FORMATH
Forest Resources & Mathematical Modeling

©2021 The Author(s) CC-BY 4.0
Scientific Category

FORMATH Vol. 20 (2021), DOI:10.15684/formath.20.001

Gamma Regression Model with Nuisance Baseline for Tree Growth Data

Tetsuji Tonda[1*], Ken-ichi Kamo[2], Masayoshi Takahashi[3]

Abstract: In regression model for longitudinal data such as forest growth, it is important to model the growth behavior and to evaluate the effects of covariates such as the impact of environmental factors on growth behavior. In this paper, we consider a situation where the primary interest is centered on the effects of covariates. Though the time trend of the baseline is non-negligible, it is out of scope for this study. Considering age dependent growth behavior as baseline, we can reconstruct the models to include a nuisance baseline. Then it is possible only to estimate the regression coefficients of covariates without modeling the shape for the baseline growth behavior. The linear regression model assumes normality and homoscedasticity for error terms. However, forest growth data sometimes fail to meet the above assumptions because of data non-negativity. In this paper, we develop a regression model for longitudinal non-negative continuous data, using gamma distribution. We apply the proposed method to a forest growth data on Sugi (*Cryptomeria japonica*) at Hoshino village in Japan.

Keywords: tree growth, gamma regression model, nuisance baseline

1. Introduction

In our paper we construct a statistical method for modeling longitudinal trend on the growth of forest stands. It is well known that several environmental factors affect the growth of forest stand. For example, Imada et al. (1997) investigated the relationships between planting density and the growth behavior based on spacing experiment for Japanese oak (*Quercus mongolica* var. *grosseserrata*) using a systematic design by Nelder (1962). Regression analysis is one of statistical methods to evaluate effects of several environmental factors on the growth behavior of forest stands. Forest growth data such as height, DBH (diameter at breast height) and volume are important factors for projecting the growth of forest resources, which can be used to determine the timing of thinning activities, as well as final harvest. To model the growth behavior of forest stands, various nonlinear models have been proposed based on mathematical formulations of the growth process. In general, the mathematical formulation of growth processes allows us to understand the role of parameters in a model. Burkhart and Tomé (2012) reviewed many growth functions commonly used for modeling tree growth data and interpretations of the parameters. Zeide (1993) introduced twelve types of nonlinear growth models. One of those nonlinear growth models is the Chapman-Richards growth function (Richards, 1958) which can be expressed as:

$$[1] \qquad E[y(t)] = \alpha \left(1 - e^{-\beta t}\right)^{\gamma}$$

where $y(t)$ is the observation on the growth of a forest stand at age t, and α,β,γ are positive unknown parameters describing the shape of the growth function. Chapman-Richards growth function is one of widely used sigmoid growth functions in modeling growth data. Others include the Gompertz function (Gompertz 1825) which is also widely used to model growth data. Tjørve and Tjørve (2020) reviewed several representations of the Gompertz function with their usefulness. These nonlinear functions are developed under some assumptions based on mathematical constraints. When these assumptions are not satisfied, it may include some element of bias in the results. It is therefore important to pay attention to these assumptions in the selection of nonlinear functions when modeling growth behaviors.

One approach to overcome the aforementioned problem is to use information criterion such as Akaike's Information Criterion (Akaike 1973) for model selection. A review of literature shows that Kamo and Yoshimoto (2013) compared three criteria: AIC, Mallows' Cp (Mallows 1973) and

Received Feb. 8, 2021; Accepted Apr. 8, 2021
[1]Prefectural University of Hiroshima, Japan
[2]Sapporo Medical University, Japan
[3]Forest Research and Management Organization, Japan
*Corresponding Author: ttetsuji@pu-hiroshima.ac.jp

Cross-validation (Stone 1974; 1977) for selecting growth function among candidate models in Zeide (1993). Kamo (2019) proposed a new clustering approach for tree growth data by combining model selection and k-means method.

