants pour l'année de référence (cette entité ne peut être sujette à des évolutions et est maintenue constante sur la période de projection) par zone et âge de l'enfant ;
- ❖ les quotients de décès par zone, sexe et âge ;
- ❖ les quotients d'émigration.

Les équations sont élaborées sur une période quinquennale. L'algorithme de projection projette ainsi, de cinq ans en cinq ans, une pyramide initiale issue de l'agrégation en mobilisant les quotients projetés. Il comprend ensuite une procédure d'annualisation (voir chapitre 6) qui, à partir des pyramides quinquennales projetées, reconstitue celles des années manquantes.

#### 4.3.1 Les données de base

À partir des données issues de la phase d'agrégation, celles portant sur les flux peuvent être élaborées. On y distingue, d'une part, les flux d'émigration et d'immigration et, d'autre part, les différentes zones d'échange.

Le solde des flux qui en découle consiste, pour une zone donnée, à soustraire les flux d'émigration des flux d'immigration. Les flux par zone d'origine, zone de destination, sexe et âge sont également issus de la première étape d'agrégation.

Ces flux sont observés sur les cinq années d'enquêtes, allant de 2005 à 2009, qui constituent le RP 2007. Ils sont représentatifs d'une fenêtre d'observation glissante allant de 2000-2005 (mouvements migratoires mesurés par l'enquête annuelle de 2005) à 2004-2009 (mouvements migratoires mesurés par l'enquête annuelle de 2009).

Contrairement à l'ancienne application, ces flux sur cinq ans ne feront pas l'objet d'une procédure d'annualisation systématique. Une telle optique implique de se limiter à des projections avec un pas de cinq ans (2007, 2012, 2017, etc.) fondées sur les données précédentes.

Cette démarche permet d'éviter les problèmes inhérents à l'annualisation et notamment le positionnement au bon âge des phénomènes migratoires.

Néanmoins, les besoins des utilisateurs, pour certaines problématiques spécifiques associées à des échéances précises, peuvent porter sur des pyramides aux années intermédiaires. Dans un tel cadre, l'application Omphale 2010 permet de calculer les pyramides annuelles à partir des pyramides projetées avec un pas quinquennal (voir partie 3.4).

### 4.3.2 Les quotients de référence

#### 4.3.2.1 Les quotients de mortalité et de fécondité

Comme l'indique le mécanisme général de projection (cf. paragraphe 2.3), ces quotients entreront en ligne de compte pour respectivement ajouter les naissances à la pyramide de l'année  $A + 5$  projetée à partir de celle de l'année  $A$ , et retirer ex-ante les décès de cette dernière pyramide pour ne conserver cinq ans plus tard que les personnes vivantes.

Les quotients de fécondité et de mortalité sont calculés pour l'ensemble des années allant de l'année de début de projection à l'année de fin.

Pour les quotients de fécondité, décès, actifs, ménages, les calculs s'opèrent au niveau de la zone, contrairement aux quotients d'émigration. Pour ces derniers, les calculs sont en effet réalisés pour chaque couple zone d'origine et zone de destination, avec la zone d'origine différente de la zone de destination.

##### 4.3.2.1.1 Les quotients de mortalité

Le quotient de mortalité à un âge donné mesure **la probabilité, pour un individu ayant atteint cet âge, de décéder avant d'atteindre l'âge suivant**. Étant donné la nature du nouveau recensement, le quotient de mortalité, détaillé par sexe et âge, sera calculé en rapportant les décès moyens observés au cours du cycle de référence à la population sur ce cycle.

$$Qd_{z,s,i} = \frac{\text{décès}_{z,s,i}}{\text{pop}_{z,s,i}}$$

La variable **DÉCÈS** désigne les décès annuels moyens selon la zone de résidence et **POP** la population de cette zone. Les quotients de mortalité s'interprètent comme des quotients annuels représentatifs d'une situation moyenne sur les cinq années d'un cycle EAR.

À la différence de l'ancien modèle, ils prennent en compte les spécificités locales, n'étant pas calqués sur le profil par âge métropolitain. Ces profils sont plus accidentés qu'avec l'ancien modèle. Pour des zones à faibles effectifs, ce risque sera toutefois réduit par l'empilement sur cinq années des données de l'état civil, ce qui a pour effet de se prémunir contre une surmortalité ou sous-mortalité conjoncturelle. Par exemple, pour une zone de 50 000 habitants (seuil minimal), le nombre d'événements de l'état civil utilisés pour le calcul des quotients sera équivalent à celui d'une zone de 250 000 habitants avec une seule année valorisée. À une telle échelle, les profils apparaissent nécessairement moins accidentés.

