We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900
Open access books available

185,000
International authors and editors

200M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Chapter

Bayesian Analysis for Random Effects Models

Junshan Shen and Catherine C. Liu

Abstract

Random effects models have been widely used to analyze correlated data sets, and Bayesian techniques have emerged as a powerful tool to fit the models. However, there has been scarce literature that systematically reviews and summarizes the recent advances of Bayesian analyses of random effects models. This chapter reviews the use of the Dirichlet process mixture (DPM) prior to approximate the distribution of random errors within the general semiparametric random effects models with parametric random effects for longitudinal data setting and failure time setting separately. In a survival setting with clusters, we propose a new class of nonparametric random effects models which is motivated from the accelerated failure models. We employ a beta process prior to tact clustering and estimation simultaneously. We analyze a new data set integrated from Alzheimer's disease (AD) study to illustrate the presented model and methods.

Keywords: beta process, Dirichlet process mixture, clustered data, longitudinal data, random effects, survival outcome, nonparametric transformation model

1. Introduction

Random effects models have been widely used as a powerful tool for analyzing correlated data [1, 2]. The model features a finite number of random terms acting as latent variables to model unobserved factors; see [3] for a comprehensive review. Some authors have further proposed semiparametric mixed effect models by allowing for infinite dimensional random effects [4, 5]. Most of the aforementioned works draw inferences using frequentist approaches, while Bayesian approaches have been largely ignored because of the lack of computational feasibility and expediency. With the advent of the "supercomputer" era, Bayesian analyses have recently sparked much interest in the setting of random effects models for clustered data or longitudinal settings. However, there is scarce literature that has systematically reviewed the Bayesian works in the area.

By extending the traditional random effects models, recent research focus has shifted to study heterogeneous random effects or nonparametric distributions of random effects, which arise because of skewness of data, missing covariates, or unmeasurable subject-specific covariates [6]. The extended random effects models, termed semiparametric random effects models, improve statistical performance with added interpretability. Bayesian techniques, which provide a convenient means to model non-Gaussian distributions, have recently been proposed for semiparametric random effects model in a variety of settings ([7, 8], among others).

IntechOpen

The discreteness of the Dirichlet process makes it impossible as a prior for estimating a density. However, as a remedy by convolving with a kernel, Dirichlet process mixture plays an important role [9].

For censored outcome data, transformation models, which transform the time-to-event responses using a monotone function and link them to the covariates of interest, have surged as a strong competitor of the Cox model [10]. Moreover, the transformation model framework is fairly general. The Cox model and the proportional odd model [11] can be viewed as nonparametric transformation linear models with some specific error terms; see [12–14]. For correlated data, the transformation model naturally extends the semiparametric random effects model by directly incorporating random effects to the transformation functions, treating them as realizations of an underlying random function. Bayesian analyses have found much use in this new area. For example, the beta process has been found to be a reasonable candidate for the prior of the monotone transformation function [15–17].

This chapter focuses on the Bayesian analysis of the transformed linear model with censored data and in a clustered setting. In many biomedical studies, the observations are naturally clustered. For example, patients in observational studies can be grouped in analysis according to a variety of factors, such as age, race, gender, and hospital, in order to reduce the confounding effects. Following Mallick and Walker [18], we explore using a mixture of beta distributions and the beta process as the candidates for the prior distribution of the random transformation function [17, 19, 20].

The rest of this chapter is structured as follows. Section 2 reviews the use of the Bayesian approach to infer parametric random effects models. In the setting of survival analysis, Section 3 proposes a beta process prior to fit random effects model with nonparametric transformation functions, and Section 4 applies the method to study the progression of Alzheimer's disease (AD). Section 5 concludes the chapter with future research directions.

2. Dirichlet process mixture prior

In parametric random effects models, we considered the situation that the distribution form of the random error term is unknown. Dirichlet process mixture (DPM) is used as the prior for the baseline distribution in that error terms used to be continuous random variables in most situations.

