

Variables dynamiques

Table of Contents

Facteur dynamique traitée de manière fixe	1
Estimation avec une variable dynamique	3
Modèle de Cox	3
Modèle à temps discret.....	5
Quelques remarques sur les problèmes de causalité avec les variables dynamiques.....	5

Cette section sera principalement traitée par l'exemple.

- Dans un modèle de durée, une variable dynamique peut-être appréhendée comme une interaction entre la durée et une variable.
- Pour un modèle de Cox, l'hypothèse de risque proportionnel ne peut donc pas être testée.
- Ne pas tenir compte du caractère dynamique d'une dimension peut conduire à des interprétations erronées.
- La façon de modéliser les dimensions dynamiques en analyse des durées peut conduire à des biais de causalité, en particulier dans sciences sociales, en omettant les effets d'anticipation. C'est une situation classique avec des covariables dynamiques de type discrètes. Les techniques standards ne peuvent modéliser que des effets d'adaptation (la cause - observée - précède l'effet).

Cette partie sera principalement traité par l'exemple, et on s'intéressera uniquement aux variables de type discrète.

Facteur dynamique traitée de manière fixe

On reprend l'exemple sur les transplantations cardiaques.

On a dans la base 2 variables: une variable binaire pour savoir si l'individu à été greffé , **transplant**, et une variable continue tronquée donnant la durée en jour jusqu'à la greffe greffe (0 si pas de greffe), **wait**.

On va estimer le modèle (de Cox) avec la variable fixe transplant.

Cox regression -- Efron method for ties

No. of subjects =	103	Number of obs =	103
No. of failures =	75		
Time at risk =	31938		
Log likelihood =	-273.21499	LR chi2(4) =	49.81
		Prob > chi2 =	0.0000

_t	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
year	-0.0909	0.0659	-1.38	0.168	-0.2201	0.0383
age	0.0579	0.0147	3.95	0.000	0.0292	0.0866
1.surgery	-0.6547	0.4475	-1.46	0.143	-1.5318	0.2224
1.transplant	-1.6484	0.2792	-5.90	0.000	-2.1957	-1.1011

Interprétation?

Au niveau des données le modèle à été estimé, pour une personne greffée, à partir de ce mapping:

id	year	age	surgery	transp~t	wait	_d	_t	_t0
10	68	42	0	1	12	0	1	0
10	68	42	0	1	12	0	2	1
10	68	42	0	1	12	0	3	2
10	68	42	0	1	12	0	5	3
10	68	42	0	1	12	0	5.0999999	5
10	68	42	0	1	12	0	6	5.0999999
10	68	42	0	1	12	0	8	6
10	68	42	0	1	12	0	9	8
10	68	42	0	1	12	0	12	9
10	68	42	0	1	12	0	16	12
10	68	42	0	1	12	0	17	16
10	68	42	0	1	12	0	18	17
10	68	42	0	1	12	0	21	18
10	68	42	0	1	12	0	28	21
10	68	42	0	1	12	0	30	28
10	68	42	0	1	12	0	32	30
10	68	42	0	1	12	0	35	32
10	68	42	0	1	12	0	36	35
10	68	42	0	1	12	0	37	36
10	68	42	0	1	12	0	39	37
10	68	42	0	1	12	0	40	39
10	68	42	0	1	12	0	43	40
10	68	42	0	1	12	0	45	43
10	68	42	0	1	12	0	50	45

10	68	42	0	1	12	0	51	50
10	68	42	0	1	12	0	53	51
10	68	42	0	1	12	1	58	53

Problème: une personne est codée transplantée avant le jour de la transplantation. L'effet causal est donc mal mesuré si sa dimension temporelle a été ignorée.

Estimation avec une variable dynamique

Il convient donc de modifier l'information avec le délai d'attente jusqu'à la greffe:

$tvc = transplant$, si $transplant = 1$ et $t < wait$ alors $tvc = 0$.