**Au-delà de 80 ans** cependant, les quotients de décès ne sont plus directement issus de l'observation du rapport entre les décès de l'état civil et la population de référence. En effet, pour certains de ces âges élevés, les effectifs soumis au risque de décès peuvent être très faibles (voire nuls), ce qui implique, en l'absence de correction, des quotients fortement erratiques. Il convient donc au préalable de vérifier si les populations ou les décès ne sont pas nuls. Si c'est le cas, le quotient de décès de base de la zone en question sera calqué sur le quotient métropolitain.

$$\text{Si } \text{POP}_{z,s,i} = 0 \text{ ou } \text{DÉCÈS}_{z,s,i} = 0 \text{ alors } Qd_{\text{base}_{z,s,i}} = Qd_{\text{frances},i}$$

Sinon, on calcule un quotient « de base » qui se confond avec les quotients de décès des âges inférieurs à 80 ans.

$$Qd\_base_{z,s,i} = \frac{décès_{z,s,i}}{pop_{z,s,i}}$$

Ensuite, ces quotients de base subiront des transformations afin de s'extraire de leur caractère erratique lié aux faibles populations en jeu. Ainsi, pour l'ensemble des quotients de décès relatifs aux âges supérieurs ou égaux à 80 ans, une transformation logarithmique sur ces quotients de base est opérée :

$$Qd\_lth_{z,s,i} = \ln(qd\_base_{z,s,i})$$

On réalise ensuite une régression linéaire du vecteur  $Qd\_lth$  sur les âges. Pour chaque zone et chaque sexe, en notant  $AGES$  le vecteur des âges allant de 80 à 99 ans et  $Qd\_lth_{z,s,i}$  le vecteur constitué des  $qd\_lth$  pour les âges en question :

$$Qd\_lth\_reg_{z,s,i} = regression(Qd\_lth_{z,s,i}, AGES)$$

L'obtention des quotients de décès pour les âges supérieurs ou égaux à 80 ans se fait alors par le passage à l'exponentielle :

$$Qd_{z,s,i} = e^{Qd\_lth\_reg}$$

**Au-delà de 99 ans**, les quotients de décès sont supposés égaux aux quotients d'âge 99 ans (qui intègre lui-même de l'information sur les plus de 99 ans, les effectifs du recensement et les décès de l'état civil pour les âges supérieurs à 99 ans étant agrégés avec ceux de l'âge 99 ans).

Enfin, ces quotients de mortalité annuels seront appliqués à une pyramide initiale afin d'estimer une sous-population de survivants cinq ans plus tard. À cette fin, il convient de déterminer des **quotients de survie sur cinq ans** ne dépendant que de la zone de résidence initiale. Ils s'obtiennent en chainant les quotients de survie annuels correspondants.

## Remarques

Les quotients de décès rapportent directement un cumul de décès annuel à une population observée en début de cette même année. Cette optique pourrait traduire un biais dans la mesure où une partie des décès est imputable aux migrants en cours d'année, ou inversement certains décès imputables à cette population ne sont pas comptabilisés, et que l'on préfère ramener généralement ces décès à une population en milieu d'année. Toutefois, il convient de remarquer que :

- d'une part, les âges de forte mortalité sont également des âges de faibles migrations ;
- d'autre part, le fait que les quotients de décès soient calculés sur une moyenne quinquennale (des décès et des populations) permet de prendre en compte en grande partie l'impact des migrations. Le dénominateur, moyenne de cinq populations en début d'année, peut facilement s'écrire en fonction de populations annuelles moyennes et d'une population moyenne entre les deux années extrêmes.

Par ailleurs, les quotients de survie « cinq ans plus tard » sont appliqués à la pyramide initiale, ce qui implique que les éventuels migrants entre ces deux dates se voient appliquer les comportements de mortalité de leur zone d'origine.

### 4.3.2.1.2 Les quotients de fécondité

Dans Omphale 2010, comme pour la mortalité, les quotients de fécondité sont également spécifiques à la zone étudiée, ce qui constitue l'un des apports importants du nouveau modèle. Les naissances par âge et zone de résidence de la mère à la date de naissance sont directement issues de la première étape d'agrégation sur les données de l'état civil.

$$Qf_{z,i} = \frac{\sum_{\text{sexe} \in \{\text{garçons}, \text{filles}\}} \text{naissances}_{z,s,i}}{\text{pop}_{z,\text{femmes},i}}$$

La variable **NAISSANCES** désigne les naissances moyennes selon l'âge de la mère et la zone de résidence de la mère à la naissance et **POP** la population légale de la zone.

