

CHAPITRE 1

PRÉSENTATION DE LA BASE

DU SUIVI DE LA DEMANDE TOURISTIQUE

Ce premier chapitre de la troisième partie est un prélude méthodologique et statistique, nécessaire à l'analyse de données.

- La première section présente le panel S.D.T (Suivi de la Demande Touristique) de la SOFRES. Comme prélude au travail d'exploitation, la première section délimite le champ de notre étude et définit les notions et concepts spécifiques au champ touristique.

- La deuxième section de ce chapitre donne au lecteur des points de repères statistiques à caractère général sur les principales variables utilisées : distributions et concentrations des distances, des fréquences et du temps de sortie.

SECTION 1 - PRESENTATION ET TRAITEMENT DE LA BASE DE DONNÉE S.D.T 1995 -2001

1.1 S.D.T : Présentation de la base de la Direction du Tourisme

Le panel S.D.T supervisé par la Direction du Tourisme (DT) et animé par l'institut Taylor Nelson Sofres (TNSO), vise à connaître l'ensemble des déplacements touristiques des Français, quels que soient leur motif et leur durée. L'enquête cherche à mesurer les déplacements touristiques des Français qui répondent à un motif personnel, professionnel ou mixte, en France ou à l'étranger.

Pour cela, l'enquête S.D.T fait appel au panel de la SOFRES dit METASCOPE. Depuis Avril 1990, le panel utilisé portait sur 10 000 foyers, représentatifs de ménages ordinaires, interrogés mensuellement. En mars 1999, le panel METASCOPE est porté à 20 000 foyers (pour un taux moyen de réponse de 75%). Les panélistes sont recrutés au cours des différentes enquêtes au téléphone ou face à face, et sur fichiers spécialisés. A la différence de l'enquête vacances de l'INSEE, qui prend en compte les seuls déplacements impliquant *quatre nuitées* passées en dehors du domicile, le panel S.D.T est plus exhaustif. En effet, il prend en compte les déplacements à partir *d'une nuitée passée en dehors du domicile*.

1.2 Champ de l'étude

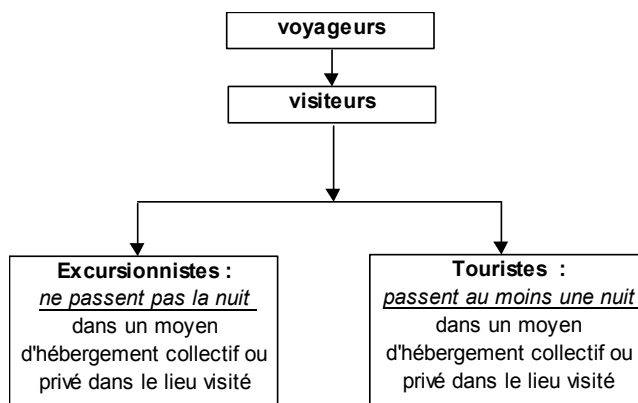
1.2.1. Définition du tourisme et distinction Séjour / Voyage : un voyage = un « mono » séjour

Selon l'Organisation Mondiale du Tourisme (O.M.T), le tourisme comprend :

« les activités déployées par les personnes au cours de leurs voyages, et de leurs séjours dans les lieux situés en dehors de leur environnement habituel, à des fins de loisirs, pour affaires et autres motifs. » (O.M.T, cité dans Mémento du Tourisme, 2002)

Tous les voyageurs qui intéressent le tourisme sont désignés par le terme de *visiteur* qui est le concept de base de l'ensemble du système des statistiques du tourisme. On distingue deux types de visiteurs, comme le précise la figure ci-dessous :

Figure 57. _ Catégories de voyageurs / visiteurs



Notre étude statistique portera exclusivement sur l'observation des comportements touristiques, et non des comportements excursionnistes. Bien que le cadre analytique mis en place n'exclut pas les excursions à la journée, le traitement statistique ne permet pas de vérifier la pertinence des hypothèses faites sur ce type de séjour ¹.

Une deuxième distinction importante adoptée par la terminologie de l'enquête S.D.T, concerne la définition du voyage et du séjour .

- *Le voyage* est défini comme tout départ du domicile, et retour à celui-ci avec au moins une nuit passée en dehors.

- *Le voyage se décompose en « séjour(s) »*. Le séjour est défini par le fait d'avoir passé au moins une nuit à l'extérieur en lieu fixe et sa durée est aussi comptabilisée en nuitées. Un voyage peut donc contenir plusieurs séjours si le vacancier a changé *au moins une fois* de lieu durant son voyage. Prenons l'exemple d'un touriste qui passerait 15 jours en dehors de son domicile. Ce dernier a pu visiter les châteaux de la Loire la première semaine, et sa famille en Bretagne durant la deuxième. Il aura donc réalisé deux séjours dans un même voyage.

Les motifs des voyages et des séjours retenus par le S.D.T sont de plusieurs types :

- *agrément* : salon, foire, exposition, croisière, manifestation sportive, stage sportif, manifestation culturelle, festival, visite, vacances, tourisme...

- *familiaux* : visite à la famille, aux amis, réunions de famille (mariage, baptême..)

- *autres raisons personnelles* : séjour linguistique, cure, thalassothérapie...

Le motif du voyage peut également être de type *professionnel* lorsqu'il s'agit de réunion, rendez-vous professionnel, stage, formation, congrès, colloque, séminaire. Les « VRP » sont exclus des observations car le mode d'enquête par voie postale rend difficile leur interrogation.

1.2.2 Traitement effectué sur la base concernant la distinction voyage - séjour

Le fichier dont nous disposons est un fichier d'enregistrement « séjour ». Chaque ligne correspond donc à un séjour effectué par un panéliste. Lorsqu'un voyage contient plusieurs séjours, il y a autant

¹ Les excursions à la journée ont été récemment intégrées au S.D.T. Une étude récente du SES de mai 2004, réalisée à partir de la base S.D.T, montre que plus du tiers des voyages à longue distance des Français est effectué dans la journée sur la période d'avril 2002 à mars 2003 et en voiture huit fois sur dix. L'étude de ce type de déplacements s'intégrerait parfaitement dans notre cadre analytique dans la mesure où la combinaison de distance et de temps qu'ils mettent en œuvre est encore plus parfaitement « séparable » dans la mesure où le temps de transport sur place est quasi nul. Par conséquent pour les excursions, il est parfaitement possible de connaître précisément et sans biais le ratio temps de transport / temps de séjour. Ce type de déplacements pourra faire l'objet d'une étude ultérieure.

de lignes que de séjours dans un même voyage à condition que les dates de départ et de retour soient identiques d'une ligne à l'autre. Chaque ligne correspond à une commune de destination.

Les double séjours dans un même voyage représentent 11% des séjours enregistrés. Leur poids est donc relativement faible. Intégrer ces doubles séjours est problématique dans la mesure où les estimations des temps de transport et des distances se compliquent : en effet, il est difficile de savoir combien de temps l'individu est resté à chacune des destinations, et donc de rattacher une distance à une durée de séjour sur un lieu donné. Aussi, pour simplifier le traitement, le choix a été fait de ne considérer que les voyages mono-séjour (soit $\text{numsej} = 1$). Avec cette simplification, le voyage devient équivalent à un séjour. Le temps passé à l'extérieur du domicile est mesuré en nombre de nuitées passées au même lieu principal d'hébergement déclaré sans changement de lieu de destination durant la durée totale du voyage. Par conséquent, l'appellation séjour et/ou voyage est donc parfaitement équivalente dans le traitement effectué.

1.2.3 Populations et motifs sélectionnés : actifs de 18 à 65 ans pour motifs loisirs au sens large

- La première sélection effectuée sur la base porte sur *les populations* : leur âge et leur activité. Seules *les populations actives de 18 à 65 ans* sont retenues. Ces populations sont triées à partir de la catégorie socioprofessionnelle (code Sofres, allant de 10 à 69). Les chômeurs et étudiants de moins de 18 ans sont exclus du traitement (catégories 71 à 86). Ce choix de ne retenir que les populations actives se justifie dans la mesure où notre première partie théorique a fondé son raisonnement sur la rareté du temps...et donc le prix implicite du temps. Les catégories inactives comme les retraités¹ ou les étudiants accordent certes une valeur à leur temps, mais leurs pratiques de loisir et de tourisme n'évoluent pas dans le même système de contraintes que celui des actifs. Aussi, les pratiques de vacances de ces populations méritent un examen à part entière qui n'est pas l'objet de notre étude.

