Bayesian Joint Modelling of Longitudinal and Time-to-Event Data

정희철

Department of Statistics, SKKU

2022/12/07

INDEX

1

논문 소개

1 논문 소개

Bayesian Joint Modelling of Longitudinal and Time-to-Event Data

여러 Bayesian Joint Modelling 기법과 분석 과정에 사용된 구조들와 가정들을 분석 상황에 따라 다양하게 소개하는 리뷰논문

논문을 리뷰한 논문을 리뷰하는 발표…

Bayesian joint modelling of longitudinal and time to event data: a methodological review

Maha Alsefri[1,2*], Maria Sudell[1], Marta García-Fiñana[1] and Ruwanthi Kolamunnage-Dona[1]

Abstract

Background: In clinical research, there is an increasing interest in joint modelling of longitudinal and time-to-event data, since it reduces bias in parameter estimation and increases the efficiency of statistical inference. Inference and prediction from frequentist approaches of joint models have been extensively reviewed, and due to the recent popularity of data-driven Bayesian approaches, a review on current Bayesian estimation of joint model is useful to draw recommendations for future researches.

Methods: We have undertaken a comprehensive review on Bayesian univariate and multivariate joint models. We focused on type of outcomes, model assumptions, association structure, estimation algorithm, dynamic prediction and software implementation.

Results: A total of 89 articles have been identified, consisting of 75 methodological and 14 applied articles. The most common approach to model the longitudinal and time-to-event outcomes jointly included linear mixed effect models with proportional hazards. A random effect association structure was generally used for linking the two sub-models. Markov Chain Monte Carlo (MCMC) algorithms were commonly used (93% articles) to estimate the model parameters. Only six articles were primarily focused on dynamic predictions for longitudinal or event-time outcomes.

Conclusion: Methodologies for a wide variety of data types have been proposed; however the research is limited if the association between the two outcomes changes over time, and there is also lack of methods to determine the association structure in the absence of clinical background knowledge. Joint modelling has been proved to be beneficial in producing more accurate dynamic prediction; however, there is a lack of sufficient tools to validate the prediction.

Keywords: Joint models, Longitudinal outcomes, Time-to-event, Dynamic prediction, Bayesian estimation

Background

Over the last decade, there has been an increasing interest in joint models for longitudinal and time-to-event outcome data, especially in medical research, due to their ability to predict individual-level patients' risks. A joint model consists of two linked sub-models. The relationship between the longitudinal and time-to-event outcomes is represented by an association structure, a function that links the longitudinal and time-to-event sub-models. A commonly used longitudinal sub-model is the linear mixed effect model, and the time-to-event sub-model is often the Cox proportional hazards model.

Joint modelling reduces the biases of parameter estimates by accounting for the association between the longitudinal and time-to-event data [1]. In clinical trials, this leads to more efficient estimation of the treatment effect on both time-to-event and longitudinal outcomes. It also quantifies the strength of the association between

* Correspondence: m.alsefri@liverpool.ac.uk
[1]Department of Health Data Science, Institute of Population Health, University of Liverpool, L69 3GL, Liverpool, UK
[2]Department of Statistics, University of Jeddah, Jeddah, Saudi Arabia

- Maha Alsefri

- Department of Health Data Science, Institute of Population Health, University of Liverpool, UK