Les quotients de fécondité s'interprètent comme des quotients annuels représentatifs d'une situation moyenne sur les cinq années d'un cycle EAR.

La répartition par sexe de l'enfant se fera a posteriori, au cours du processus de projection, selon un coefficient métropolitain.

### Remarques

Ces quotients intervenant en année N + 5, des évolutions souhaitées sur la fécondité lors de la définition de scénarii démographiques (cf. chapitre 5) seront répercutées tous les cinq ans contre annuellement pour les composantes mortalité et migrations.

Comme pour la mortalité, la population de femmes au dénominateur n'est pas corrigée d'un demi-solde migratoire annuel qui permettrait d'obtenir une population en milieu d'année et donc de nature à prendre en compte l'impact des migrations sur le cumul annuel des naissances de l'état civil. Toutefois, même si les âges de forte fécondité sont également des âges de fortes migrations, le raisonnement en moyenne quinquennale des naissances et de la population des femmes permet d'atténuer ce biais. Enfin, à travers ce ratio, **on ne cherche pas à exprimer un indicateur conjoncturel de fécondité**, mais bien à calculer un nombre de naissances dans cinq ans, à partir d'une population féminine. Il y a donc une légère différence entre cet indicateur retenu et celui qui conduirait, par sommation sur les âges, à l'indicateur conjoncturel de fécondité stricto-sensu.

Enfin, ce mode opératoire contient l'hypothèse implicite que les comportements de fécondité des femmes migrantes s'adaptent à ceux de la zone qu'elles rejoignent et non de la zone qu'elles quittent. Cette hypothèse se rapproche davantage de la réalité démographique que ne l'aurait fait l'hypothèse inverse. Dans les zones universitaires, les étudiantes ont peu d'enfants. À âge égal, la fécondité y est plus faible qu'ailleurs. Une femme rejoignant une zone universitaire aura ainsi une probabilité moindre d'avoir des enfants. À l'inverse, une femme quittant une zone universitaire pour un premier emploi aura une probabilité plus forte d'avoir des enfants, ce qui n'aurait pas été pris en compte si on lui avait affecté les comportements de fécondité des zones universitaires.

#### 4.3.2.1.3 La population des enfants de 0 à 4 ans

Le processus de projection étant défini sur une base quinquennale, les pyramides des âges ne sont calculées, avant annualisation, que pour les années charnières. Le quotient de fécondité projeté en N + 5 à partir du quotient de fécondité de l'année N permet de calculer les naissances relatives à N + 5, ce qui correspond à la population d'âge - 1 en N + 5. Il reste donc à déterminer les populations des enfants âgés de 0 à 4 ans pour compléter cette pyramide.

À cette fin, des « **ratios naissances-enfants** » vont être calculés à partir de l'année de référence. Ces ratios permettent donc, à partir du nombre de naissances d'une année donnée sur un territoire donné, d'en déduire le nombre d'enfants d'âge 0, 1, 2, 3 et 4 ans présents sur ce territoire cette même année. Les différences d'effectifs entre ces naissances et les âges de 0 à 4 ans reflètent à la fois une légère mortalité à ces âges, les comportements migratoires des jeunes enfants, ainsi que des différences d'effectifs par génération.

À partir des données de l'année de référence, les ratios naissances-enfants correspondent donc pour chaque zone et pour chaque âge de 0 à 4 ans à :

$$ratio\_nais\_enf_{z,i} = \frac{\sum_s pop_{z,s,i}}{\sum_{s,i\_mere} naissances_{z,s,i\_mere}}$$

Au même titre que les naissances, les enfants de 0 à 4 ans sont « réinjectés » en fin de processus pour chaque pas quinquennal de projection. Pour l'année de fin de période N + 5, pour chaque zone et chaque sexe :

$$Naissances_s = part_s * \sum_{i\_mere} Pop\_proj_{z,femmes,i} * Qf_{N+5,z,i\_mere}$$

$$Enfants_{s,i} = Naissances_s * ratio\_nais\_enf_{z,i}$$

### Remarque

Les naissances étant ajoutées en fin de processus, une fois tous les cinq ans, d'éventuelles modifications sur les comportements de fécondité ne peuvent être pris en compte qu'à cette fréquence (pas de cinq ans) et non de façon progressive annuellement.