2.1 Linear mixed effects model

With a longitudinal data set $\{Y_i, x_i, z_i\}$, we posit a mixed effects model with an AR(1) serial correlation structure:

$$\begin{aligned} &\mathbf{y}_i = \mathbf{x}_i\boldsymbol{\beta} + \mathbf{z}_i\mathbf{b}_i + \mathbf{w}_i, i = 1, \ldots, m; \\ &\mathbf{w}_i = (w_{i1}, \ldots, w_{in_i})^T; w_{ij} = \rho w_{i,j-1} + \epsilon_{ij}, j = 2, \ldots, n_i, \end{aligned} \tag{1}$$

where $\mathbf{y}_i = \left(y_{i1}, \ldots, y_{in_i}\right)^T$ with y_{ij} being the jth response of the ith subject for $i = 1, \ldots, m$, β is a $p \times 1$ vector of fixed effect parameters, $\mathbf{b}_i$ a $q \times 1$ Gaussian random vector representing the subject-specific random effects, $\mathbf{x}_i$ and $\mathbf{z}_i$ are $n_i \times p$ and $n_i \times q$ design matrices linking β and $\mathbf{b}_i$ to $\mathbf{y}_i$, respectively, $\mathbf{w}_i = (w_{i1}, \ldots, w_{in_i})^T$ is an $n_i \times 1$ vector of model errors, ρ is the autoregressive coefficient, and ϵ_{ij},s are i.i.d. noises. When $\{\epsilon_{ij}\}$ is non-normal, we assume a mixture model:

$$f_G(\epsilon|\sigma^2) = \int \varphi(\epsilon|u, \sigma^2) dG(u), \tag{2}$$

where $\varphi(\cdot|u, \sigma^2)$ is the probability density function for a normal random variable with mean u and variance σ^2 and G is an unspecified probability distribution of u satisfying $\int u dG(u) = 0$, which ensures that ϵ comes from a mean-zero mixture distribution.

Replacing the Dirichlet process by an equivalent Pólya urn representation, [8] employed an empirical likelihood approach with the moment constraints and developed a posterior adjusted Gibbs sampler for more precise estimation. The algorithm is computationally feasible.

2.2 Accelerated failure time model

We shift gears to study survival outcomes with a cluster structure. Denote the data set by $(T_{ij}, X_{ij}), i = 1, \cdots, K, j = 1, \cdots, n_i$, where T_{ij} is the failure time of the jth subject in the ith cluster and X_{ij} is a vector of associated covariates. To accommodate such data, we utilize a general accelerated failure time model:

$$\log T_{ij} = \mathbf{X}_{ij}^T \boldsymbol{\beta} + \varepsilon_{ij}, \qquad i = 1, \cdots, K \text{ and } j = 1, \cdots, n_i, \tag{3}$$

where β is a vector of p-dim regression coefficients of interest and ε_{ij} are independent random errors following the distribution with density f_i. [7] posed an exponential tilt on the distributions of error terms to incorporate the cluster heterogeneity. That is,

$$\frac{f_i(t)}{f_1(t)} = \exp(\theta_{0i} + \boldsymbol{\theta}_i^T q(t)), \qquad i = 2, \cdots, K, \tag{4}$$

where $q(t)$ is a q-dimensional prespecified functions containing potential covariate information and $\boldsymbol{\theta}_i$ is the corresponding parameter vector with $\theta_{0i} = \log\left[\left(\int \exp(\boldsymbol{\theta}_i^T q(t)) f_1(t) dt\right)^{-1}\right]$. Thus, $\boldsymbol{\theta}_i$ represents the parametric random effects in the model. Li et al. [7] place the DPM prior on the baseline density f_1 to develop a set of procedures which improves estimation efficiency through information pooling.

3. Beta process prior

We now present a nonparametric random effects model for the clustered survival data with nonparametric monotone link functions. We employ a beta process as the prior for the baseline function.