- La deuxième sélection porte sur les *motifs du déplacement*. En effet, les différents motifs que proposent le S.D.T n'ont pas tous été retenus. Le tableau ci-après décrit les motifs retenus comme faisant partie du loisir, au sens large, et mesure leur part relative dans le total des mono-séjours :

¹ Sur les pratiques des seniors, nous renvoyons le lecteur à la récente étude (octobre 2004) réalisée par Abdel Khiati et Hélène Denizet du Département de la Stratégie, de la Prospective, de l'Evaluation et des Statistiques de la Direction du Tourisme. Cette étude a pour objet l'analyse de l'ensemble des séjours des Français de 55 ans et plus.

Tableau 4 . _ Pourcentage des motifs déclarés dans le total des mono-séjours. Actifs de 18 à 65 ans.

Code SOFRES	motifs	% du total des mono-séjours (1995 – 2001)	N : mono-séjours (1995 – 2001)
11	Manifestations culturelles	1,2%	971
12	Manifestations sportives	1,3%	1078
14	famille	50,6%	41173
15	Vacances, tourisme, loisir	40,5 %	32988
16	Etapas chemin vacances	1,1%	859
20	Visite amis	5,3%	4352

Source statistique : DT – TNSO ; SDT 1995-2001 ; calculs, V. Bagard

Le premier motif est celui de famille avec 50% du total et 41 000 mono-séjours, le deuxième motif est celui des vacances, du tourisme et des loisirs avec 40,5% du total et 33 000 mono-séjours. Si ces deux blocs de motifs sont bien identifiables, la frontière qui les sépare n'est pas étanche. Pour s'en rendre compte, il suffit de croiser chacun des motifs avec le type d'hébergement : 15% des séjours « déclarés » en motifs vacances sont en réalité passés dans un lieu « déclaré » famille. Ces 15% représentent 4800 séjours, ce qui n'est pas négligeable. De même, 6% des séjours déclarés en vacances, loisir, tourisme, sont passés chez des amis et 20% des séjours déclarés en motif famille sont passés dans un lieu d'hébergement autre que la famille (ce qui s'explique si la famille sur place ne peut héberger ses proches).

Ces chevauchements impliquent de retenir une définition large des loisirs en intégrant la famille et les amis. Cette sélection garde ainsi un échantillon important qui par la suite permettra de conserver une robustesse aux estimations lorsque les désagréations seront plus fines.

1.3 Nouvelles variables et traitement longitudinal du panel

La principale valeur ajoutée à la base S.D.T est l'introduction des distances à vol d'oiseau pour l'ensemble des modes motorisés, y compris le mode qui nous intéresse le plus : l'avion. Sur ce point, le traitement des données sur les distances aériennes est un enjeu clé dans toutes les études de prospective de la mobilité dans la mesure ou la soutenabilité du secteur des transports nécessite de mesurer et comprendre les logiques de la mobilité récréative de longue distance. Le potentiel de développement de cette mobilité semble *a priori* plus élevé que la mobilité dite non contrainte. Le recours fréquent à la vitesse aérienne pour les motifs de loisir change considérablement les échelles des distances annuelles habituellement constatées sur route ; de même que les émissions de CO₂ ...

Pour estimer les distances sur terre et dans l'air, les calculs suivants ont été menés :

- Lorsque les points de départ et de destination sont nationaux (la route, le fer, l'eau, l'air), la formule utilisée pour calculer les distances à vol d'oiseau est classique : de centre *commune départ* (x_1, y_1) à centre *commune destination* (x_2, y_2) , la distance entre les points à vol d'oiseau est calculée par la formule : $\sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$. Les coordonnées géographiques utilisés sont de type Lambert 2 étendu.

- Lorsque les points de départ sont nationaux et les points d'arrivée sont hors des frontières de la France métropolitaine, pour la route, le fer, l'avion et l'eau, le calcul est différent et n'utilise plus les coordonnées géographiques nationales en Lambert 2 étendu. La base de données utilisée provient du travail de G.Gaulier et S.Zignago au CEPII¹. Cette base fournit, *au départ de Paris*, les distances à vol d'oiseau géodésiques pour l'ensemble des principales villes du monde. Les données sont bilatérales. Par conséquent, sur l'ensemble des déplacements transfrontaliers, nous faisons l'hypothèse d'un départ de Paris et d'une arrivée dans la ville principale du pays dont nous faisons l'hypothèse qu'elle correspond pour l'avion à l'aéroport principale et pour le train à la gare principale.

Après avoir décrit les modalités de construction de la variable distance, nous abordons maintenant le deuxième travail spécifique réalisé sur la base S.D.T, à savoir la construction d'une variable longitudinale de contrôle.

Le S.D.T offre la possibilité d'un traitement longitudinal des données et présente l'avantage de pouvoir examiner l'évolution des déplacements de loisir et de tourisme *de 1995 à 2001*. Si l'objet de notre étude n'est pas de faire un suivi longitudinal, et de traiter l'évolution de la mobilité sur la période 1995-2001, la possibilité d'un traitement longitudinal permet de donner *des valeurs moyennes sur la période*. Par exemple, pour ce qui concerne les distances parcourues, nous entendons par valeur moyenne, le résultat de la division des distances parcourues sur plusieurs années par le nombre d'années de présence dans le panel. Cette même opération peut être effectuée pour le temps de sortie et la fréquence des sorties. Plus le panéliste est resté longtemps dans le panel, plus cette valeur moyenne garantira une certaine « récurrence » de son comportement et tempèrera les comportements extrêmes, non récurrents, dans le temps. La question est plus problématique pour les panélistes restés *une seule* année dans la mesure où aucun moyen ne nous permet de tempérer d'éventuels comportements extrêmes.

Pour utiliser ces valeurs moyennes il est toutefois nécessaire de vérifier au préalable la fiabilité du panéliste dans le temps. C'est à ce niveau que se situe notre travail. Lorsque le total des distances

¹ Pour un descriptif technique : http://www.cepii.fr/anglaisgraph/bdd/distance/noticedist_en.pdf

parcourues sur plusieurs années est divisé par le nombre d'années de fidélité, il est nécessaire d'être certain que l'identifiant panéliste concerne le même individu ou du moins, un « clone » de celui-ci. Le risque étant que ce « clone » n'en soit pas vraiment un et que le panéliste ait vu ses caractéristiques socio-économiques changer radicalement au cours du temps, comme par exemple son statut professionnel. Cette base étant toutefois largement contrôlée et suivie, la qualité et la fiabilité des données sont sans doute ce qui se fait de mieux en la matière. Néanmoins, le fait de travailler sur des actifs peut justifier ce contrôle. En effet, les actifs sont légèrement moins représentés et moins fidèles que les inactifs, plus nombreux et plus fidèles. En outre, le profil d'un actif est plus susceptible de changer dans le temps que celui d'un inactif (revenu, lieu d'habitation). Par conséquent, pour assurer un éventuel traitement longitudinal du panel, un dépistage des panélistes multi-profil a été entrepris.

La base de départ comprenait 192 072 enregistrements sur 6 ans soit en moyenne 32 012 séjours par an. Après sélection des actifs de 18 à 65 ans, pour les motifs cités ci-dessus, la base comprend 93 496 enregistrements séjours (mono-séjours et multi-séjours) sur une durée de 6 ans soit en moyenne 15 580 séjours par an pour 9846 panélistes sur 6 ans. Le travail est effectué sur la base multi-séjours afin de ne pas exclure des critères de contrôle les panélistes qui pratiquent les multi-séjours. La distinction sera faite par la suite avec les mono-séjours.

Le test réalisé sur cette base est le suivant. Dans un premier temps, ce test est réalisé sur l'ensemble des panélistes qu'ils soient restés 1 an ou 7 ans dans le panel.

