

2021-2022

M2 Expertise Sciences des Populations

ANALYSE DES DONNÉES LONGITUDINALES

Arno Muller

arno.muller@ined.fr

(Ined - Service Méthodes Statistiques)

Introduction aux données longitudinales

- Quels outils ?
- Quelles questions ?
- Quelles données ?
 - *Prospectif vs rétrospectif*
 - *Exemples de dispositifs d'enquêtes*

Quels outils ?

- Analyses descriptives
 - *Durée moyenne de survenue d'un évènement*
- Régression logistique
 - *Facteurs expliquant la survenue d'un évènement ...*
- Classification des trajectoires
 - *Analyse de séquences*
- Analyse de survenue d'un évènement
 - *Modèles de durée (régression qui tient compte du temps)*
- Et d'autres qu'on ne verra pas dans ce cours : Modèle de panel (sauf si on a le temps), modèle Multi-états, Chaîne de Markov

Quelles questions ?

■ Fécondité, Nuptialité, Mise en couple

- *Analyse de séquence : Trajectoire familiale, de conjugalité, ...*
- *Analyse de durée : Temps avant d'avoir un (1^{er}) enfant, durée avant de se (re)marier, longévité d'une union ...*

■ Logement

- *Analyse de séquence : Trajectoire migratoires, d'habitations, ...*
- *Analyse de durée : Devenir propriétaire, Effectuer une mobilité résidentielle ...*

■ Emploi

- *Analyse de séquence : Trajectoire d'emploi, de salaire, ...*
- *Analyse de durée : Trouver un 1^{er} emploi, Analyse des changements d'emploi ...*

■ Mortalité, Morbidité

- *Analyse de séquence : Trajectoire de santé, ...*
- *Analyse de durée : Durée avant décès après diagnostic, survivre à un traitement, ...*

Données prospectives

- Différentes sources de données
 - *Panel (plutôt anglo-saxon)*
 - « Photos » à différents instants
 - *Suivi de cohorte (ELFE)*
 - *Registre de population*
 - *Système de suivi démographique*
- Avantages
 - *Qualité des données : observation « en direct »*
- Inconvénients
 - *Coût important*
 - *Délais pour une analyse longitudinale*
 - *Attrition*
 - *L'instrument de mesure doit être identique à chaque vague*
 - Comparabilité des questions

Exemples données prospectives

- Suivi de cohorte ELFE
 - 18 000 enfants né en 2011 (1/50)
 - Pendant 20ans
 - Sujets : Enfance, alimentation, éducation
- Observatoires de population en Afrique, au Sénégal et au Mali
 - Suivi de population dans une région rurale, à Mlomp
 - https://www.ined.fr/fichier/s_rubrique/19434/103.fr.pdf
- Panel d'étudiant.e.s (DEP)
 - Ministère, réussite scolaire
 - <https://www.education.gouv.fr/les-panels-d-eleves-de-la-depp-source-essentielle-pour-connaître-et-evaluer-le-systeme-educatif-4871>
- Les enquêtes « Générations » du Cereq
 - Suivi de générations de diplômé.e.s, insertion professionnelle.
 - <https://www.cereq.fr/enquetes-et-donnees/insertion-professionnelle-generation>
- Panel Study of Income Dynamics
 - Suivi de ménages sur très long termes
 - <https://psidonline.isr.umich.edu/default.aspx>



elfe

Etude Longitudinale
Française depuis
l'Enfance

<https://www.elfe-france.fr/>

Données rétrospectives

■ Principe

- *Echantillon d'individus interrogés une seule fois*
- *Recueil de biographies thématiques depuis une origine jusqu'au moment de l'enquête (Bloc de questions occurrence)*
- *Recueil d'informations complémentaires à la date de l'enquête.*

■ Avantages

- *Données immédiatement exploitables*
- *Quantitatif ou qualitatif*
- *Faible coût*

■ Inconvénients

- *Données qui peuvent être incomplètes, tronquées, fragmentaires, hétérogènes*
- *Données qui font appel à la mémoire de l'enquêté*
- *Questionnaire long, cohérence informatique complexe*

La fiche AGEVEN

Grille chronologique

- *Grille âge-événement-s*
- *Permet de resituer les types d'évènements de manière relative*
- *Repère(s) temporel(s) en ligne (âge, année)*
- *Les éléments biographiques sont situés en colonne, remplis de manière progressive.*
- *Exemples de domaines renseignés : Histoire professionnelle, familiale, résidentielle*

Références

- *Antoine, P., X. Bry, et P.D. Diouf. 1987. La fiche « Ageven » : un outil pour la collecte des données rétrospectives. Statistiques Canada, Techniques d'Enquêtes, 13, no 2 : 173-81.*
- *Vivier G, « Comment collecter des biographies ? De la fiche Ageven aux grilles biographiques, Principes de collecte et Innovations recentes », Acte des colloques de l'AIDELF, 2006*
- *GRAB, 1999, « Biographies d'enquêtes : bilan de 14 collectes biographiques », Paris, INED.*

Exemple AGEVEN (1)

LOGEMENTS									Périodes transitoires Logements et Remarques
Année	Âge	Famille	L1	Dépt. ou pays L2	Commune (arr ¹) L3	L4	L5	L6	
943	00		1	93	CLICHY SS BOIS	22	3	1	
944	01								
945	02								
946	03								
947	04								
948	05								
949	06								
950	07								
951	08								
952	09							4	Installation du confort sanitaire juste avant mon départ en pension
953	10		2	93	SAINT-DENIS	30	-	-	PENSIIONNAT
954	11	NF1							
955	12								
956	13								
957	14		3	75	PARIS 9 ^{ème}	12	4	4	
958	15	NF2							
959	16								
960	17		4	75	PARIS 9 ^{ème}	12	1	1	même immeuble que logmt précédent
961	18								
962	19		5	ALGERIE	-	30	-	-	CASERNE
963	20								
964	21	C1	6	93	LIVRY GARGAN	12	2	1	
965	22	M1							
966	23	NE1	7	93	LIVRY GARGAN	14	3	4	
967	24								
968	25								
969	26								
970	27								
971	28								
972	29								
973	30								
974	31	NE2	8	93	LIVRY GARGAN	22	3	4	
975	32								
976	33								
977	34								
978	35								