Another approach is to treat growth behaviors over time as a nuisance baseline trend. Satoh and Tonda (2016) introduced nuisance baseline in growth curve models (Kollo and von Rosen, 2005). Kamo et al. (2017) proposed regression model with nuisance baseline for forest growth data. These approaches were based on linear regression models using data on continuous outcomes. These linear regression models assume normality and homoscedasticity for the error terms. However, forest growth data sometimes fail to meet these assumptions because the forest growth data are non-negative. In our study, we develop a method to estimate regression coefficients only for covariate of interest by regarding the growth behavior as nuisance trend under the framework of Gamma regression. This is a type of semi-parametric approach because there are no assumptions on specific functional form for nuisance baseline trend.

2. Materials and Methods

2.1. Methods

We adopt the gamma distribution for random component in a regression model. The gamma distribution is a family of continuous probability distribution with two parameters. Note that exponential distribution and chi-squared distribution are special cases of gamma distribution. There are several parameterizations to describe the gamma distribution. For example, the characterization using shape parameter ν and scale parameter θ (Johnson et al. 1994). The probability density function of gamma distribution is given by:

$$[2] \qquad f(y) = \frac{1}{\Gamma(\nu)} \theta^{-\nu} y^{\nu-1} e^{-\frac{y}{\theta}}.$$

where $\Gamma(\nu)$ is the gamma function defined by:

$$[3] \qquad \Gamma(\nu) = \int_0^\infty u^{\nu-1} e^{-u} du.$$

From equation [2], the mean and variance of the gamma distribution can be expressed as $\nu\theta$ and $\nu\theta^2$, respectively. Faraway (2005) suggests the usefulness of characterization using mean parameter μ and scale parameter ν instead of shape and scale parameters. For regression analysis, it is convenient to reparametrize by substituting $\theta = \mu\nu^{-1}$ into equation [2]:

$$[4] \qquad f(y) = \frac{1}{\Gamma(\nu)} \left(\frac{\nu}{\mu}\right)^{\nu} y^{\nu-1} e^{-\frac{\nu y}{\mu}}.$$

We represent this parameterization of gamma distribution as Gamma (μ, ν). The mean and variance of Gamma (μ, ν) are μ and $\mu^2\nu^{-1}$, respectively. In particular $\phi = \nu^{-1}$ is called the dispersion parameter in terms of the generalized linear model (McCullagh and Nelder 1989). For large shape parameter ν (or small dispersion parameter ϕ) the gamma distribution can be approximated by normal distribution with mean μ and variance ϕ.

Let $y_i(t)$ $(t = t_1, \ldots, t_m)$ be longitudinal measurements of non-negative continuous outcomes for a sample i at given age t. Here we assume that $y_i(t)$ follows a Gamma distribution, Gamma $(\mu_i(t), \nu)$. To describe the relationship between longitudinal measurements and covariates, we consider the following log linear model:

$$[5] \qquad E[y_i(t)] = \mu_i(t) = \lambda(t) \exp\left(\boldsymbol{a}_i' \boldsymbol{\beta}\right),$$

where $\boldsymbol{a}_i$ is a vector of covariates, $\boldsymbol{\beta}$ is the regression coefficient on the effect of $\boldsymbol{a}_i$, and $\lambda(t)$ is the nuisance baseline growth. The maximum likelihood estimator of $\boldsymbol{\beta}$ and ν can be obtained by maximizing the log-likelihood function:

$$[6] \qquad \log L = g(\nu) - \nu n \sum_{j=1}^{m} \log \left\{ \frac{1}{n} \sum_{i=1}^{n} \frac{y_i(t_j)}{e^{\boldsymbol{a}_i'\boldsymbol{\beta}}} \right\} - \nu m \sum_{i=1}^{n} \boldsymbol{a}_i'\boldsymbol{\beta}$$

where $g(\nu)$ is the function of shape parameter ν. Note that the equation [6] does not include the nuisance baseline $\lambda(t)$. This means that estimates of $\boldsymbol{\beta}$ and ν can be obtained without modeling the shape of $\lambda(t)$. This is analogous to the proportional hazards model of Cox (1972) in survival analysis, which estimates regression coefficients without specifying the shape of the baseline hazard function. Equation [6] can be derived using an approach similar to the construction of partial likelihood in Cox (1975). The derivation for equation [6] is shown in the Appendix.