- un premier contrôle opéré sur la variable de l'âge : en théorie, un panéliste resté par exemple de 1995 à 1999 voit son âge augmenté de 4 ans en 1999. Si son âge a bien augmenté de 4 ans **et** qu'aucune des variables socio-économiques se rapportant à son numéro de panéliste n'a changé, alors le numéro de panéliste cache le même individu ou en tous cas son clone parfait. Combien de panélistes sont dans ce cas de figure ?

Sur l'échantillon de 93 496 enregistrements, 39 244 enregistrements sont déclarés par des panélistes dont l'ensemble des caractéristiques socio-économiques (agglomération, niveau d'instruction, revenu, activité) n'ont absolument pas changé dans le temps sauf leur âge, ce qui représente 42% de panélistes que l'on peut qualifier avec certitude de mono-profil.

Cela dit, parmi ces 39 244 enregistrement séjours, il est *logique* que les panélistes restés une année soient nécessairement mono-profil. Or ces derniers représentent 17 024 enregistrements sur les 39 244, il reste donc $39\,244 - 17\,024 = 22\,220$ enregistrements correspondant à des panélistes restés *au moins deux ans* sur un total initial de 93 496 enregistrements soit 24% du total. Ce pourcentage est faible, mais il répond au strict critère de la progression de l'âge en maintenant constantes toutes les variables de contrôles.

Il existe toutefois des panélistes mono-profils ayant connu une évolution de leur profil à la marge. Or, ces derniers ne peuvent être exclus, d'autant qu'ils apportent de la variance à l'échantillon. Il reste à s'assurer qu'ils sont bien mono-profils. Ainsi, pour les panélistes ayant vu leur profil changer à la marge, un changement d'ordre 1 est toléré sur une des trois variables suivantes : agglomération, niveau d'instruction, revenu, tout en contrôlant en parallèle que l'activité reste identique et que l'âge suit une progression logique.

Ce faisant, nous récupérons 832 séjours sur le critère de l'agglomération, 1285 sur le critère du niveau d'instruction et 20 668 sur le critère du revenu... ce bon score sur le revenu est logique pour des actifs, d'autant plus que la fin de la décennie a été marquée en France par une forte croissance économique. Au total, 22 785 enregistrements sont donc récupérés.

L'échantillon se porte alors à $22\ 220 + 22\ 785 = 45\ 005$ enregistrements déclarés par des panélistes mono-profils restés au moins deux ans sur les 93 496 initialement déclarés par l'ensemble des panélistes soit 48%. Ce pourcentage est déjà plus élevé mais là encore il élimine l'ensemble des panélistes restés une seule année qui constituent un nombre d'autant plus important que l'échantillon a été doublé en 1999.

En tenant compte des panélistes restés une année, nous obtenons : $45\ 005 + 17\ 024 = 62\ 029$ enregistrement voyages mono-séjours et multi-séjours soit 66% de l'échantillon initial. Par extension des critères de contrôles nous sommes donc passés de 24% à 48% puis 66%. Le tableau ci-dessous détaille la structure de l'échantillon selon les années validées par le panéliste :

Tableau 5. _ Structure de l'échantillon (62 029) selon le nombre d'années de fidélité validées selon les critères de cohérence de profil évoqués ci-dessus

années validées	nb de séjours	%	Pourcentage cumulé
1	17 024	27	27
2	45005	31	59
3		12	71
4		9	80
5		9	89
6		7	96
7		4	100
<i>Total séjours actifs validés</i>		62 029	100
<i>Total séjour actifs</i>		93 496	

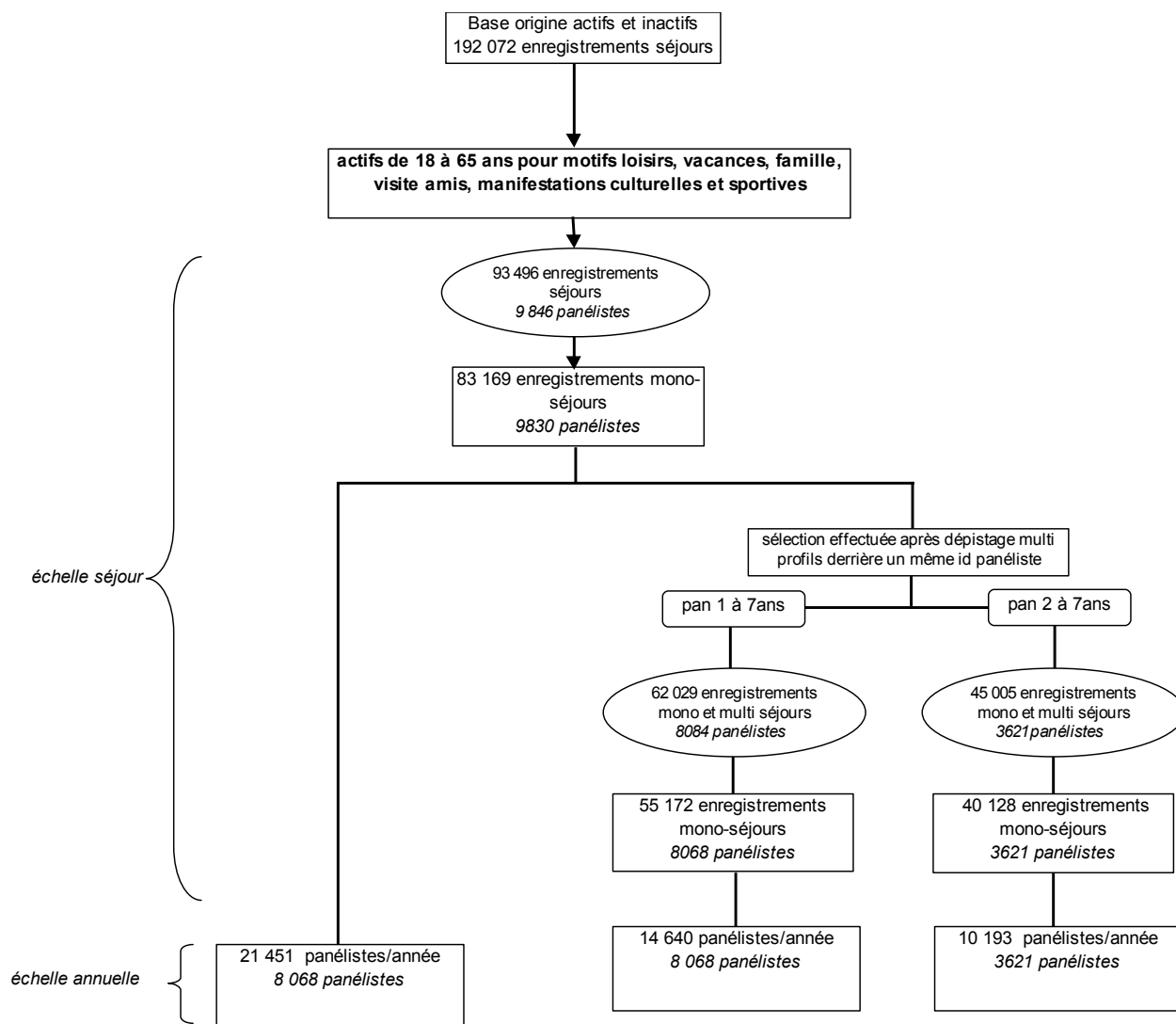
Source statistique : DT – TNSO, SDT 1995 – 2001 : 9830 actifs de 18 à 65 ans, calculs, V. Bagard

Sur l'échantillon 45 005, le calcul de moyennes longitudinales est possible. Les comportements extrêmes sont étalés au moins sur deux années. Quant à l'intégration des panélistes restés une seule année dans le traitement longitudinal, elle présente un avantage, celui d'un échantillon plus important

et un inconvénient, la perte de contrôle longitudinal sur les panélistes restés un an et ayant eu des pratiques de vacances atypiques.

