A Regression Model For Recurrent Events With Distribution Free Correlation Structure

J. Pénichoux(1), A. Latouche(2), T. Moreau(1)

(1) INSERM U780
(2) Université de Versailles, EA2506

ISCB - 2009 - Prague

Funding :
2008 doctoral fellowship from AXA research fund

Background

Each subject i ($i = 1, ..., N$) can experience K_i repeated occurrences of the same type of event during his follow-up.

2 issues

- Estimation of covariates effect on the risk of occurrence
- Assessment of the strength of correlation between events

Motivation : On-going cohort study on the risk of recurrent infections in patients with cystic fibrosis (364 patients). Being aware of the presence of correlation can influence therapeutical decisions.

Worked example : infection times for patients with chronic granulotomous disease (CGD)

Handling of recurrent events (1)

Extensions of Cox proportional hazards model

Timescale for a conditional model

- Gap time : $\tilde{T}_{i(k)} = T_{i(k)} - T_{i(k-1)}$
- Calendar time : $T_{i(k)}$

with $T_{i(1)}, \ldots, T_{i(K)}$ the successive event times for individual i

Our framework : Gap time

Handling of correlation :

- Stratification of the baseline hazard on the rank of event to account for event dependence
- Random effect (frailty) to account for heterogeneity across individuals

[1] Kelly, P.J. and Lim, L.L.(2000) *Stat Med*

Handling of recurrent events (2)

The conditional frailty model for gap time

$$\lambda_{ik}(t|T_{i(k-1)} = t_{i(k-1)}, X_i) = u_i \lambda_{0k}(\tilde{t}) exp({}^t\beta X_i)$$

where $\tilde{t} = t - t_{i(k-1)}$

u_i : random effect, shared for all intervals of individual i, with expectation 1 and variance φ (usually gamma).

⇒ Accounts for heterogeneity across individuals

λ_{0k} : interval-specific hazard to account for event dependence [1].

Some issues :

- The estimation of the variance of the random effect is complex → assessment and estimation of correlation might not be straightforward.
- The history of an individual is not used to evaluate his current risk.

[2] Box-Steffensmeier, J. and De Boef, S. (2006) *Stat Med*

Proposed approach (1)

Hazard of event k for individual i

$$\lambda_{ik}(t|T_{i(k-1)} = t_{i(k-1)}, X_i) = \lambda_{0k}(\tilde{t}) \exp\left[\theta \underbrace{(1-\hat{\Lambda}_{k-1}(\tilde{t}_{i(k-1)}))}_{Z_{ik}} + {}^t\beta X_i\right]$$

with $\tilde{t} = t - t_{i(k-1)}$,
Z_{ik} interval-specific covariate ($Z_{ik} = 0$ if $k = 1$),
$\hat{\Lambda}_{k-1}$: cumulative hazard estimated with Breslow's formula computed on the $(k-1)$ intervals,
X_i a vector of covariates for individual i

Estimation of parameters θ,β : simple partial likelihood maximization.

Proposed approach (2)

Interpretation of Z_{ik} : $(Z_{ik} = 1 - \hat{\Lambda}_{k-1}(\tilde{t}_{i(k-1)}))$
Difference between

- 1 : the observed number of events during the previous interval $[t_{i(k-2)}, t_{i(k-1)}]$
- $\hat{\Lambda}_{(k-1)}(\tilde{t}_{i(k-1)})$: the expected number of events in this interval estimated under a Poisson process with intensity λ_{k-1}

Interpretation of parameter θ
The proposed approach for two successive events is an approximation of a shared-gamma frailty model close to $\varphi = 0$ (ie close to the independence), with $\theta \approx \varphi$[1]
To keep the comparability with a frailty, $\theta > 0$ in the proposed approach ($\hat{\theta} = \max(0, \hat{\theta}_{ml})$).

[3] LEFFONDRÉ, K., LELLOUCH, J.,COM-NOUGUÉ, C. AND MOREAU, T. (2001) *Commun. Stat., Theory Methods*

Simulation outline

In the sequel, we'll consider :

2 events per subject, no censoring

Outline

- Estimation of covariates effects using simulations from gamma and lognormal frailty models
- Estimation of the variance of the frailty using simulations from a gamma frailty model
- Test of independence, type I error and power using simulations from gamma frailty models
- Example : CGD data

Models compared

- "PWP-GT" (Usual stratified Cox)
- Conditional Gamma Frailty
- Proposed approach

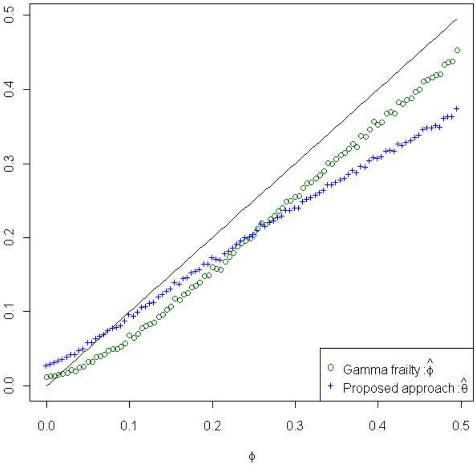
Estimation of covariates effect

500 repetitions, $N = 100$

			$\beta = 0.7$ ($RR \approx 2$)		$\beta = 1.1$ ($RR \approx 3$)	
			$\overline{\hat{\beta}}$	$\overline{\hat{\sigma}_\beta}$	$\overline{\hat{\beta}}$	$\overline{\hat{\sigma}_\beta}$
Gamma	$\varphi = 0.2$	PWP-GT	0.596	0.150	0.939	0.157
		CGFM*	**0.662**	0.174	**1.050**	0.182
		PA*	0.605	0.150	0.955	0.157
	$\varphi = 0.6$	PWP-GT	0.446	0.147	0.730	0.150
		CGFM*	**0.664**	0.231	**1.071**	0.235
		PA*	0.484	0.148	0.793	0.153
Lognormal	$\varphi = 0.2$	PWP-GT	0.614	0.150	0.979	0.158
		CGFM*	**0.668**	0.168	**1.056**	0.176
		PA*	0.623	0.151	0.990	0.158
	$\varphi = 0.6$	PWP-GT	0.558	0.149	0.868	0.154
		CGFM*	**0.682**	0.198	**1.053**	0.200
		PA*	0.576	0.150	0.893	0.155