- Department of Statistics, University of Jeddah, Saudi Arabia

Bayesian Joint Modelling of Longitudinal and Time-to-Event Data

Joint Modelling

- 경시적자료와 생존자료를 결합하여 두 자료간 의존성과 관계를 파악하는 분석기법
- 바이오 분야에서 중용

Time-to-Event Data 생존자료

- 어떠한 사건이 일어났는가 와 해당 사건이 언제 일어났는지에 대한 정보가 담긴 자료

Frequentist 방법론에 비해 Bayesian 방법론에 대한 리뷰논문이 부족

Bayesian Joint Modelling of Longitudinal and Time-to-Event Data

Joint Modelling

- 경시적자료와 생존자료를 결합하여 두 자료간 의존성과 관계를 파악하는 분석기법
- 바이오 분야에서 중용

Time-to-Event Data 생존자료

- 어떠한 사건이 일어났는가 와 해당 사건이 언제 일어났는지에 대한 정보가 담긴 자료

Frequentist 방법론에 비해 Bayesian 방법론에 대한 리뷰논문이 부족

1 논문 소개

Bayesian Joint Modelling of Longitudinal and Time-to-Event Data

Joint Modelling

- 경시적자료와 생존자료를 결합하여 두 자료간 의존성과 관계를 파악하는 분석기법
- 바이오 분야에서 중용

Time-to-Event Data 생존자료

- 어떠한 사건이 일어났는가 와 해당 사건이 언제 일어났는지에 대한 정보가 담긴 자료

Frequentist 방법론에 비해 Bayesian 방법론에 대한 리뷰논문이 부족

1 논문 소개

Bayesian Joint Modelling

Bayesian Joint Modelling of Longitudinal and Time-to-Event Data

- 분석 상황을 상당히 세부적으로 분류하고, 각 상황마다 매우 다양한 방법들을 소개하여, 이에 대해 하나씩 이론적 설명을 하기에는 불가능
- 논문에 충실하게 단순히 각 케이스별로 사용된 방법들을 나열하기에는 리뷰 발표의 성격과 맞지 않는다고 판단
- 케이스별로 사용된 방법들을 소개하되, 이 중 흥미로웠던 방법들에 대해 이론적 설명을 할 예정

Joint Modelling

- 경시적자료와 생존자료를 결합하여 두 자료간 의존성과 관계를 파악하는 분석기법
- 바이오 분야에서 중요

Time-to-Event Data 생존자료

- 어떠한 사건이 일어났는가 와 해당 사건이 언제 일어났는지에 대한 정보가 담긴 자료

Frequentist 방법론에 비해 Bayesian 방법론에 대한 리뷰논문이 부족

2

Modelling

사용된 모형별 비율

Model	
GLM, Partially LME[a]	2(9.1%)
Multivariate GLM	4(18.2%)
Multivariate mixed effect models	5(22.7%)
ZAB, Proportional-odds cumulative logit model[a]	2(9.1%)
GLM and CR mixed-effects model, Mixed-effect model and CR mixed-effects model, LME and continuous latent variable model, LME and a mixed-effects beta regression model, ZOIB[a]	5(22.7%)
MLIRT	2(9.1%)
MLLTM, MLTLM[a]	2(9.1%)

- 제일 많이 쓰인 모형 : Linear Mixed Model
- 흥미로웠던 모형 : Latent Variable Model

Latent Variable Model

Latent Variable

관측되지 않았지만 response variable에 영향을 줄 것이라 생각되는 변수

기존 변수로 설명되지 않은 분산을 잡아내는 것

measurement errors

summarizing different measurements

Latent Variable Model

Latent Variable

관측되지 않았지만 response variable에 영향을 줄 것 이라 생각되는 변수

기존 변수로 설명되지 않은 분산을 잡아내는 것

예측 변수에 직접적으로 영향을 주는 변수를 만드는 것

summarizing different measurements

Latent Variable Model

Latent Variable

관측되지 않았지만 response variable에 영향을 줄 것이라 생각되는 변수

기존 변수로 설명되지 않은 분산을 잡아내는 것

예측 변수에 직접적으로 영향을 주는 변수를 만드는 것

summarizing different measurements

Latent Variable Model

Latent Variable

관측되지 않았지만 response variable에 영향을 줄 것이라 생각되는 변수

Factor Analysis, Item Response Theory Model, Generalized Linear Mixed Model, Finite Mixture Model, Latent Class Model, Finite Mixture Regression Model, Latent Markov Model, Latent Growth/Curve Model, etc.

기존 변수로 설명되지 않은 분산을 잡아내는 것

예측 변수에 직접적으로 영향을 주는 변수를 만드는 것

표본 간 차이를 더욱 명확하게 나타낼 수 있는 변수를 만다는 것

3 Random Effect Distribution

사용된 가정별 비율

Random effect distribution	
Normal	12 (54.5%)
Multivariate normal	7(31.8%)
Dirichlet process prior	3(13.7%)

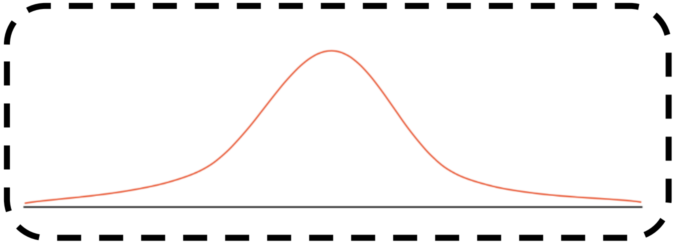
- 제일 많이 쓰인 분포 : Normal Distribution
- 흥미로웠던 분포 : Dirichlet Process Prior