Pour clarifier la structure l'échantillon, et intégrer la distinction entre les effectifs mono-séjours et multi-séjours, un arbre est présenté ci-dessous :

Figure 58. Structure de la base S.D.T 1995-2001/ DT - TNSO



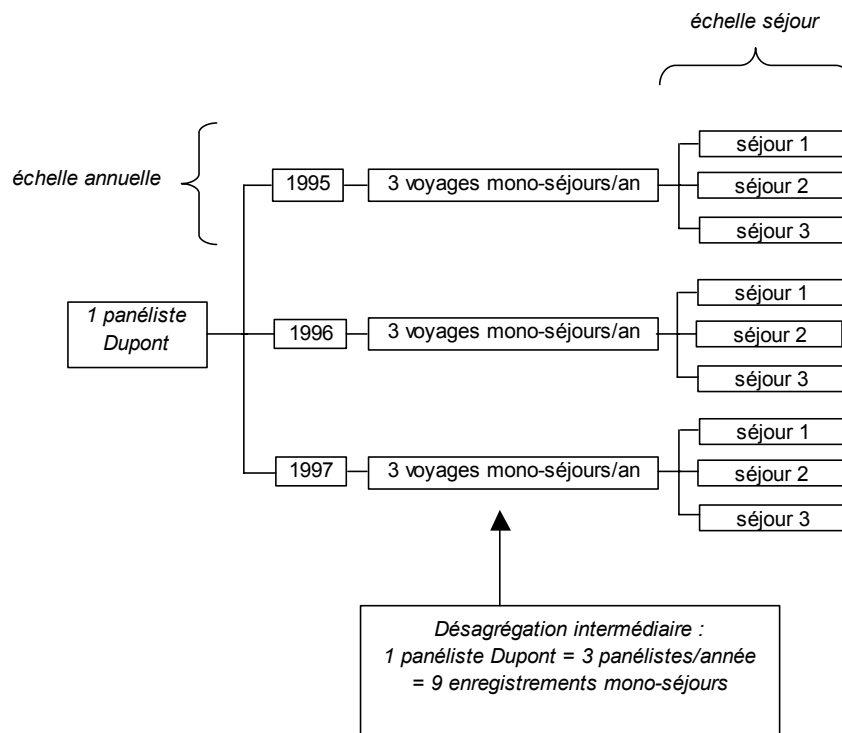
Cette désagrégation appelle deux remarques :

- Sur le passage a priori curieux de 9846 à 9830 panélistes et parallèlement de 93 496 à 83 169 enregistrements séjours. Les 10 327 séjours de différence ne sont pas effectués par 16 panélistes... Ces 16 panélistes sont en effet ceux n'ayant déclaré que des multi-séjours sur l'ensemble de leurs enregistrements. Or lorsque tous les multi-séjours sont pris en compte, c'est l'ensemble des panélistes ayant effectué au moins un multi-séjour parmi des mono-séjours qui est pris en compte. Cette

différence est donc prélevée sur un ensemble beaucoup plus large de panélistes puisque 4 467 panélistes ont fait au moins un ou plusieurs voyages multi-séjours.

- La deuxième remarque porte sur les avantages de travailler au niveau intermédiaire de désagrégation. La variable des mono-séjours par panéliste et par année laisse les années indépendantes et permet d'agréger l'ensemble de l'échantillon sur l'unité de temps annuelle. A ce niveau intermédiaire, les séjours réalisés sur une même année par un même panéliste sont tous rattachés à un même numéro de panéliste/année. L'arbre ci-dessous permet de situer ce niveau intermédiaire :

Figure 59. _ Niveau intermédiaire de désagrégation



Ce schéma permet de mieux comprendre notre nécessité de travailler sur deux niveaux de désagrégation pour être cohérent avec le cadre analytique.

1.4 Cadre analytique sur deux échelles : travail sur deux niveaux

En effet, il est nécessaire de travailler sur l'échelle du séjour qui implique les variables $D_i, TS_i, TT_i, \overline{V}_{ij}$, soit la respectivement la distance par séjour, le temps de séjour en nuitées, le temps de transport par séjour, la vitesse moyenne du mode j.

Il est également nécessaire de travailler sur l'échelle annuelle qui implique les variables $DA, TAS, TTA, \overline{VA}$, soit respectivement la distance annuelle, le temps de sortie annuel en nuitées, le temps de transport annuel, la part des modes utilisée dans les distances annuelles, la fréquence annuelle de sortie.

Le travail sur deux échelles de temps nécessite de travailler sur deux niveaux d'échantillon :

- *pour travailler à l'échelle du séjour* : les variables moyennées longitudinales n'ont aucun sens. Elles ne peuvent être utilisées comme variables de contrôle sans gommer la relation fine entre une durée de séjour et une distance. L'échantillon 83 169 est utilisé. Ce dernier est éventuellement ramené à 55 172 si le fait de sélectionner les panélistes fidèles entraîne une différence significative. Ce qui n'est pas forcément le cas lorsque les variables socio-économiques n'entrent pas en ligne de compte.

- *pour travailler à l'échelle annuelle* : le niveau intermédiaire est adéquat. Les variables panélistes/année peuvent être obtenues de deux manières : soit en passant par les variables moyennées longitudinales \overline{DA} , \overline{TSA} , \bar{i} pour un gage de fiabilité, (échantillon 10 193 panélistes/année et 3621

paénélistes), soit en passant par les variables/année non modifiées tel que $DA = \sum_n^i D_i$ et

$TSA = \sum_n^i TS_i$, l'échantillon est alors de 21 451 panélistes/année et 8 068 panélistes.

L'utilisation des valeurs moyennées sur plusieurs années pose toutefois un problème car elle empêche d'effectuer les calculs qui croisent les deux échelles comme par exemple la part que représente chaque séjour selon sa durée dans le temps de sortie annuel total, ou encore la part des distances effectuées dans les distances annuelles selon la durée des séjours. (cf. chapitre 3 sur l'échelle annuelle).

Par conséquent, la solution la plus souple et rigoureuse consiste à utiliser les variables longitudinales moyennes comme des variables de contrôle chaque fois que cela est possible. Le traitement de données à proprement parler est effectué sur l'échantillon le plus large possible.

- Sur l'échelle séjour, l'échantillon 83 169 est utilisé avec un contrôle par l'échantillon le plus fiable 45 005.

- Sur l'échelle annuelle, l'échantillon 21 451 est utilisé, ce dernier est issu de la même structure que l'échantillon 83 629. Il est contrôlé, lorsque cela est possible, par l'échantillon 10 193.

SECTION 2 - DISTANCE ET TEMPS DE SORTIE ANNUEL : DISTRIBUTION ET CONCENTRATION

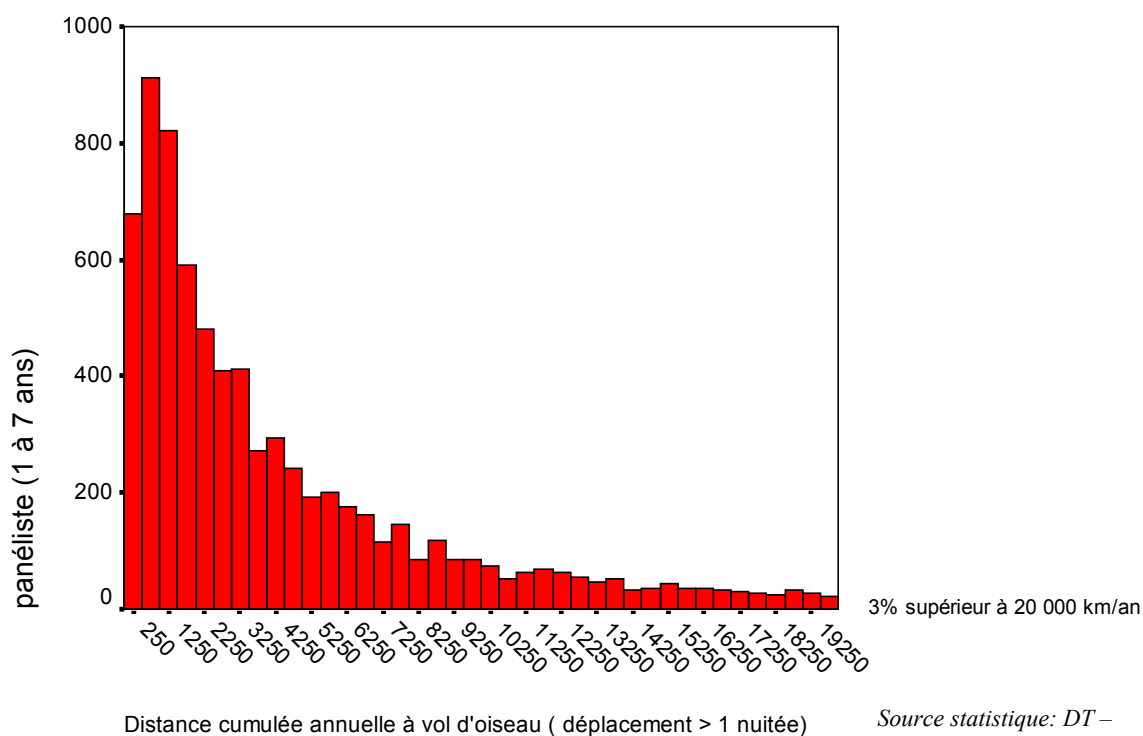
Cette deuxième section présente les distributions du temps de sortie annuel et des distances annuelles. L'échantillon utilisé dans cette section est le plus robuste : 10 193 mono/séjours année pour 3621 panélistes fidèles de 2 à 7 ans. Nous avons toutefois ajouté séparément les panélistes représentés une seule année afin d'évaluer leur impact sur la moyenne. Cet impact traduit en partie le passage à partir de 1999 à 20 000 panélistes.