Toute modification des comportements de fécondité se répercute instantanément sur la population de 0 à 4, car ses effectifs dépendent directement des naissances, et non de ceux des enfants de 5 ans.

#### 4.3.2.1.4 Les quotients d'émigration bi-localisés

Ces quotients interviennent en milieu de processus, une fois les décès et émigrations vers l'étranger ôtés, et avant l'injection des immigrants et des enfants de moins de 5 ans. Ces flux s'observent sur une population résidant en France et ayant mentionné une zone de résidence antérieure en France, et donc implicitement vivante aux deux périodes. Il est donc logique d'appliquer ces probabilités de migrer à ce même champ.

Les quotients d'émigration bi-localisés, détaillés par sexe et par âge, peuvent s'interpréter comme la probabilité pour un individu d'être sur une zone d'origine ZO, de migrer et de se retrouver cinq ans plus tard sur une zone de destination ZD :

$$Qe_{zo,zd,s,i-5} = \frac{flux_{zo,zd,s,i}}{stables_{zo,s,i} + flux_{zo,.,s,i}}$$

La variable **FLUX ZO.** traduit les flux de la zone d'origine vers l'ensemble des autres zones et **FLUX. ZO** représente les flux de toutes ces autres zones vers la zone d'origine. Le dénominateur peut s'interpréter comme une population de survivants présente cinq ans auparavant sur la zone d'origine. De la même manière que pour les quotients de décès, ces quotients d'émigration ne feront l'objet d'aucune procédure de lissage.

Le quotient d'émigration est un quotient quinquennal (relatif à la question sur les migrations résidentielles « *Où habitiez-vous cinq ans auparavant ?* ») représentatif d'une situation moyenne sur les cinq années du cycle de recensement. S'agissant du recensement 2007 par exemple, l'information sur les flux peut provenir de l'enquête de recensement 2005 et être relative à une migration intervenue en 2000. Elle peut également provenir de l'enquête de recensement 2009, relativement à une migration ayant eu lieu très récemment, en 2008 voire dans les tout premiers jours de 2009. C'est pourquoi les quotients d'émigration associés au recensement 2007 traduisent en fait des comportements migratoires moyens de la période 2000-2008, avec surreprésentation des années médianes (les migrations survenues aux années médianes ont plus de chance d'être observées que celles intervenues aux bornes de l'intervalle).

Par ailleurs, il convient de prendre garde au décalage d'âge inhérent à la périodicité retenue : un quotient d'émigration calculé à partir des données moyennes labélisées année N est relatif à un comportement migratoire sur la période  $[N - 5, N]$ . Ainsi, le quotient d'émigration estimé à partir des données concernant les hommes ayant 30 ans l'année N correspond à un quotient d'émigration relatif aux individus âgés de 25 ans voici cinq ans.

Le modèle est mis en œuvre sur l'ensemble du territoire (le zonage constituera nécessairement une partition de France métropolitaine + DOM) et présente un triple avantage :

- ❖ garantir la nullité de la somme des flux internes à la France ;
- ❖ tenir compte de la démographie des principales zones d'échange du territoire étudié ;
- ❖ apporter, aux utilisateurs des projections, une information nouvelle et pertinente en phase de projection sur ces flux d'échange : intensité et orientation (immigration et émigration) des flux, ventilation par sexe et par classes d'âge, et évolution de ces flux dans le temps.

Ensuite, le recours à des flux sur cinq ans permet d'annihiler une difficulté inhérente au positionnement à l'âge exact des migrations, caractéristique indispensable pour réaliser des projections avec un pas annuel. L'âge auquel s'est effectuée la migration est en effet inobservable. Dans l'ancien modèle, cette question était résolue en considérant que les individus étaient soumis à X quotients de migration successifs (X correspondant à la durée de la période intercensitaire). Dans le nouveau modèle, l'âge auquel s'est effectuée la migration importe peu : les seules informations utiles sont l'âge en début de période (a fortiori l'âge en fin de période) et l'existence d'un éventuel flux. En d'autres termes, pour les personnes de 18 ans en début de période dans la zone Z, on projette la probabilité d'être là où sur une autre zone à 23 ans, sans se soucier de l'âge précis auquel sont survenues les éventuelles migrations.