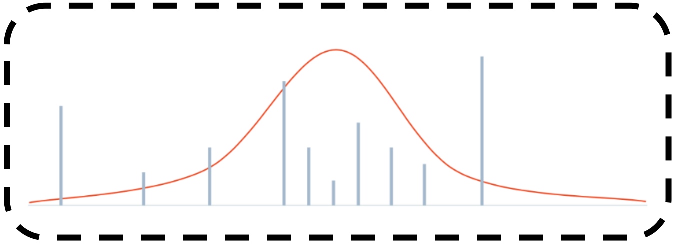
Dirichlet Process

아래와 같은 process G가 있다고 하자,

$$G \sim DP(\alpha, G_0)$$

(G_0는 sampling하게 될 Base Distribution이라고 하며,
$\alpha (> 0)$는 Scaling Parameter라고 한다.)
이 때, G는 Base Distribution G_0와 같은 Support를
가지는 Random Probability Measure이다.





Dirichlet Process

$$X_n|G \sim G \quad \text{for } n = \{1, \dots, N\}$$

$$G \sim DP(\alpha, G_0)$$

→ Marginalizing out G introduces dependencies between X_i's,

$$P(X_1, \dots, X_N) = \int P(G) \prod_{n=1}^{N} P(X_n|G)\, dG$$

→ $X_n|X_1, \dots, X_{n-1} = \begin{cases} X_i, \ \ with\ probability \frac{1}{n-1+\alpha} \\ new\ draw\ from\ G_0, \ \ with\ probability \frac{\alpha}{n-1+\alpha} \end{cases}$

Dirichlet Process

$$X_n|X_1, \dots, X_{n-1} = \begin{cases} X_k^*, & with\ probability \dfrac{num_{n-1}(X_k^*)}{n-1+\alpha} \\ new\ draw\ from\ G_0, & with\ probability \dfrac{\alpha}{n-1+\alpha} \end{cases}$$

$\rightarrow P(X_1, \dots, X_N) = P(X_1)P(X_2|X_1) \dots P(X_N|X_1, \dots, X_{N-1})$

$$= \frac{\alpha^K \prod_{k=1}^K (n_k - 1)!}{\alpha(1+\alpha)\dots(N-1+\alpha)} \prod_{k=1}^K G_0(X_k^*)$$

Dirichlet Process

$$X_n | X_1, \dots, X_{n-1} = \begin{cases} X_k^*, & with\ probability\ \dfrac{num_{n-1}(X_k^*)}{n-1+\alpha} \\ new\ draw\ from\ G_0, & with\ probability\ \dfrac{\alpha}{n-1+\alpha} \end{cases}$$

$$\rightarrow P(X_1, \dots, X_N) = P(X_1)P(X_2|X_1) \dots P(X_N|X_1, \dots, X_{N-1})$$

$$= \boxed{\frac{\alpha^K \prod_{k=1}^{K} (n_k - 1)!}{\alpha(1+\alpha)\dots(N-1+\alpha)}} \boxed{\prod_{k=1}^{K} G_0(X_k^*)}$$

새로운걸 뽑을 확률 기존걸 뽑을 확률

3 Random Effect Distribution

Dirichlet Process

$$X_n|X_1,\dots,X_{n-1} = \begin{cases} X_k^*, & with\ probability\ \dfrac{num_{n-1}(X_k^*)}{n-1+\alpha} \\ new\ draw\ from\ G_0, & with\ probability\ \dfrac{\alpha}{n-1+\alpha} \end{cases}$$

$$\rightarrow P(X_1,\dots,X_N) = P(X_1)P(X_2|X_1)\dots P(X_N|X_1,\dots,X_{N-1})$$

$$= \boxed{\frac{\alpha^K \prod_{k=1}^{K}(n_k-1)!}{\alpha(1+\alpha)\dots(N-1+\alpha)}}\boxed{\prod_{k=1}^{K} G_0(X_k^*)}$$

새로운걸 뽑을 확률 기존걸 뽑을 확률

관측치에 순서를 가정하고 전개했지만 순서가 의미가 없어짐

4 Error Distribution

사용된 가정별 비율

Error distribution	
Normal	18(48.6%)
N/I, SN[a]	3(8.1%)
t-distribution	1(2.8%)
ST	6(16.2%)
Multivariate ST	6(16.2%)
ALD	3(8.1%)