NB : Conformément aux critères retenus dans la première section, les individus représentés sont des actifs de 18 à 65 ans se déplaçant pour motifs loisirs au sens large (motifs famille, amis, vacances, manifestations sportives et culturelles). Ces séjours comportent au moins une nuitée hors du domicile.

2.1 Distribution de la distance cumulée annuelle

Les distributions pour les distances et les temps de séjours sont estimées pour l'ensemble des années 1995-2001. Ces distances moyennes sont obtenues par panélistes, ce qui signifie que nous utilisons la base avec le contrôle longitudinal (les panéliste fidèles mono-profil).

Figure 60. _ Distribution des distances cumulées annuelles (8068 panélistes de 1 à 7 ans)



Ex. calcul de distance : si le panéliste est resté deux années dans le panel et a parcouru sur l'ensemble des modes 4000 km, sa distance moyenne annuelle sera de $4000 / 2 = 2000$ km.

Comme on le constate sur le diagramme en bâton, la diminution des distances est très progressive et ne présente pas de rupture marquée : 3% de l'effectif des actifs, non pris en compte sur le diagramme, parcourent plus de 20 000 km/an pour les loisirs tous modes confondus.

Le tableau ci-dessous présente l'ensemble des caractéristiques de la distribution pour plusieurs séries de panélistes : (pour le coefficient correcteur proposé (1,15) cf. note p. suivante).

Tableau 6. _ Tableau récapitulatif des distances annuelles avec et sans correction du vol d'oiseau (coef. standard 1,15)

	3621 panélistes de 2 à 7 ans	8068 panélistes de 1 à 7 ans	3621 panélistes de 2 à 7 ans	8068 panélistes de 1 à 7 ans	3621 panélistes de 2 à 7 ans	8068 panélistes de 1 à 7 ans
	Distance annuelle tous modes à vol d'oiseau		Distance annuelle <u>routes</u> à vol d'oiseau		Distance annuelle dont distance terrestres (route + fer) corrigées du vol d'oiseau (coef. 1,15)	
moyenne	4 085 km/an	3 785 km/an	2 158km/an	2 119km/an	4 454km/an	4 083km/an
écart type	5 256 km/an	5 542 km/an	2 235 km/an	2 431 km/an	5 582 km/an	5 773km/an
moyenne tronquée à 5%	3 363 km/an	2 988 km/an	1 891km/an	1 813km/an	3 700km/an	3 268km/an
médiane	2 311 km/an	1 960 km/an	1 557 km/an	1 381 km/an	2 619 km/an	2 198 km/an
Maximum	112 074 km/an	166 115 km/an	45 404 km/an	54 102 km/an	126 700 km/an	167 305 km/an
décile supérieur	9 559 km/an	8 847 km/an	4 421 km/an	4 532 km/an	10 312 km/an	9 356km/an

Source statistique : DT – TNSO, SDT 1995 – 2001 : 9830 actifs de 18 à 65 ans, calculs, V. Bagard

La distance annuelle moyenne parcourue pour l'ensemble des modes est de 4 085 km/an à vol d'oiseau et 4 454 km lorsque les distances terrestres (soit la route et le fer) sont corrigées du coefficient (1,15). Dans ce cas, les distances aériennes restent par définition en vol d'oiseau même si ce n'est pas le cas dans la réalité. Cette moyenne de 4454 km représente 1,7 fois la médiane (2619 km), signe d'une distribution tirée vers le haut par les forts kilométrages et l'aérien. De même, pour l'ensemble des modes, la moyenne tronquée à 5% est inférieure de 722 km.

Cette estimation donne donc une idée de la distance totale parcourue pour le tourisme et les vacances sur tous les modes, et non de la seule distance parcourue sur route qui est de 2 158km (environ 2 500 km en réel avec un coefficient de 1,15). Les distances sur route comprennent également les autres modes (camping car, minibus, car, moto) auxquels s'ajoutent la voiture particulière et la voiture de localisation. Lorsque l'on isole la distance moyenne effectuée en voiture particulière seulement, la moyenne annuelle est 1 955 km à vol d'oiseau par an soit environ 2 250 km en réel (sur le même échantillon 3 621). Que représente ces 2 250 km dans la distance moyenne annuelle

parcourue par un véhicule particulier en France ? Selon les comptes du transport 2003 (INSEE), la distance moyenne est de 14 100 km, et se révèle relativement stable depuis 1995, du moins pour les véhicules essence. On peut en déduire que les distances réalisées en voitures particulières pour les déplacements d'une nuitée ayant pour motifs le loisirs, le tourisme et la famille représenteraient en moyenne 16% de cette distance moyenne annuelle par véhicule en France ((1 955 km x 1.15) / 14 100 km). Il faudrait toutefois ajouter aux 1955 km les déplacements sur les lieux de destination des séjours que nous ne pouvons estimer ici. Par conséquent, les 16 % sont sous-estimées pour la voiture particulière.

On peut faire trois autres remarques concernant ce tableau :

- Sur la concentration dans le temps de cette mobilité. En effet, lorsque la distance moyenne tous modes (4 454 km) est ramenée à son temps moyen effectif de sortie sur l'année (soit 24 nuitées), cela équivaut à une moyenne de 185 km par jour. Ces 4 454 km sont donc concentrés sur 7% du temps annuel, ce qui donne une idée de la concentration temporelle de cette mobilité.

- Sur la mobilité des résidents de l'agglomération parisienne. Ces derniers doublent la moyenne nationale en parcourant 7 280 km /an *tous modes confondus* pour les loisirs. Ces derniers se déplacent plus que la moyenne nationale en VP avec 2787 km/an à vol d'oiseau contre 2 193 km pour la moyenne nationale. L'écart à la moyenne (7 280 km) est donc imputable au train et surtout à l'avion. Il est très probable que cet écart de distance s'explique par un effet revenu, un effet d'instruction, un « besoin d'évasion », facilité par une meilleur desserte autoroutière et un recours accru à l'aérien.