Let T_{ij} denote the failure time of the jth subject in the ith cluster, X_{ij} be the covariate vector for the subject, and C_{ij} be the potential censoring time to the jth subject in the ith cluster. Assume that C_{ij} is independent of the failure time T_{ij}. Let $Z_{ij} = \min(T_{ij}, C_{ij})$ and let $\delta_{ij} = I(T_{ij} < C_{ij})$ be the censoring indicator. Then the observed data can be described as

$$(Z_{ij}, \delta_{ij}, X_{ij}), i = 1, \cdots, n; \quad j = 1, \cdots, n_i. \tag{5}$$

Within each cluster, T_{ij} is linked to X_{ij} via the following transformation model:

$$\ln H_i(T_{ij}) = X_{ij}^T\boldsymbol{\beta} + \ln\varepsilon_{ij}, i = 1, 2\cdots, n, \tag{6}$$

where ε_{ij} are i.i.d. variables with a known density function $f_\varepsilon(\cdot)$ and $H_i(t)$ are unknown cluster-specific monotone functions, which are i.i.d. realizations of a random function and can be viewed as a nonparametric version of random effects for independent clusters. In a parametric setting, if we set $H_i(t) = t\exp(-b_i)$ with b_i being a cluster-specific random effect, Eq. (6) reduces to a classical random effects model, which has been discussed in Section 2.2. The challenge, however, lies in how to draw inferences in such a nonparametric setting.

To proceed, let the coefficient vector $\boldsymbol{\beta}$ be a p-dim unknown vector of interest. We further assume H_i's are differentiable with derivative $h_i(t) = H_i'(t)$, and then the likelihood based on the observed data is

$$L(\boldsymbol{\beta}, H_1, \cdots, H_n|data) = \prod_{i=1}^{n}\prod_{j=1}^{n_i} p\left(T_{ij}, X_{ij}, \delta_{ij}|H_i, \beta\right), \tag{7}$$

where

$$p(t, x, \delta|H, \beta) = \left(f_\varepsilon\left(H(t)e^{-x^T\beta}\right)h(t)e^{-x^T\beta}\right)^{\delta} S_\varepsilon\left(H(t)e^{-x^T\beta}\right)^{1-\delta}.$$

Here S_ε is the survival function of *varepsilon* defined by $S_\varepsilon(s) = P(\varepsilon \geq s)$.

We develop a Bayesian inference procedure based on model (6). We assume that the regression coefficient β follows a normal prior:

$$\beta \sim N_p\left(0, \sigma_\beta^2 I_p\right), \tag{8}$$

where I_p is the $p \times p$ dimensional identity matrix. Since H_i is assumed differentiable, we model it with a kernel convolution:

$$H_i = \int \Phi_\sigma(\cdot - s)dB_i(s),$$

where B is an increasing function and Φ_σ is the zero-mean normal distribution with variance σ^2. Hence, the derivative of H_i is

$$h_i = \int \phi_\sigma(\cdot - s)dB_i(s)$$

with $\phi_\sigma(t) = \frac{1}{\sigma}\phi\left(\frac{t}{\sigma}\right)$. This actually mimics the idea of DPM to smooth beta process by convolution.

We are in a position to select an appropriate stochastic process used as the prior of B_i. Beta process, as studied by [16, 17], is an ideal candidate for the prior of a monotone function. Specifically, beta process $\mathrm{BP}(\gamma, B_0)$ with concentration parameter γ and a base measure B_0 is an increasing Lévy process with independent increments of the form

$$dB(t) \sim \mathrm{Beta}\left(\gamma dB_0(t), \gamma(1 - dB_0(t))\right).$$

Teh et al. [20] showed that a sample from $\mathrm{BP}(\gamma, B_0)$ could be represented as

$$B_i(y) = \sum_{l=1}^{\infty} p_{il} I(\theta_{il} \le y), \tag{9}$$

where $p_{il} = \prod_{j=1}^{l} \nu_{il}$ and (θ_{il}, ν_{il}) follows

$$\theta_{il} \sim B_0(\theta), \nu_{il} \sim \mathrm{Beta}(\gamma, 1) \quad l = 1, 2, \cdots.$$