- 제일 많이 쓰인 분포 : Normal Distribution
- 흥미로웠던 분포 : Skew Normal Distribution

4 Error Distribution

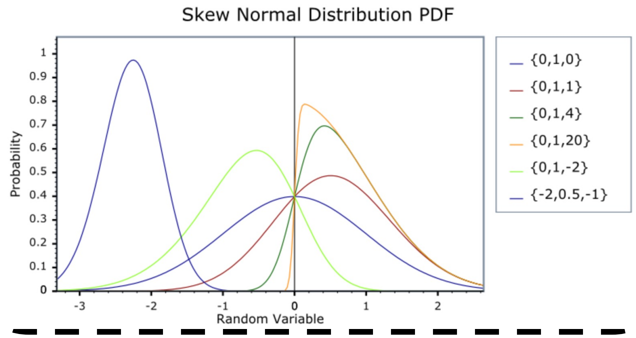
Skew Normal Distribution

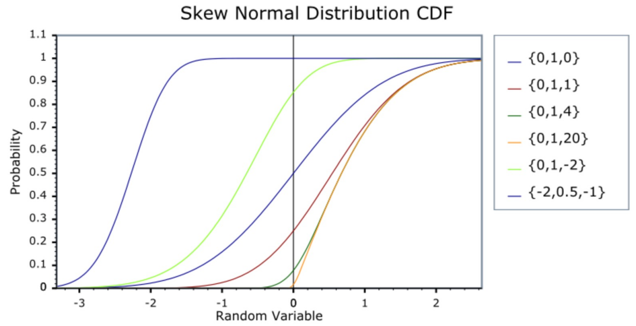
- 정규분포에서 파생된 분포 (shape parameter = 0)

$$f(x|\alpha) = 2\phi(x|\xi,\omega)\Phi(\alpha x)$$

→ $\xi : location, \omega : scale, \alpha = shape$

$$pdf : \frac{1}{(\omega\pi)^e}\int_{-\infty}^{\alpha(\frac{x-\xi}{\omega})} \exp\{-\frac{t^2}{2}\}\,dt$$





Skew Normal Distribution

- 정규분포에서 파생된 분포 (shape parameter = 0)

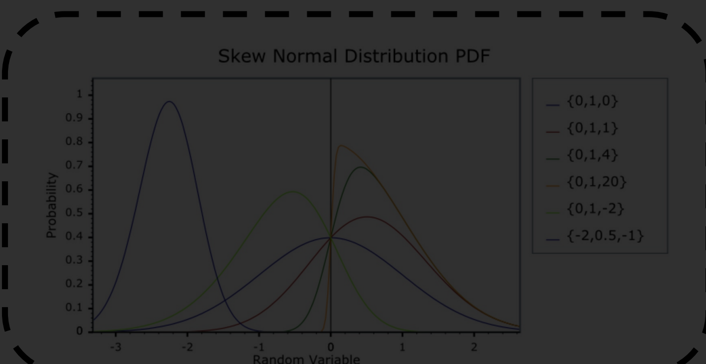
$$f(x|\alpha) = 2\phi(x|\xi,\omega)\Phi(\alpha x)$$

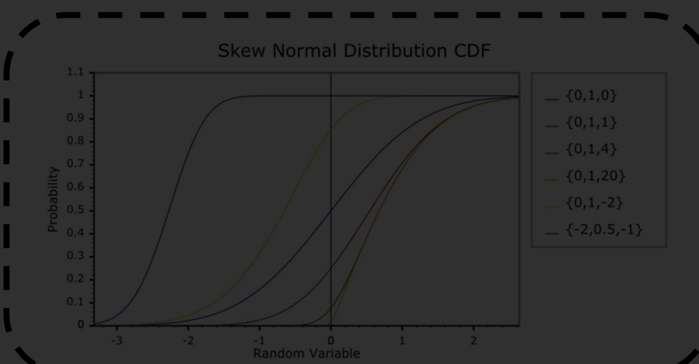
→ $\xi : location, \omega : scale, \alpha : shape$

이상치가 많을수록 Skew Normal Distribution이

결과에 대한 Robustness가 증가

$$pdf : \frac{1}{(\omega\pi)^e}\int_{-\infty}^{\frac{x-\xi}{\omega}} \exp\{-\frac{t^2}{2}\}\, dt$$