- Sur l'introduction des panélistes restés un an dans le panel et leur impact sur la moyenne des distances parcourues. On remarque en effet que la distance moyenne annuelle parcourue sur *l'ensemble des modes ou seulement sur la route est inférieure lorsque ces panélistes sont ajoutés à la moyenne*. Cette différence se retrouve sur la médiane et le décile supérieur. En revanche, l'écart type est supérieur, signe d'une dispersion plus grande. Cette dispersion est-elle due à l'élargissement de l'échantillon en 1999 dont proviennent la très grande majorité des panélistes restés un an ? Il est fort probable que nous retrouvions ici le constat de Christophe Terrier¹ :

« On enregistre des différences de résultats entre l'échantillon d'origine (10.000 individus) et l'échantillon total (20.000 individus). Le taux de départ est plus faible sur l'échantillon de 20.000 individus. Ce phénomène est observé à partir de juin 1999. » (Terrier, 2002, p.32)

Avons-nous ici une répercussion de cet effet sur les distances parcourues ? C'est probable. En outre, concernant les dispersions plus fortes autour de la moyenne, elles pourraient refléter une légère

¹ Direction du Tourisme. Présentation Taylor Nelson Sofres : « Le dispositif de Suivi de la Demande Touristique. Présentation des principaux points d'actualité méthodologique ». (2002)

réduction de l'intervalle de confiance concernant l'échantillon de 20.000 individus. L'hypothèse de Christophe Terrier est la suivante ; :

« Il semblerait que les écarts de structure d'échantillon ne soient pas les éléments essentiels pour expliquer cette différence, l'introduction de 10.000 nouveaux panélistes, avec leurs caractéristiques propres et non contrôlables induirait l'essentiel de ce choc en niveau. » (Terrier, 2002, p.40)

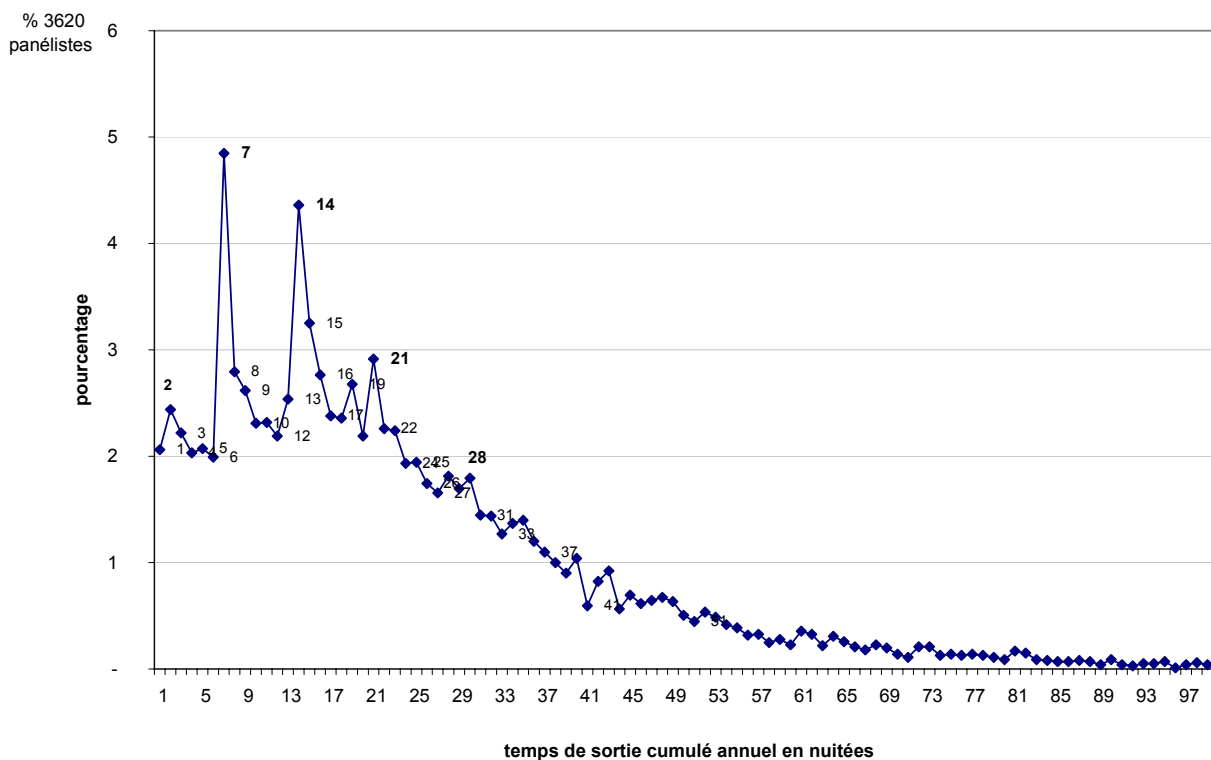
Encadré 5. _ Le coefficient correcteur

Il n'existe pas, à notre connaissance, de coefficient correcteur spécifique pour la longue distance en voiture particulière adapté aux déplacements de loisirs et de tourisme. Le coefficient correcteur généralement utilisé pour le milieu urbain (ville centre) de 1,3 surestimerait la distance réelle. Le coefficient correcteur généralement utilisé dans la littérature pour ramener la distance à vol d'oiseau du centre à la périphérie pour 80 km est de 20%. Cela équivaut à un coefficient de 1,2 soit 100 km en distance réelle pour 80 km à vol d'oiseau (cf. les études de Orfeuil (2002) et Madre (1997b), 1999). Dans notre cas, nous pensons qu'un coefficient de 1,2 est encore légèrement au dessus de la réalité. En effet, dans la base SDT, plus de 77% des déplacements sur route dépassent les 100 km aller à vol d'oiseau avec une moyenne de 232 km. Par conséquent, dans des déplacements longue distance, la sortie d'agglomération de départ et l'entrée dans l'agglomération de destination, pèsent moins sur la totalité du trajet alors qu'ils représentent les distorsions les plus significatives relativement à un tracé au vol d'oiseau. Un coefficient de 1,15 semble plus approprié à titre indicatif pour l'ensemble des distances terrestres y compris le train. Toutefois, ce coefficient n'a pas de valeur officielle et pour cette raison nous laissons la plupart du temps les distances à vol d'oiseau afin que nos résultats puissent être exploités avec d'autres coefficients.

2.2 Distribution du temps annuel moyen de sortie

La deuxième variable sur laquelle nous aimerions donner au lecteur une vue d'ensemble et des points de repère concerne le temps annuel moyen de sortie mesuré en nuitées. La distribution des durées cumulées de sortie est présentée dans la figure ci-dessous.

Figure 61. _ Distribution des durées annuelles de sortie pour motifs de loisirs



Source : DT – TNSO ;SDT 1995-2001 ; calculs, V.Bagard

La moyenne du temps cumulé de sortie annuel se situe à 24 nuitées. Quatre pics sont clairement identifiables dans cette distribution. Respectivement à 2, 7, 14, 21, 28 nuitées. Y-a-t-il réellement un effet semaine ? ou bien est-ce un effet de « chiffre rond » lorsque le panéliste remplit le questionnaire et arrondi la durée de ses séjours ? A première vue, la semaine passée hors du domicile apparaît comme « l'étalon temporel de référence », le « métronome » des vacances. Ce métronome est-il appelé à disparaître avec la moindre prégnance des rythmes sociaux ? Cette question forte est déterminante dans la construction des modèles de prospective¹ et suscite des débats. Le tableau ci-dessous présente les caractéristiques de cette distribution des temps de sortie annuel :

¹ Ces quatre pics ont-ils toujours existé, du moins depuis les années soixante ? Dans notre réflexion prospective, nous avons fait l'hypothèse d'une certaine stabilité de la taille des créneaux temporels. Ces quatre pics nous sont apparus comme les constantes du système vacancier français. A notre avis, ce sont les quantités de distances injectées dans chacun de ces créneaux qui constitue, le changement important, et ce grâce à la vitesse puisque l'équilibre temps de transport / temps de séjour reste en deçà d'un seuil maximal comme nous le verrons. Cette intensification s'explique d'un point de vue économique par la double baisse du prix du kilomètre en équivalent temps et en équivalent pouvoir d'achat. Dans cette

Tableau 7. _ Tableau récapitulatif des temps de sortie annuels

	3621 panélistes de 2 à 7 ans	8068 panélistes de 1 à 7 ans
	Temps cumulé de sortie annuel en nuitées	
moyenne	23,7 nuitées/an	21,8 nuitées/an
écart type	17,3 nuitées/an	17,8 nuitées/an
moyenne tronquée à 5%	22,1 nuitées/an	20,1 nuitées/an
médiane	19 nuitées/an	17 nuitées/an
décile supérieur	47 nuitées/an	45,5 nuitées/an

Source : DT – TNSO ; SDT 1995-2001 ; calculs V.Bagard

La moyenne du temps annuel de sortie est de 24 nuitées. Les panélistes restés une seule année n'ont quasiment pas d'impact sur cette moyenne, en revanche, ils diminuent la médiane d'une nuitée. En outre, 10% des panélistes passent plus de 47 jours en dehors de leur domicile.