* CGFM : Conditional gamma frailty model
PA : Proposed Approach

Estimation of the variance of the frailty



Simulations :
Gamma shared frailty models with variance φ
1000 repetitions for each value of φ
$N = 200$ for each dataset

Before $\varphi = 0.25$, $\hat{\theta}$ estimated by the proposed approach is closer to the variance of the frailty.
After $\varphi = 0.25$, $\hat{\varphi}$ estimated by a gamma shared frailty model is closer.

Test of independence and type I error

1000 repetitions

		Nominal level for type I error			$\overline{\hat{\varphi}}$	$\overline{\hat{\theta}}$
		5%	1%	1‰		
N=50	CGFM*	1.7%	0.5%	0.1%	0.020	.
	PA*	3.7%	0.4%	0.1%	.	0.055
N=100	CGFM*	1.7%	0.4%	0.1%	0.015	.
	PA*	3.9%	0.6%	0.0%	.	0.039
N=200	CGFM*	1.9%	0.6%	0.1%	0.012	.
	PA*	4.6%	0.9%	0.0%	.	0.028

* CGFM : Conditional gamma frailty model
PA : Proposed approach

Test of independence and power

1000 repetitions

		Power (%)			
		CGFM*	PA*	CGFM* ($\overline{\hat{\varphi}}$)	PA* ($\overline{\hat{\theta}}$)
$\varphi = 0.2$	N=50	17.9	**28.0**	0.11	**0.19**
	N=100	34.4	**47.3**	0.14	**0.17**
	N=200	60.5	**73.8**	0.16	**0.17**
$\varphi = 0.4$	N=50	52.7	**64.0**	0.28	**0.33**
	N=100	81.8	**88.6**	0.32	0.32
	N=200	97.7	**98.7**	**0.36**	0.31
$\varphi = 0.6$	N=50	74.9	**83.0**	0.43	**0.45**
	N=100	97.0	**97.6**	**0.50**	0.44
	N=200	100	100	**0.54**	0.43

* CGFM : Conditional gamma frailty model
PA : Proposed Approach

Application to the CGD data

CGD data :
Data from a placebo controlled trial examining the effect of gamma interferon in chronic granulotomous disease (CGD), which manifests in recurrent infections. 135 patients included

Model	$\hat{\beta}_{treat}$	p-value	$\hat{\varphi}$	$\hat{\theta}$
PWP-GT	-0.8760	0.0016	.	.
Conditional shared gamma fraitly model	-0.8760	0.0016	$0,00$	.
Proposed approach	-0.8562	0.0023	.	0*

* $\hat{\theta}_{ml} = -0.5434$

Conclusion

- Covariates effect estimation
  - GF : robust estimation of covariates effect
- Assessment of correlation and test of independence
  - Correct Type I error for both approaches
  - Greater power for the proposed approach
  - Better estimation of the variance of a frailty for small datasets and small values of φ

Perpectives

- Easy to develop for non stratified analysis and calendar timescale
- Extending the analysis to non-shared frailties which would combine the two sources of correlation

Bibliography

[1] P. Kelly and L. Lim, “Survival analysis for recurrent event data : an application to childhood infectious diseases,” *Stat Med*, vol. 19, pp. 13–33, Jan 2000.

[2] J. Box-Steffensmeier and S. De Boef, “Repeated events survival models : the conditional frailty model,” *Stat Med*, vol. 25, pp. 3518–3533, Oct 2006.

[3] K. Leffondré, J. Lellouch, C. Com-Nougué, and T. Moreau, “Optimality of nonparametric tests for association between survival time and continuous covariates.,” *Commun. Stat., Theory Methods*, vol. 30, no. 5, pp. 913–929, 2001.

[4] T. M. Therneau and P. M. Grambsch, *Modeling survival data : extending the cox model.* Statistics for biology and health, Springer, 2000.

Partial likelihood for our approach

$$L = \prod_{i=1}^{n} \prod_{k=1}^{n_i} \frac{\exp\left({}^t\beta X_i + \theta\{1 - \hat{\Lambda}_{k-1}(t_{i(k-1)})\}\right)}{\sum_{l=1}^{n} Y_l(T_{ik}) \exp\left({}^t\beta X_l + \theta\{1 - \hat{\Lambda}_{k-1}(t_{l(k-1)})\}\right)}$$

where

$$\hat{\Lambda}_{k-1}(t) = \sum_{T_{j(k-1)} \leq t} \frac{\delta_j}{\sum_{l | Y_l(T_{j(k-1)})=1} \exp({}^t\hat{\beta}^{\underline{k-1}} X_l(T_{j(k-1)}))} \exp({}^t\hat{\beta}^{\underline{k-1}} X)$$

Example of results with censorship and any number of events

Modèle	$\overline{\hat{\beta}}$	$\overline{\hat{Var}(\hat{\beta})}$
Stratified Cox model	-0.9097	0.0610
Stratified shared gamma frailty model	-0.9463	0.0657
Proposed approach	-0.9271	0.0618

TAB.: Simulations under a gamma frailty, $n = 50$, $\lambda_0 = 1$, $\beta = -1$