In practice, we need to approximate samples of $\mathrm{BP}(\gamma, B_0)$ with a finite dimensional form. Since beta process $\mathrm{BP}(\gamma B_0)$ can be represented by a stick-breaking process defined in Eq. (9), a natural approximation is obtained by retaining its first L components. That is,

$$B_i^* = \sum_{l=1}^{L} p_{il} \delta_{\theta_{il}},$$

with $p_{il} = \prod_{j=1}^{l} \nu_{il}, l = 1, \cdots, L$. Denote $\boldsymbol{\xi}_i = (\nu_{i1}, \cdots, \nu_{iL}, \theta_{i1}, \cdots, \theta_{iL})^T$ and define

$$H_\sigma^*(z, \boldsymbol{\xi}_i) = \sum_{l=1}^{L} p_{il} \Phi_\sigma(z - \theta_{il}), h_\sigma^*(z, \boldsymbol{\xi}_i) = \sum_{l=1}^{L} p_{il} \phi_\sigma(z - \theta_{il}).$$

The approximated posterior based on the truncated DP is

$$\pi(\beta) \prod_{i=1}^{n} \left[\pi^{\xi}(\boldsymbol{\xi}_i) \prod_{j=1}^{n_i} f\left(Z_{ij}, X_{ij}, \delta_{ij} | \beta, \boldsymbol{\xi}_i\right) \right], \tag{10}$$

where

$$f(z, x, \delta | \beta, \boldsymbol{\xi}) = \left(p_\varepsilon\left(H_\sigma^*(z, \boldsymbol{\xi}) \exp\left(-x^T \boldsymbol{\beta}\right)\right) h_\sigma^*(z, \boldsymbol{\xi}) \exp\left(-x^T \boldsymbol{\beta}\right)\right)^{\delta}$$

$$\times \left(P_\varepsilon\left(H_\sigma^*(z, \boldsymbol{\xi}) \exp\left(-x^T \boldsymbol{\beta}\right)\right)\right)^{1-\delta}.$$

The samples for $\boldsymbol{\beta}$ and $(\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_n)$ based on the posterior can be obtained with Markov chain Monte Carlo (MCMC) [21]. In our simulation, we use the R-package MCMC (https://cran.r-project.org/web/packages/mcmc/index.html) to draw samples for $\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_n$ and β and use the Metropolis algorithm with a normal working distribution.

4. An application to Alzheimer's disease neuroimaging initiative

Alzheimer's Disease Neuroimaging Initiative (ADNI) is a multisite cooperative study for the purpose of improving the prevention and treatment of Alzheimer's disease. The subjects in the study fall into three groups, cognitively normal (CN) individuals, mild cognitive impairment (MCI) patients, and early AD patients. ADNI provides a rich array of patients' information, including functional magnetic resonance imaging (fMRI), positron emission tomography (PET), longitudinal functional cognitive tests scores, blood samples, genetics data, and censored failure time outcomes. Details of the study can be found at http://adni.loni.usc.edu.

We focus on the MCI group. MCI is recognized as a transitional stage between normal cognition and Alzheimer's disease. The failure time is defined to be the time that a MCI patient is diagnosed with AD, which will be censored if a MCI patient remains at the MCI stage at the end of the follow-up time. Wide heterogeneities are exhibited among the failure times, which may be due to demographics and a variety of functional clinical biomarkers, such as the brain areas of the hippocampus, ventricles, and entorhinal cortex. The goal of the analysis is to study the impact of risk factors on progression to AD.

Using the same data as analyzed by [14], we demonstrate our methodology by modeling the failure time (the observed time of AD diagnosis from MCI stage in year) of 281 MCI patients on gender (0 = female, 1 = male), years of education, the number of apolipoprotein E alleles (0, 1, or 2), and the baseline hippocampal volume.

As age is a strong confounder but the functional form of its impact has not reached consensus, we elect to model its impact nonparametrically. Specifically, we use age to form two strata (below and above the median age) and use model (6) to estimate the stratum-specific transformation functions and the effects of other covariates. For comparisons, we also fit model (6) with age as a continuous variable and with a common transformation function. That is, we do not assume the data are clustered. For both models, the regression errors ε's are assumed to follow an exponential distribution with mean 10. In our calculation, we approximate the BPs by a finite truncation with $L = 20$. We assume the precision parameter $\alpha = 1$ and scale parameter $\sigma^2 \sim 1/\sigma^2$.

