

# L'analyse de séquences pour étudier les trajectoires individuelles

Un état des lieux, quelques limites et des promesses

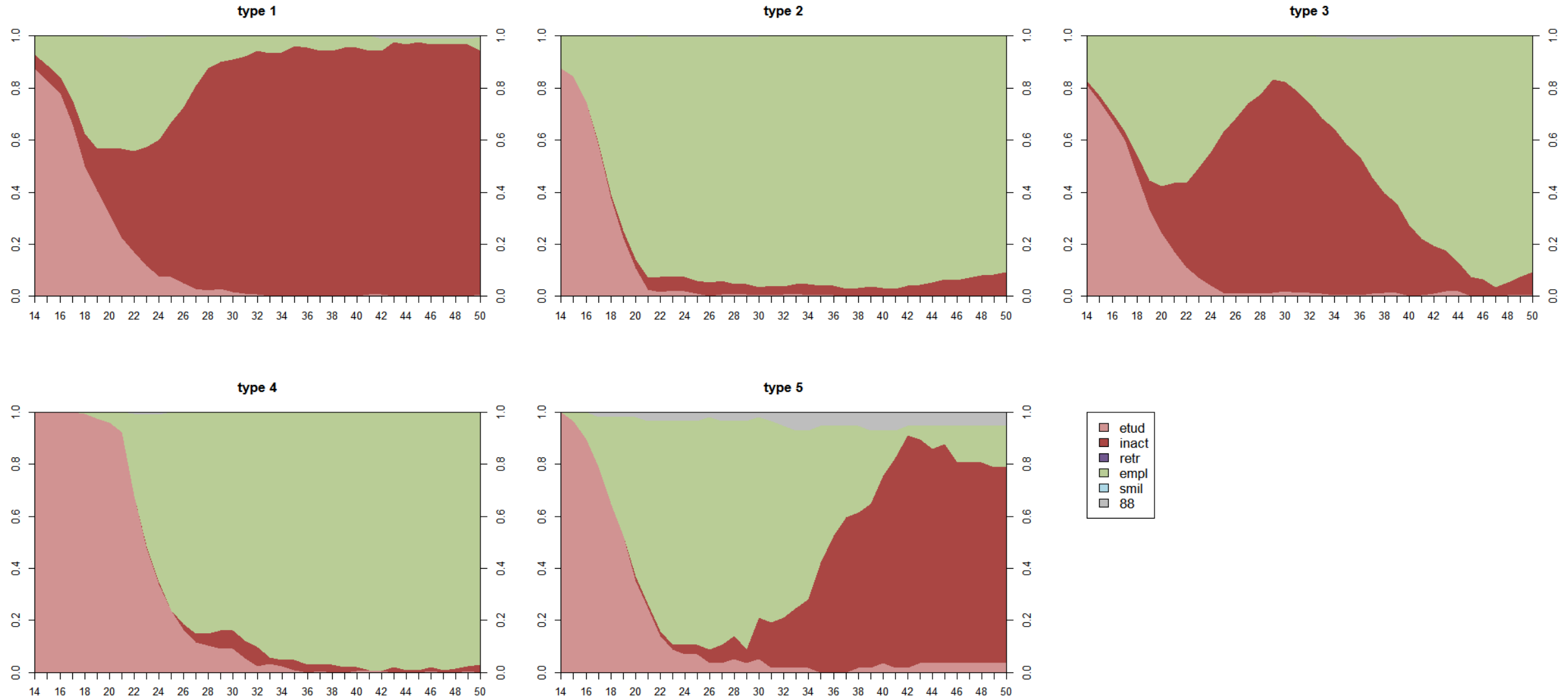
# Petit historique

- Introduction en SHS par Andrew Abbott dans les 80s
- Diffusion assez soutenue à partir de 2000, sur les carrières et les parcours de vie
- Disponibilité de données et de logiciels

# Quels usages ?

- Un usage dominant = les typologies de séquences
  - Résumer les données
  - Identifier des parcours-types

# Trajectoires d'activité des femmes franciliennes des générations 1934-1950



Source : *Biographies et entourage* (2001)

# Quels usages ?

- Un usage dominant = les typologies de séquences
  - Résumer les données
  - Identifier des parcours-types
- Mais utilisation souvent assez routinisée...
- ... et peu d'applications marquantes ?

# Une typologie est utile :

- 1) Si elle constitue un résumé des données qui sert à aller plus loin (étape de l'analyse).