2.3 Ces distributions cachent-elles de fortes concentrations ?

Ces distributions relativement régulières pour les distances annuelles et spécifiques pour le temps annuel de sortie cachent en contrepartie de fortes concentrations. Estimer ces effets de concentration est d'autant plus un enjeu lorsque nous intégrons les distances parcourues en avion et dans une moindre mesure en train. La suite de cette section consiste à donner au lecteur une vue d'ensemble des phénomènes de concentration des distances et des fréquences. Il est en effet très important de garder à l'esprit, comme nous allons le voir, que la très grande partie des distances et des sorties est accomplie par une minorité de panélistes, du moins pour les populations que nous étudions, c'est-à-dire les actifs de 18 à 65 ans.

L'indicateur utilisé est classique. Il s'agit de l'indice de concentration à partir duquel peut être dérivé une courbe de Lorenz et un indice de Gini. Du point de vue du calcul réalisé, une précision s'impose. Il s'agit toujours du même échantillon robuste de 3621 panélistes avec toutefois des données

logique, si le temps total annuel de congé reste stable, deux solutions se présentent pour augmenter et diversifier le nombre de sorties :

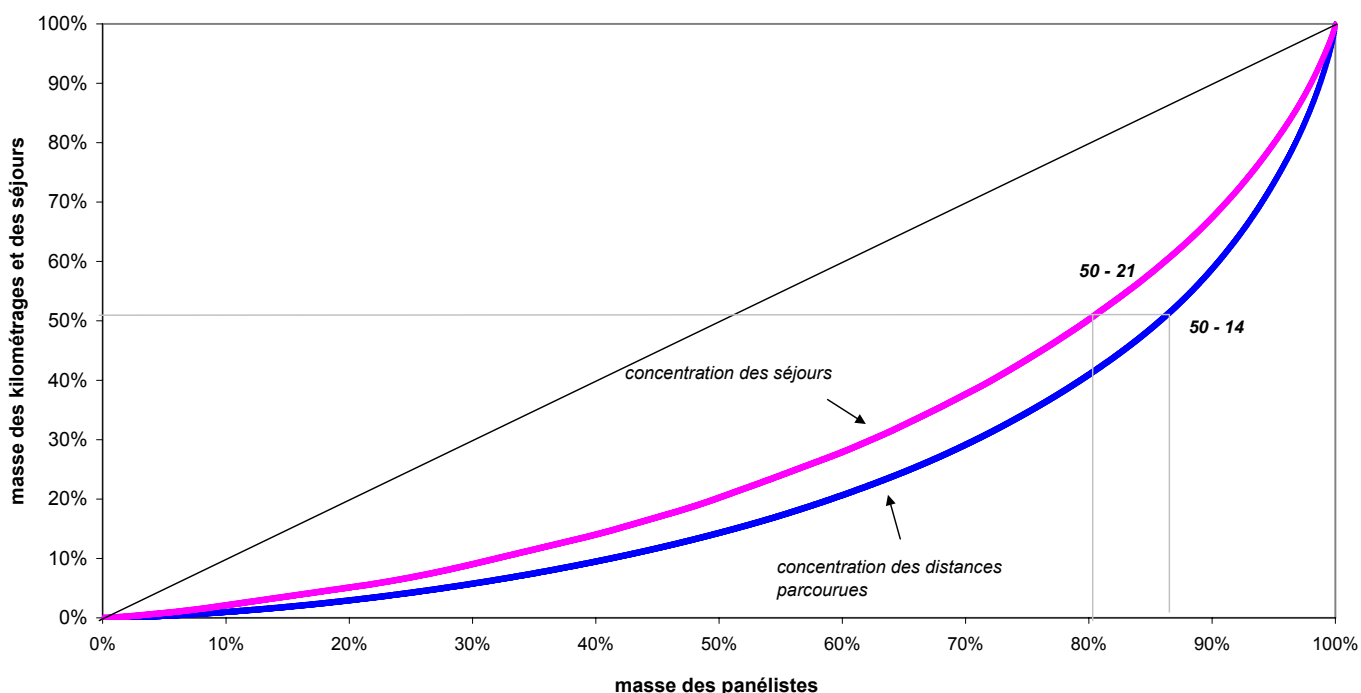
- d'une part, utiliser plus souvent les petits créneaux (excursions, week-end, ponts),
- d'autre part, fractionner le temps total (éliminer le long créneau de 21 nuitées ou le réserver aux destinations très lointaines, utiliser chaque année un créneau de 15 nuitées, puis fractionner le reste en semaine).

Selon nous, ce n'est pas l'unité de temps que la croissance économique remet en question, mais le rythme, le tempo, et la portée des déplacements. Ainsi, les évolutions les plus rapides ne sont pas là où l'ont croit qu'elles sont : c'est à dire dans les temporalités...mais plutôt dans les vitesses.

moyennées sur le nombre d'années durant lesquelles le panéliste est resté fidèle sous un même numéro. Toutes les distances annuelles sont donc en réalité des distances moyennes annuelles, de même que le nombre de sortie annuel est en réalité un nombre moyen de sortie annuel. L'utilisation du panel longitudinal est ici particulièrement utile pour fiabiliser les indices de concentration. En effet les valeurs moyennes contrôlent les comportements extrêmes. Les graphiques ci-dessous présentent les courbes de Lorenz pour les distances et le nombre de sorties :

2.2.1 Distances et nombre de séjours annuels moyens par panéliste

Figure 62. _ Concentration des kilométrages parcourus, et des séjours (3621 panélistes)



Source statistique: DT – TNSO ; SDT 1995-2001 ; calculs: V.Bagard

Le principe de la courbe de Lorenz est simple. Si la masse des panélistes se répartissait de manière totalement équitable la masse des distances parcourues et la masse des sorties, il y aurait alors « équirépartition ». Dans cette situation fictive, 50% des panélistes parcouraient 50% des distances totales et totaliseraient 50% des sorties. Nous serions alors le long de la droite qui coupe le rectangle en deux parties de surface égale.