Modèle de Cox

id	year	age	surgery	tvc	wait	_d	_t	_t0
10	68	42	0	0	12	0	1	0
10	68	42	0	0	12	0	2	1
10	68	42	0	0	12	0	3	2
10	68	42	0	0	12	0	5	3
10	68	42	0	0	12	0	5.0999999	5
10	68	42	0	0	12	0	6	5.0999999
10	68	42	0	0	12	0	8	6
10	68	42	0	0	12	0	9	8
10	68	42	0	1	12	0	12	9
10	68	42	0	1	12	0	16	12
10	68	42	0	1	12	0	17	16
10	68	42	0	1	12	0	18	17
10	68	42	0	1	12	0	21	18
10	68	42	0	1	12	0	28	21
10	68	42	0	1	12	0	30	28
10	68	42	0	1	12	0	32	30
10	68	42	0	1	12	0	35	32
10	68	42	0	1	12	0	36	35
10	68	42	0	1	12	0	37	36
10	68	42	0	1	12	0	39	37
10	68	42	0	1	12	0	40	39
10	68	42	0	1	12	0	43	40
10	68	42	0	1	12	0	45	43
10	68	42	0	1	12	0	50	45
10	68	42	0	1	12	0	51	50
10	68	42	0	1	12	0	53	51
10	68	42	0	1	12	1	58	53

Estimation du modèle avec la variable dynamique (Cox)

Cox regression -- Efron method for ties

No. of subjects =	103	Number of obs =	3,668
No. of failures =	75		
Time at risk =	31938.1		
Log likelihood =	-289.27058	LR chi2(4) =	17.70
		Prob > chi2 =	0.0014

_t	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
year	-.1202612	.0673414	-1.79	0.074	-.252248	.0117256
age	.0304498	.0138998	2.19	0.028	.0032068	.0576929
1.surgery	-.9829386	.4365524	-2.25	0.024	-1.838566	-.1273116
1.tvc	-.0826682	.3047751	-0.27	0.786	-.6800165	.51468

Interprétation?

Solutions avec les logiciels :

- R, Stata, Python: la base doit être transformée en format long aux temps d'évènement (survsplit avec R, stsplot avec Stata) avant la création de la variable dynamique.
- Sas: la base n'est pas modifiée et la création de la TVC est faite "en aveugle" dans la procédure phreg

Modèle à temps discret

Même principe pour la construction de la variable dynamique

+-----+							
	id	year	age	surgery	tvc	mwait	t

	13	68	54	0	0	2	1
	13	68	54	0	1	2	2
	13	68	54	0	1	2	3
+-----+							
Logistic regression				Number of obs = 1,127			
				LR chi2(7) = 90.73			
				Prob > chi2 = 0.0000			
Log likelihood = -230.32152				Pseudo R2 = 0.1645			

	e	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	

	t	-.365048	.0915105	-3.99	0.000	-.5444052	-.1856907
	t2	.0139226	.0053256	2.61	0.009	.0034846	.0243606
	t3	-.000162	.0000815	-1.99	0.047	-.0003217	-2.27e-06
	year	-.1324928	.0737516	-1.80	0.072	-.2770433	.0120577
	age	.033829	.0149503	2.26	0.024	.004527	.0631311
	surgery	-1.007795	.4490177	-2.24	0.025	-1.887854	-.1277365
	tvc	-.0543011	.3114096	-0.17	0.862	-.6646528	.5560505
	_cons	7.060589	5.305278	1.33	0.183	-3.337566	17.45874

Quelques remarques sur les problèmes de causalité avec les variables dynamiques

- Rappel: la cause précède toujours l'effet.
- Lorsque l'évènement étudié n'est pas intraséquement de type absorbant (décès), la "cause" peut se manifester après la survenue de l'évènement étudié.
- Les modèles de durée standards ne peuvent pas gérer ces situations car l'observation sort du risque après la survenue de l'évènement.
- Logique d'adaptation: la "cause" identifiée est mesurée avant l'évènement étudié **[OK]**.
- Logique d'anticipation: la "cause" identifiée est mesurée après l'évènement étudié **[Problème]**. L'origine causale est bien antérieure à l'évènement, mais elle n'est pas observable.