



Universiteit  
Leiden  
The Netherlands

## Statistical modelling of time-varying covariates for survival data

Spreafico, M.

### Citation

Spreafico, M. (2022, October 12). *Statistical modelling of time-varying covariates for survival data*. Retrieved from <https://hdl.handle.net/1887/3479768>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/3479768>

**Note:** To cite this publication please use the final published version (if applicable).

# Sommario

Nella ricerca clinica, le *covariate tempo-dipendenti* sono di grande interesse in quanto racchiudono informazioni sull'evoluzione di elementi dinamici che possono avere un impatto sullo stato di salute del paziente e sulla progressione della malattia. In letteratura, queste caratteristiche vengono solitamente analizzate utilizzando un approccio di “baseline” o “costante a tratti”, perdendo così il potenziale contenuto informativo che potrebbero fornire se l'associazione tra i dati tempo-dipendenti e quelli di *tempo all'evento* venisse catturata in maniera appropriata. Data la complessità dei fenomeni, è necessario formulare adeguatamente i modelli per l'evoluzione della malattia e caratterizzare in un contesto analitico avanzato la relazione tra di essa e la natura dinamica delle caratteristiche tempo-varianti. Questi aspetti sono raramente affrontati e possono fornire nuovi spunti per la ricerca medica, rappresentando una sfida sia di rilevanza clinica che di interesse statistico.

Questa tesi si concentra sullo sviluppo di metodi matematici e statistici per rappresentare appropriatamente i processi tempo-dipendenti ed includerli in un contesto di analisi di tempo all'evento, mediante modelli di sopravvivenza di tipo “Cox”. Lo scopo principale è quello di arricchire le informazioni disponibili per studiare la sopravvivenza con caratteristiche rilevanti legate alle covariate tempo-varianti di interesse. Questo lavoro affronta quindi le difficoltà sia di (i) sviluppare adeguate caratterizzazioni dinamiche dei processi oggetto di studio (ovvero il problema di *rappresentazione*), sia di (ii) identificare e quantificare l'associazione tra i processi tempo-dipendenti e la sopravvivenza dei pazienti (ovvero il problema della *modellazione* del tempo all'evento). In entrambi i casi, la sfida principale consiste nel trattare fonti di dati complesse, sia tenendo conto della natura dei processi e dei loro aspetti peculiari (come le dinamiche temporali, i livelli categorici o gli eventi ricorrenti), sia bilanciando in modo adeguato l'interpretabilità clinica e la formulazione matematica.

Lungo il corso della tesi vengono proposti approcci diversificati a seconda del contesto dello studio. In termini di *rappresentazione*, grazie all'integrazione di fonti di dati complesse e all'utilizzo di tecniche di analisi dei dati funzionali, vengono elaborate nuove covariate tempo-dipendenti (longitudinali o funzionali) in grado di riflettere l'evoluzione dei processi nel tempo. Inoltre, vengono introdotte nuove metodologie statistiche per estrarre dai dati informazioni aggiuntive, impiegando, tra gli altri, la teoria dei processi stocastici, i modelli di Markov latenti e l'analisi dei dati composizionali. In termini di *modellazione* del tempo all'evento, i modelli di regressione di tipo “Cox” vengono estesi al fine di includere le covariate dinamiche nell'analisi di sopravvivenza. Le tecniche di riduzione della

dimensionalità ed i modelli proposti (“joint” o “marginal structural”) sono implementati a secondo del tipo di studio e di dati a disposizione.

Oltre ad avere un impatto sulla comunità dei ricercatori in matematica e statistica, questo lavoro mira a fornire strumenti utili per supportare i medici nella pratica clinica. Tutti i temi di ricerca che vengono affrontati sono motivati da domande cliniche specifiche riguardanti i trattamenti personalizzati in pazienti cardiologici ed oncologici. La parte I di questa tesi si concentra sullo studio dei processi di re-ospedalizzazione, consumo ed aderenza ai farmaci in pazienti con scompenso cardiaco. La parte II tratta lo studio degli aspetti tempo-varianti della chemioterapia, come le modifiche del dosaggio, i biomarcatori o l’accumulo di tossicità, in pazienti affetti da osteosarcoma.

In conclusione, lo sviluppo di nuove metodologie per dati di tempo all’evento che includano informazioni dinamiche può rappresentare un significativo passo avanti nella definizione di nuovi strumenti di monitoraggio personalizzati, che potrebbero poi essere applicati allo studio di diverse patologie caratterizzate da fonti di dati complesse.