5 Association Structure

사용된 구조별 비율

Parameterisation	Latent association		Number of articles (%)[a]
Random effect	Univariate	$W_i(t)=\alpha b_i$	31(41%)
	Multivariate	$W_i(t)=\sum_{k=1}^{K}\alpha_k b_{ik}$	8(10.7%)
Current Value parameterisation	Univariate	$W_i(t)=\alpha m_i(t)$	14(18.7%)
	Multivariate	$W_i(t)=\sum_{k=1}^{K}\alpha_k m_{ik}(t)$	13(17.3%)
Correlated random effect	Univariate	$W_i(t)=\varphi$ with $B_i=\{b_i,\varphi_i\}\sim H$	2(2.7%)
	Multivariate	$W_i(t)=\varphi$ with $B_i=\{b_i,\varphi_i\}\sim H_{\alpha_k}$	1(1.3%)
Random effect with fixed effect	Multivariate	$W_i(t)=\sum_{k=1}^{K}\alpha_k(\beta_k+b_{ik})$	2(2.7%)
Time-dependent slope	Univariate	$W_i(t)=\alpha^{(1)}m_i(t)+\alpha^{(2)}\frac{d}{dt}m_i(t)$	1(1.3%)
	Multivariate	$W_i(t)=\sum_{k=1}^{K}\{\alpha_k{}^{(1)}m_{ik}(t)+\alpha_k{}^{(2)}\frac{d}{dt}m_{ik}(t)\}$	2(2.7%)
Cumulative effect	Univariate	$W_i(t)=\alpha\int_0^t m_i(s)ds$	1(1.3%)

- 제일 많이 쓰인 구조 : Random Effect
- 흥미로웠던 구조 : Current Value Parameterization

사용된 알고리즘별 비율

Sampling algorithm	Number of articles (%)
Markov Chain Monte Carlo (MCMC)	28(38.8%)
Gibbs sampler and Metropolis Hastings (MH)	24(33.3%)
Gibbs sampling	9(12.5%)
Gibbs sampling with adaptive rejection and MH	3(4.2%)
Block Gibbs sampling and MH	2(2.8%)
Bayesian Lasso	1(1.4%)
Newton-Raphson procedure and a derivative-based MCMC	1(1.4%)
No-U-Turn sampler	2(2.8%)
Hamiltonian Monte Carlo (HMC)	1(1.4%)
HMC and No-U-Turn sampler	1(1.4%)

- 제일 많이 쓰인 알고리즘 : MCMC
- 흥미로웠던 알고리즘 : Hamilton Monte Carlo

Hamilton Monte Carlo

- Pdf의 미분을 이용한 사후분포(Target Density) 간 transition 생성
- Leapfrog Integrator로 적분근사

1. Auxiliary Momentum Variable

- ρ : Auxiliary Momentum Variable
- $\mathrm{P}(\rho, \theta) = P(\rho|\theta)P(\theta)$에서 sampling
- $\rho \sim MultiNormal(0, \Sigma),\ where\ \Sigma = I$ assumed mostly

Hamilton Monte Carlo

2 The Hamiltonian

- $H(\rho, \theta) \coloneqq Hamiltonian\ of\ P(\rho, \theta)$
- $H(\rho, \theta) = -\log P(\rho, \theta) = -\mathrm{logP}(\rho|\theta) - \mathrm{logP}(\theta)$

$$= T(\rho|\theta) + V(\theta)$$

- $T(\rho|\theta) \coloneqq \text{Kinetic Energy}$
- $\mathrm{V}(\theta) \coloneqq \text{Potential Energy}$

Hamilton Monte Carlo

3 Generating Transitions

- Start with the current value $\theta^{(i)}$
- Draw $\rho^{(i)} \sim MultiNormal(0, \Sigma)$
- $\frac{d\theta}{dt} = +\frac{\partial H}{\partial \rho} = +\frac{\partial T}{\partial \rho}$
- $\frac{d\rho}{dt} = -\frac{\partial H}{\partial \theta} = -\frac{\partial T}{\partial \theta} - \frac{\partial V}{\partial \theta} = -\frac{\partial V}{\partial \theta}$

Hamilton Monte Carlo

4. Leapfrog Integrator

- With $\rho^{(i)}$, update $\rho^{(i+1)} = \rho^{(i)} - \frac{\epsilon}{2}(\frac{\partial V}{\partial \theta})$

  $\theta^{(i+1)} = \theta^{(i)} + \epsilon \Sigma \rho^{(i+1)}$ for L times

- Define $\rho^{(L)} = \rho^*, \theta^{(L)} = \theta^*$

5. Metropolis Accept Step

- Accept ρ^* and θ^* if $\exp\left(H\left(\rho^{(1)}, \theta^{(1)}\right) - H(\rho^*, \theta^*)\right) \geq 1$
- If rejected, use θ^* as the initial value to repeat the above procedure

감사합니다