Figure 1 illustrates the estimated transformation function H of the failure time without clustering. The posterior means (PM) and standard errors (SE) of the regression coefficients in the model are reported in **Table 1**. We run the MCMC for

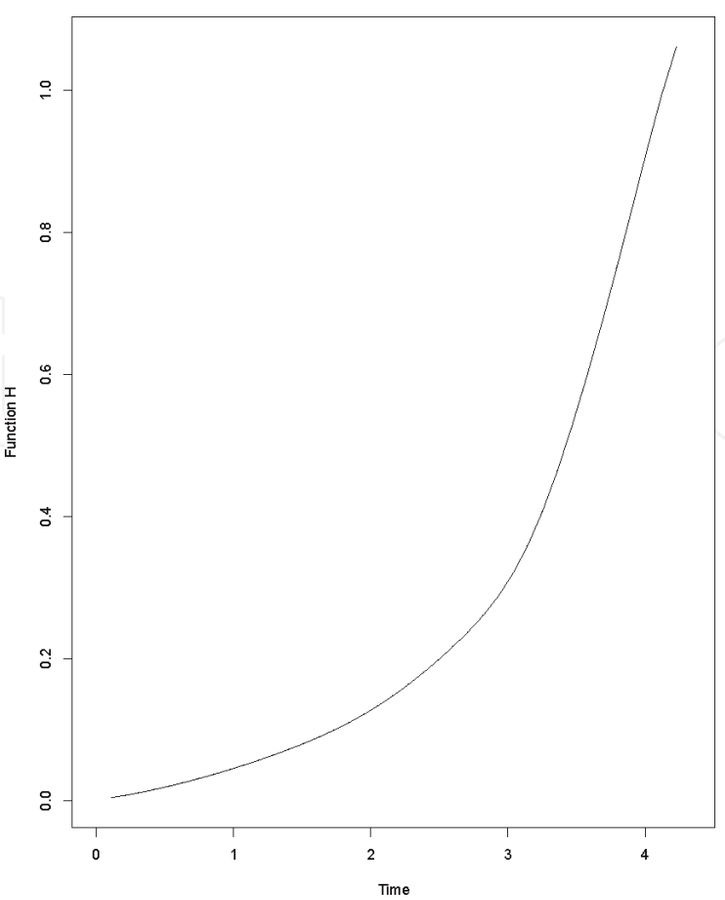


Figure 1.
Smoothed transformation function without clustering.

20,000 iterations with the first 4000 draws discarded as burn-in samples and use Geweke's statistic to ensure the convergence of the chains.

Figure 2 illustrates the estimated transformation functions with age-stratified data, and **Table 2** summarizes the posterior means and standard errors of the other regression coefficients.

The left curve is relatively flat, while the right curve has a sharper slope. This is consistent with the recognition that AD is an aging disease: elder people above a certain age threshold tend to progress faster from MCI to AD.

Both **Tables 1** and **2** show that none of the biomarkers are significant, whereas they are statistically significant in the analysis of [14]. One possible conjecture is that our nonparametric transformation functions may have well captured the effects of unobserved confounders, which may leave little to be explained by the observed covariates. More thorough investigation is warranted.

	RID	AGE	PTGENDER	PTEDUCAT	APOE4	Hipp.
PM	−0.9635	0.0069	−0.1453	−0.0231	−0.1817	0.2710
SE	1.3288	0.0841	1.2331	0.1835	0.8616	0.5333

Table 1.
Posterior estimates of regression coefficients with standard errors.