Source : Enquête Biographie & Entourage (BIOENT)

Exemple AGEVEN (2)

Section 1						Section 2					Section 3					Section 4	Section 5	Section 6	Section 7					
Year	Age	Family Events				Place of Residence		Co-residents			Age	Education/Occupation				Financial	Life Crisis	Health	Reprod. Health 1st Mens/ 1st Sex	HIV Test/HIV Infection			Age	Comments
		Parents & siblings	Marital life	Children	Happiness	District	Province	Relationship	Knowledge	Attitude		Education	Results	Main Occupation	Extra jobs	Status				Tests/ INF	Reasons	Results		
1	2	3	4	5	6	7.1	7.2	8.1	8.2	8.3	2	9.1	9.2	9.3	9.4	10	11	12	13	14.1	14.2	14.3	2	
25__	0										0												0	
25__	1										1												1	
25__	2										2												2	
25__	3										3												3	
25__	4										4												4	
25__	5										5												5	
25__	6										6												6	
25__	7										7												7	
25__	8										8												8	
25__	9										9												9	
25__	10										10												10	
25__	11										11												11	
25__	12										12												12	
25__	13										13												13	
25__	14										14												14	
25__	15										15												15	
25__	16										16												16	
25__	17										17												17	
25__	18										18												18	
25__	19										19												19	
25__	20										20												20	
25__	21										21												21	
25__	22										22												22	
25__	23										23												23	
25__	24										24												24	
25__	25										25												25	
25__	26										26												26	
25__	27										27												27	
25__	28										28												28	
25__	29										29												29	

Source : Le Cœur Sophie, Im-Em Wassana, Koetsawang Suporn et al., « Vulnérabilité et vie avec le VIH en Thaïlande : l'apport de l'approche biographique », *Population*, 2005/4 (Vol. 60), p. 551-568.

Exemple de dispositif : MAFE

- Questionnaire Biographique
 - *Fiche AGEVEN*
 - *Retrace les parcours migratoires :*
 - Pas unidirectionnelle
 - Essaye de capter les migrations de retours
 - 6 sites entre l'Europe et l'Afrique
 - *Données disponibles pour la recherche*
 - <https://mafeproject.site.ined.fr/>

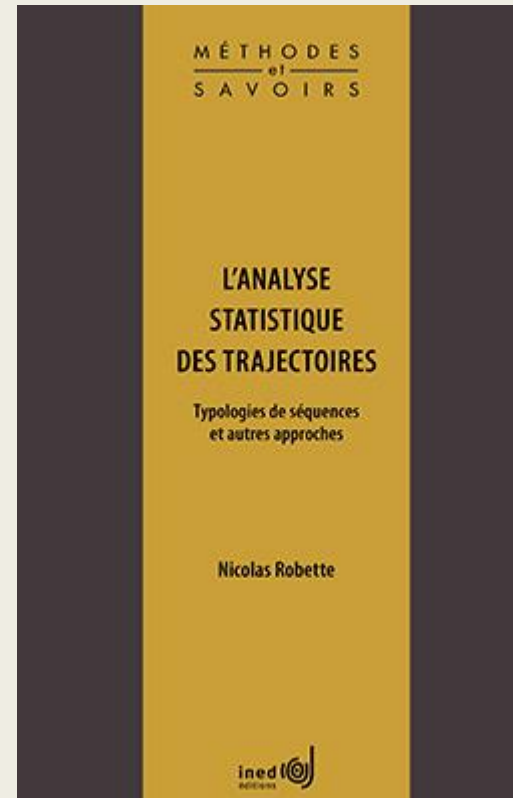
Les données administratives

- Entre prospective et rétrospective
- Données MASSIVES
- Possibilité de reconstruire des trajectoires
- Exemple :
 - Archive de l'aide sociale à l'enfance :
 - Frechon Isabelle, Robette Nicolas, « Les trajectoires de prise en charge par l'Aide sociale à l'enfance de jeunes ayant vécu un placement », Revue française des affaires sociales, p. 122-143. DOI : 10.3917/rfas.125.0122.
 - Assurance maladie et parcours de santé
 - Recensement pour les structures familiales

Truc que je sais pas où placer :
En histoire, on parle de prosopographie pour l'étude des biographies des membres d'un même milieu.
ex. : la bourgeoisie

Un bouquin top :

- L'analyse statistique des trajectoires, Typologies de séquences et autres approches
 - *Nicolas Robette*
 - *Ined editions*
 - *Collection : Méthodes et savoirs*



TP Titanic (suite)

- Séance du jour :
 - *Retour sur la séance dernière*
 - Questions (svp j'adore ça)
 - *Tableaux croisés*
 - χ^2 , Odds Ratio (brut et ajusté)
 - *Régression logistique*
 - But
 - Construire un modèle
 - *Choix des variables (logiques & fonctions automatisées)*
 - *Outil de comparaison*
 - AIC, R^2
 - types 3 analyse des effets
 - Interpréter le modèle

Pour la semaine prochaine

- Refaire les analyses et interprétation sur une autre base de données
 - *De préférences celles qu'on utilise pour les séquences.*