# Trajectoires d'activité des femmes franciliennes des générations 1934-1950

<i>Type de trajectoire</i>	<i>1934-1944</i>	<i>1945-1950</i>
<b>Activité tôt</b>	47,7	49,9
<b>Activité tard</b>	12,2	19,5
<b>Interruption</b>	20,6	14,0
<b>Inactivité</b>	16,0	11,2
<b>Sortie d'activité</b>	3,5	5,3
<i><b>Total</b></i>	<i>100,0%</i>	<i>99,9%</i>

Source : Biographies et entourage (2001)

# Trajectoires d'activité des femmes franciliennes des générations 1934-1950

<i>Type de trajectoire</i>	<i>Jusqu'au secondaire</i>	<i>secondaire</i>	<i>Niveau bac</i>	<i>Études sup.</i>	<i>Total</i>
<b>Activité tôt</b>	8	58	20	15	100
<b>Activité tard</b>	3	8	17	71	100
<b>Interruption</b>	12	56	17	16	100
<b>Inactivité</b>	21	47	19	13	100
<b>Sortie d'activité</b>	4	46	23	28	100

Source : Biographies et entourage (2001)



# Une typologie est utile :

1) Si elle constitue un résumé des données qui sert à aller plus loin (étape de l'analyse).

Mais :

- on n'a pas toujours l'espace pour le faire
- la gamme d'outils utilisés est limitée
- résultats souvent sans surprise et/ou peu marqués

# Une typologie est utile :

- 2) Si on n'a pas de connaissance a priori, ou si des types de séquences « surprenants » émergent (effet de dévoilement)

# Une typologie est utile :

- 2) Si on n'a pas de connaissance a priori, ou si des types de séquences « surprenants » émergent (effet de dévoilement)
- 3) Si on souhaite *quantifier* précisément les types de séquences  
→ sans doute à éviter, car sensibilité aux paramètres et aux méthodes de classification

# Faire autre chose que SA / OMA ?

- On s'intéresse à des événements  
→ analyse de survie (EHA)...
- On a peu de points dans le temps (ou peu d'observations)  
→ « à la main », tris croisés...
- On s'intéresse à un aspect particulier des trajectoires  
→ indicateurs ad hoc
- Selon priorité à ordre / calendrier / durée  
→ paramétrage OMA, autres algorithmes

# Faire SA sans typologie

*(à partir de la matrice de distances)*

- ANOVA, MANOVA, arbres de régression
- Multidimensional scaling → scores, projection dans un espace factoriel
- Distance à une ou plusieurs séquence(s) de référence
- Distribution des distances par groupes d'individus

# Innovations récentes

- Séquences multidimensionnelles
  - MSA, GIMSA, ...
  - Association entre dimensions
- Articulation SA - EHA
  - Sequence-Analysis Multistate Model (patterns de sous-séquences comme variable dépendante)
  - Competing Trajectory Analysis (idem)
  - Sequence History Analysis (patterns de sous-séquences comme « time-varying covariate »)

# Conclusion provisoire

- Ne pas trop attendre de l'analyse de séquences
- Encourager des usages réflexifs et imaginatifs

# Références bibliographiques

- **Un manuel d'analyse de séquences:**

Robette N., 2011, *Explorer et décrire les parcours de vie: les typologies de trajectoires*, CEPED ("Les Clefs pour"), 86 p.

- **Historique et discussion sur les usages de l'analyse de séquences:**

Robette N., 2016, « Du prosélytisme à la sécularisation. Le processus de diffusion de l'Optimal Matching Analysis ». in D. Demazière, M. Jouvenet (eds), *Andrew Abbott et l'héritage de l'école de Chicago*, éditions de l'EHESS, vol. 2, p. 173-194.

Robette N., Bry X., Lelièvre E., 2015, « Positivism and big time fishing », *Sociological Methodology*, vol.45(1), p. 88-100.

- **Comparaisons d'algorithmes d'analyse de séquences:**

Robette N., Bry X., 2012, « Harpoon or bait? A comparison of various metrics to fish for life course patterns », *Bulletin de Méthodologie Sociologique*, 116(1), p. 5-24.

Studer, M. & Ritschard, G. (2016), "What Matters in Differences between Life Trajectories: A Comparative Review of Sequence Dissimilarity Measures", *Journal of the Royal Statistical Society, Series A*. Vol. 179(2), pp. 481-511.

- **Analyse de séquences multidimensionnelles:**

Robette N., Bry X., Lelièvre E., 2015, « A "Global Interdependence" Approach to Multidimensional Sequence Analysis », *Sociological Methodology*, vol.45(1), p.1-44

Piccarreta R., 2017. "Joint Sequence Analysis. Association and clustering." *Sociological Methods & Research*, vol. 46(2), p.252-287.

- **Analyse de séquences et event-history analysis:**

Studer M., Struffolino E., Fasang A. 2018. « Estimating the Relationship between Time-varying Covariates and Trajectories: The Sequence Analysis Multistate Model Procedure ». *Sociological Methodology*, p. 1-34.

Rossignon F., Studer M., Gauthier J.-A., & Le Goff J.-M. (2016). "Childhood co-residence structure and homeleaving". In G. Ritschard & M. Studer (Eds.). *Proceedings of the International conference on sequence analysis and related methods (LACOSA II)*, p.383-428).

Studer M., Liebroer A., Mooyaart J. 2018. "Understanding trends in family formation trajectories: An application of competing trajectories analysis". *Advances in Life Course Research*, vol. 36, p. 1-12.