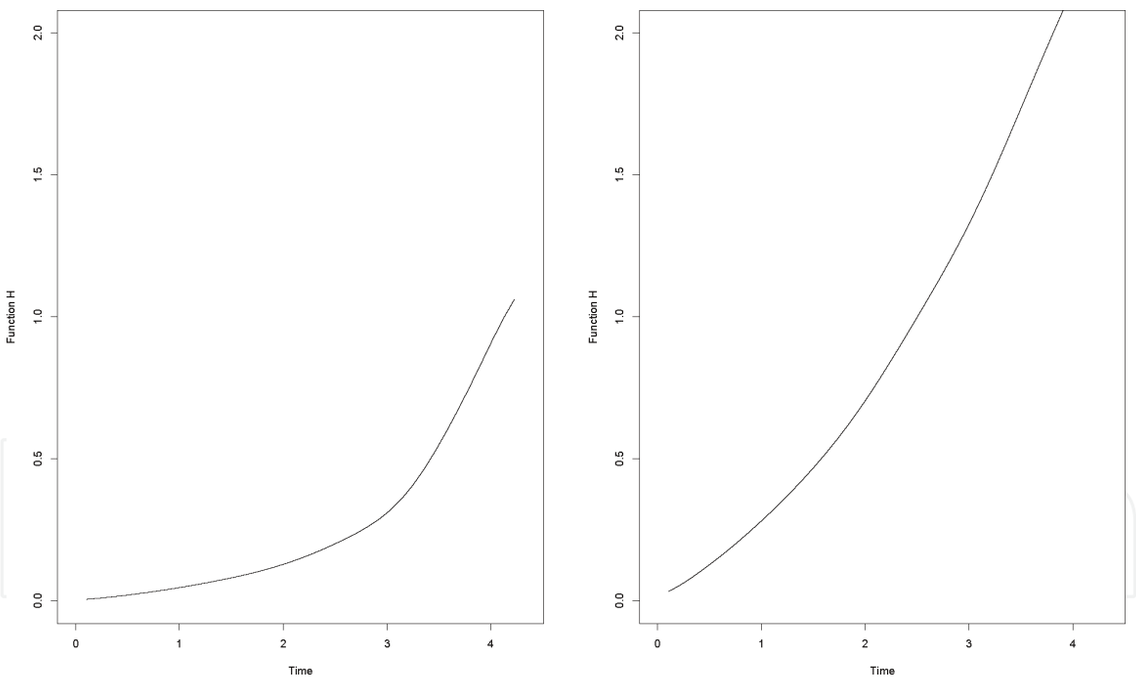


Figure 2.
Smoothed transformation functions with two age-strata: The left curve is the smoothed transformation function for group aged below the average age; the right curve is the smoothed transformation function for the group aged over the average age.

	RID	PTGENDER	PTEDUCAT	APOE4	Hipp.
PM	−0.6399	−0.0706	−0.0072	−0.1349	0.1919
SE	0.9273	0.8491	0.1267	0.6098	0.3716

Table 2.
Posterior estimators of regression coefficients with standard errors.

5. Future directions

Following [12], we can extend the transformation model (6) by allowing the error function f_ε to be unspecified. In this case, we need to specify the regression coefficient β to obey some constraints such as $\beta_1 = 1$ or $\|\boldsymbol{\beta}\| = 1$ for identifiability. We will propose to model the error function using a Dirichlet processes mixture model:

$$f_\varepsilon(t) = \int \varphi(t|\mu, \sigma^2) dG(\mu, \sigma^2), \quad G \sim \mathrm{DP}\left(\alpha, G_0 = \mathrm{N}\left(\mu|\mu_0, \sigma_0^2\right) \times \mathrm{IG}\left(\alpha_1, \alpha_2\right)\right),$$

where $\varphi(t|\mu, \sigma^2)$ is a normal kernel with mean μ and variance σ^2 and G are samples from a Dirichlet process $\mathrm{DP}\left(\alpha_1, G_0 = \mathrm{N}\left(\mu|\mu_0, \sigma_0^2\right) \times \mathrm{IG}\left(a, b\right)\right)$, where α_1 is the mass parameter and $\mathrm{IG}\left(\cdot|a, b\right)$ is the inverse gamma distribution with shape parameter a and scale parameter b.

In a slightly different context, we may also consider clustering observations by developing a new nested beta-Dirichlet process prior with companion MCMC algorithms. As there are limited works on functional random effects models that accommodate clustering structures observed, for example, from neural studies, we may propose a nested Dirichlet process [19] as the prior of Dirichlet process to cluster cumulative distribution functions successfully. We envision that such a nested Bayesian procedure will provide substantial computational expedience for practitioners and can certainly be applied to studies that cover beyond the neurodegenerative and aging diseases.

Acknowledgements

Shen's research is partially supported by Beijing Natural Science Foundation 1192006 and National Natural Science Foundation of China; Liu's research is partially supported by General Research Fund, Research Grants Council, Hong Kong, 15327216, and the Hong Kong Polytechnic University grant YBTR. Data collection and sharing for this project were funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research Development, LLC.; Johnson & Johnson Pharmaceutical Research Development LLC.; Lumosity; Lundbeck; Merck Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

Data used in preparation of this article were obtained from the Alzheimers Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf.