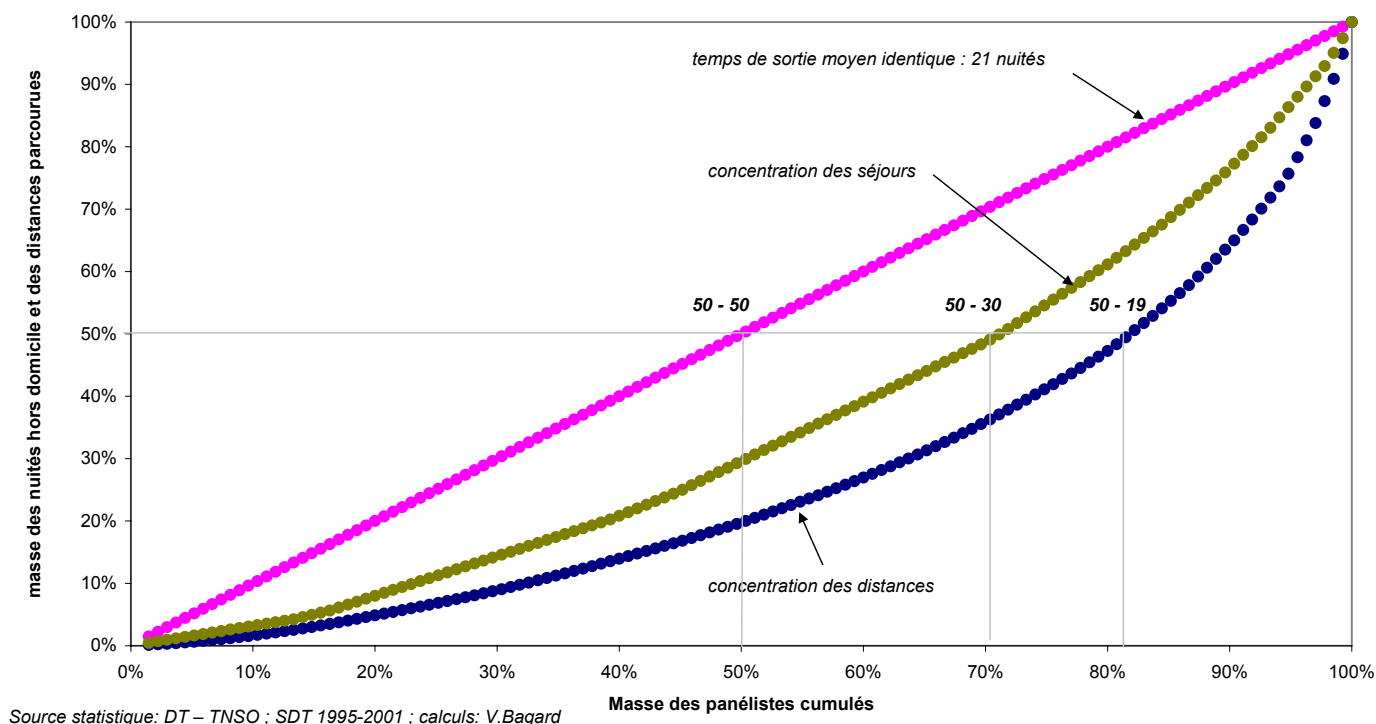
Or, on constate d'une part que 50% des distances (tous modes) sont effectuées par seulement 14% des panélistes. D'autre part, que 50% des séjours (mono-séjours) sont effectués par 21% des panélistes. Nous sommes donc très loin de l'équirépartition. Dans ces 14% qui cumulent la moitié des distances et ces 21% qui effectuent la moitié des sorties, combien sont en fait les mêmes individus ? Cette question implique aussi de se demander si plus de distances entraîne plus de sorties ? Il s'avère que sur les 3621 panélistes, 2762 panélistes se retrouvent **dans les deux échantillons** soit 76%. Nous

pourrions ainsi dire que le cumul des sorties sur l'année entraîne 7 fois sur 10 un cumul des distances sur l'année (cf. la relation mis en lumière dans le paragraphe 3.4 de la première partie).

2.2.2 Distances et nombre de séjours annuels moyens par panéliste avec contrainte temporelle

Le deuxième graphique présente cette fois la concentration des distances parcourues pour un temps de sortie annuel identique. Le temps de sortie est bloqué sur la valeur modale de l'échantillon soit 21 jours en moyenne par an soit trois semaines. L'intérêt est d'observer si la concentration des distances et des sorties est plus ou moins forte selon que le temps de sortie est bloqué ou non (pour une durée de sortie annuelle donnée). L'échantillon utilisé est logiquement plus restreint car il porte sur 135 panélistes et 644 séjours.

Figure 63. _ Concentration des kilométrages parcourus pour un temps de sortie annuel identique à 21 jours (135 panélistes, 644 séjours)



Tout d'abord, et conformément à notre hypothèse, il est logique de retrouver une équirépartition sur le temps de sortie, tel que 50% des panélistes sélectionnés concentrent 50% des jours passés en dehors du domicile.

Les concentrations sur les séjours et les distances sont logiquement moins fortes lorsque le temps de sortie est bloqué sans pour autant être négligeables : 30% des panélistes effectuent 50% de la masse totale des séjours, contre 21% lorsque le temps de sortie annuel n'est pas bloqué. En outre, 19% des panélistes se répartissent 50% de la masse totale des distances contre 14% lorsque le temps de sortie annuel n'est pas bloqué. Là encore, le cumul des séjours et des distances va de pair : 95 panélistes sur 135 se retrouvent dans les deux groupes, ce qui porte également à 7/10 la probabilité de cumuler en

moyenne des sorties nombreuses avec des distances importantes sur l'année. Nous retrouvons la même probabilité de cumul sur deux échantillons de tailles différentes.

Cette deuxième étape dans la réflexion montre que le temps annuel de sortie n'est pas le seul déterminant de la hausse des distances sur l'année puisqu'il y a bien une concentration des distances sur les panélistes pour un temps de sortie cumulé annuel constant. A un premier niveau d'explication, ceci s'explique par le type de mode utilisé (notamment l'avion) et le degré de fractionnement du temps de congé annuel. A un deuxième niveau, le choix d'une vitesse et de la durée des séjours s'expliquent par des variables de type sociologique, économiques, démographiques, spatiales.

Pour conclure ce paragraphe sur les indices de concentration, le tableau ci-dessous récapitule les principales informations à retenir :

Tableau 8 . _ Récapitulatif des indices de concentration

N	% des panélistes		moyennes annuelles par panéliste		% de cumul
8049 panélistes pour 62029 séjours	21% des panélistes	cumulent	50% des séjours	←	2762 panélistes sur 3621 sont présents dans les deux groupes
	14% des panélistes	cumulent	50% des distances		
135 panélistes pour 644 séjours	30% des panélistes avec un temps de sortie moyen annuel de 21 jours	cumulent	50% des séjours	←	95 panélistes sur 135 sont présents dans les deux groupes
	19% des panélistes avec un temps de sortie moyen annuel de 21 jours	cumulent	50% des distances		

Source statistique: DT – TNSO ; SDT 1995-2001 ; calculs: V.Bagard

Il nous paraît essentiel de ne pas perdre de vue ces indices de concentration :

- d'une part pour les politiques publiques et les décisions d'aménagements d'infrastructures (aéroports notamment) qui souvent légitiment le financement des projets par le principe de l'intérêt général. A qui profitent essentiellement ces infrastructures à court et moyen terme ? A qui profiteront-elles à plus long terme, et quelles en seront les conséquences sur la mobilité de loisir ? Là encore, la hausse générale des revenus n'infléchit pas de la même manière les structures de consommation. Il n'est pas dit que des revenus plus élevés poussent systématiquement à des déplacements aériens plus fréquents.

- d'autre part pour les exercices prospectifs qui devront tenter de mieux comprendre comment évolueront les comportements de la minorité des individus qui réalisent plus de la moitié de la masse des distances. En outre, il faut aussi ajouter que ces indices sont peut être sous-estimés si le panel S.D.T sous-représente les individus les plus actifs.

Conclusion du chapitre 1

Ce premier chapitre a posé les fondements méthodologiques de notre étude statistique et précisé l'usage qui sera fait de la base SDT – TNSO sur les années disponibles 1995 –2001.

L'étude concerne donc les populations actives de 18 à 65 ans se déplaçant en mono-séjour pour une durée minimum d'une nuitée. Les motivations des déplacements sont au nombre de sept et concernent les loisirs au sens large : famille, vacances, loisirs, tourisme, manifestations culturelles et sportives, visite à des amis, étape sur le chemin des vacances. Le traitement détaillé de l'échantillon a permis de dépister les panélistes multi-profils. De ce fait, la constitution d'une base longitudinale a été possible pour les panélistes ayant répondu sur une durée supérieure à l'année. Disposer de variables « moyennées annuelles » permettra de contrôler les valeurs extrêmes et de fiabiliser les traitements. Ce premier chapitre a également donné les premiers éléments statistiques de cadrage concernant la distribution et la moyenne des temps annuel de sortie et des distances annuelles. Le temps de sortie annuel est en moyenne de 24 nuitées. La distance moyenne annuelle oscille entre 4000 et 5000 km pour l'ensemble des modes et pour les actifs de 18 à 65 ans. Enfin, le dernier point a mis en exergue l'importance des phénomènes de concentration en matière de mobilité touristique. Les courbes de Lorenz pointent en effet les inégalités fortes en matière d'accès aux vacances et à la mobilité touristique. A travers ces indices, ce sont toutes les différences et les inégalités dans la pratique des vacances qui ressortent.

A cette première approche macroscopique qui présente les variables constituant les deux axes de notre grille de lecture, succède une approche plus microscopique, à notre sens microéconomique, dont l'objectif est de mettre en lumière les différents équilibres observés pour les combinaisons $S_i(D_i, TS_i, \varepsilon_i)$ ainsi que les logiques qui sous-tendent ces équilibres.