Il convient de garder à l'esprit que ces flux quinquennaux ne sauraient s'annualiser simplement par une division par cinq. Ils sont en fait une résultante de flux intervenus au sein de la période, certains correspondant à des migrations multiples (les zones intermédiaires sont alors inobservables), d'autres à des allers-retours (inobservables également). Toutefois, l'existence de migrations multiples et des phénomènes d'allers-retours ne sont pas de nature à remettre en cause la portée explicative du modèle de flux. En effet, les flux nous intéressent parce qu'ils contribuent à l'évolution de la population de la zone, **l'objectif premier n'étant pas d'estimer les migrations**. Ainsi, considérer des migrants même apparents (au sens où l'on ne tient compte que de la résultante des trajectoires entre deux dates espacées de cinq ans, et non des parcours intermédiaires) suffit, car c'est bien le solde migratoire (la différence entre émigrations de ZO vers ZD et émigrations de ZD vers ZO) qui, in fine, fera évoluer la pyramide des âges, et que dans ce solde, la question des migrations multiples n'entre plus en ligne de compte.

#### **4.3.2.1.5 Les échanges migratoires avec l'étranger**

Il s'agit de déterminer, pour chaque zone considérée, des immigrants en provenance de l'étranger et des émigrants vers l'étranger par sexe et âge détaillé, de manière à intégrer l'impact des échanges migratoires avec l'étranger sur la dynamique de population de chacune de ces zones. La question relative aux migrations résidentielles du recensement permet de connaître, à tout échelon géographique supra-communal, les immigrants en provenance de l'étranger. Par contre, aucune information du recensement, et plus généralement aucune source mobilisable, permettant d'appréhender de manière systématique les émigrants vers l'étranger n'a pu être identifiée.

Par ailleurs, les projections métropolitaines réalisées par le département de la démographie de l'Insee introduisent des hypothèses sur le solde métropolitain avec l'extérieur.

L'exigence de cohérence avec ces projections métropolitaines impose que, localement, les échanges avec l'étranger puissent être vus comme une ventilation des effectifs de l'hypothèse métropolitaine.

Enfin, le processus de projection, de par sa construction, conduit à privilégier une approche permettant de distinguer immigrants de l'étranger et émigrants vers l'étranger aux méthodes qui modéliseraient uniquement leur résultante sous forme de solde : **les émigrants sont en effet retranchés en début de projection et les immigrants imputés en fin de projection**.

L'identification des échanges avec l'étranger au niveau local repose donc sur la double contrainte :

- ❖ valorisation de l'information locale sur les immigrants de l'étranger issue du recensement ;
- ❖ déclinaison locale du solde métropolitain avec l'étranger issu du Bilan démographique.

La première étape consiste à calculer le nombre d'émigrants vers l'étranger au niveau métropolitain. Ces émigrants sont calculés en retranchant le solde modélisé par le département de la démographie aux immigrants en provenance de l'étranger observés au niveau métropolitain à partir du recensement. À l'instar des données des EAR et de l'état civil, ce solde avec l'étranger est un solde sur cinq ans.

Pour certains âges peu nombreux, une correction s'impose afin de réconcilier les deux sources. Les émigrants vers l'étranger au niveau métropolitain sont dès lors recalculés en retranchant le solde aux immigrants corrigés. La correction appliquée aux immigrants vers la métropole est ensuite ventilée localement, au prorata des immigrants sur la zone dans les immigrants France entière par sexe et âge. Cette ventilation permet d'obtenir des immigrants corrigés pour chaque zone proches de ceux observés au recensement et compatibles, une fois agrégés, avec le solde métropolitain.

Au terme de cette procédure, nous disposons donc :

- ❖ d'immigrants en provenance de l'étranger au niveau métropolitain et pour chaque zone ;
- ❖ d'émigrants vers l'étranger au niveau métropolitain.

Déterminer les effectifs des émigrants pour chaque zone consiste à spécifier une clé de répartition permettant de ventiler les émigrants vers l'étranger estimés au niveau de la métropole.

La clé de répartition retenue consiste à calquer la répartition de ces émigrants sur l'intensité d'immigration de chaque zone (tous sexes et âges confondus). Ainsi, si une zone Z accueille  $x$  % des immigrants métropolitains alors elle concerne  $x$  % des émigrants métropolitains de chaque âge.

Pour les âges inférieurs à 5 ans, la commune de résidence antérieure n'est pas connue. Il n'est donc pas possible de distinguer les enfants immigrants de l'étranger des autres enfants. Ainsi, les « immigrants en provenance de l'étranger » ne comprennent que des observations correspondants à des âges supérieurs ou égaux à 5 ans. Pour les âges 0 à 4 ans, c'est la modélisation des naissances et les ratios naissances-enfants qui permettent de calculer les populations correspondantes.