IntechOpen

Author details

Junshan Shen[1†] and Catherine C. Liu[2*†]

1 Capital University of Economics and Trade, Beijing, China

2 The Hong Kong Polytechnic University, Hong Kong SAR

*Address all correspondence to: macliu@polyu.edu.hk

† These authors contributed equally.

IntechOpen

© 2020 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. CC BY

References

[1] Harville DA. Extension of the Gauss–Markov theorem to include the estimation of random effect. The Annals of Statistics. 1976;**4**:384-395

[2] Laird NM, Ware JH. Random-effects models for longitudinal data. Biometrics. 1982;**38**(2):963-974

[3] Li Y. Random effect models. In: Pham H, editor. Springer Handbook of Engineering Statistics. London: Spring-Verlag; 2006. pp. 687-704

[4] Zeger SL, Diggle PJ. Semiparametric models for longitudinal data with application to CD4 cell numbers in HIV seroconverters. Biometrics. 1994;**50**(3): 689-699

[5] Li Y, Lin X, Müller P. Bayesian inference in semiparametric mixed models for longitudinal data. Biometrics. 2010;**66**(1):70-78

[6] Li Y, Mller P, Lin X. Center-adjusted inference for a nonparametric Bayesian random effect distribution. Statistica Sinica. 2011;**21**:1201-1223

[7] Li Z, Xu X, Shen J. Semiparametric Bayesian analysis of accelerated failure time models with cluster structures. Statistics in Medicine. 2017;**36**(25): 3976-3989

[8] Shen J, Yu H, Yang J, Liu C. Semiparametric Bayesian analysis for longitudinal mixed effects models with non-normal AR(1) errors. Statistics and Computing. 2019;**29**(3):571-583

[9] Ghosal S, van der vaart A. Fundamentals of Nonparametric Bayesian Inference. Cambridge: Cambridge University Press; 2017

[10] Cox DR. Regression models and life tables (with discussion). Journal of the Royal Statistical Society: Series B. 1972; **34**(2):187-220

[11] Hanson T, Yang M. Bayesian semiparametric proportional odds models. Biometrics. 2007;**63**(1):88

[12] Horowitz JL. Semiparametric estimation of a regression model with an unknown transformation of the dependent variable. Econometrica. 1996;**64**(1):103-137

[13] Linton O, Sperlich S, Keilegom IV. Estimation of a semiparametric transformation model. Annals of Statistics. 2008;**36**(2):686-718

[14] Li K, Luo S. Functional joint model for longitudinal and time-to-event data: An application to Alzheimer's disease. Statistics in Medicine. 2017;**36**(25): 3560-3572

[15] Müller P, Mitra R. Bayesian nonparametric inference why and how. Bayesian Analysis. 2013;**8**(2):269-302

[16] Kalbfleisch JD. Non-parametric bayesian analysis of survival time data. Journal of the Royal Statistical Society. 1978;**40**(2):214-221

[17] Hjort N. Nonparametric bayes estimators based on beta processes in models for life history data. Annals of Statistics. 1990;**18**(3):1259-1294

[18] Mallick B, Walker S. A Bayesian semiparametric transformation model incorporating frailties. Journal of Statistical Planning and Inference. 2007; **112**(1):159-174

[19] Rodriguez A, Dunson DB, Gelfand AE. The nested Dirichlet process. Journal of the American Statistical Association. 2008;**103**(483): 1131-1154

[20] Teh YW, Görür D, Ghahramani Z. Stick-breaking construction for the Indian buffet process. In: Proceedings of

the Eleventh International Conference on Artificial Intelligence and Statistics; Vol. 2; 2007. pp. 556-563

[21] Geyer CJ. Introduction to Markov chain Monte Carlo. In: Brooks S, Gelman A, Jones G, Meng X-L, editors. Handbook of Markov Chain Monte Carlo. Boca Raton: Chapman and Hall/CRC; 2011. pp. 3-48

IntechOpen

IntechOpen