Based on asymptotic theories of generalized linear models (see e.g., McCullagh and Nelder 1989), the maximum likelihood estimator $\hat{\boldsymbol{\beta}}$ has an asymptotic normality:

$$[7] \qquad \hat{\boldsymbol{\beta}} \sim \mathrm{N}\left(\boldsymbol{\beta}, \boldsymbol{I}(\boldsymbol{\beta})^{-1}\right)$$

where $\boldsymbol{I}(\boldsymbol{\beta})$ is Fisher information matrix defined by:

$$[8] \qquad \boldsymbol{I}(\boldsymbol{\beta}) = -E\left[\frac{\partial^2 \log L}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}'}\right].$$

A partial differential and matrix calculation leads to:

$$[9] \qquad \boldsymbol{I}(\boldsymbol{\beta}) = \nu m \boldsymbol{A}'\boldsymbol{A}$$

where $\boldsymbol{A} = \begin{bmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \dots & \boldsymbol{a}_n \end{bmatrix}'$ is a design matrix of explanatory variables. The derivation for Fisher information matrix is shown in the Appendix.

Once the regression coefficient $\boldsymbol{\beta}$ is estimated, the nonparametric estimation of $\lambda(t)$ can be posteriorly obtained as:

$$[10] \qquad \hat{\lambda}(t_j) = \frac{1}{n} \sum_{i=1}^{n} \frac{y_i(t_j)}{e^{\boldsymbol{a}_i'\hat{\boldsymbol{\beta}}}}$$

by using $\hat{\boldsymbol{\beta}}$. Equation [10] can be interpreted as the covariate adjusted means of response for each measurement point in time.

2.2. Materials

The proposed method is applied to a forest growth data on individual trees of 23 years old Sugi (*Cryptomeria japonica*) plantation. The data was collected during the period of 2002-2003 at Hoshino village in the southeastern part of Fukuoka Prefecture of Japan (Yoshimoto et al. 2012). Figure 1 shows the location of Hoshino village in Japan. The village is covered with 84 percent mountain forests and is located at an altitude between 200 and 1000 meters. According to the weather database in Japan Meteorological Agency, the average annual precipitation between 1981 and 2010 was 2018.7 mm.

The growth data is made up of three variables on growth index: diameter at breast height (DBH), tree height and tree volume. Detailed information on the content and measurement method can be found in Yoshimoto et al (2012). Figure 2 shows the observed individual growth of DBH, tree height and volume.

From Figure 2, we can see that the growth behaviors of DBH and tree volume show a nonlinear trend. It can be seen that the variances of DBH and volume increase with age. This fails to meet the assumption in Kamo et al. (2017). We adopt the altitude for explanatory variable as in Kamo et al. (2017) and analyze the growth behaviors of these three variables using the proposed regression model:

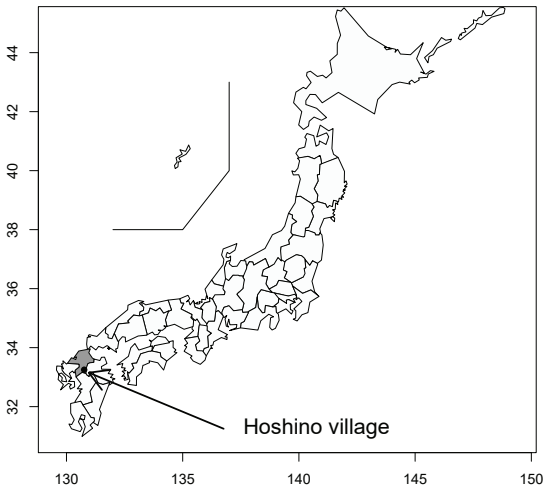


Figure 1. Location of Hoshino village in Japan.

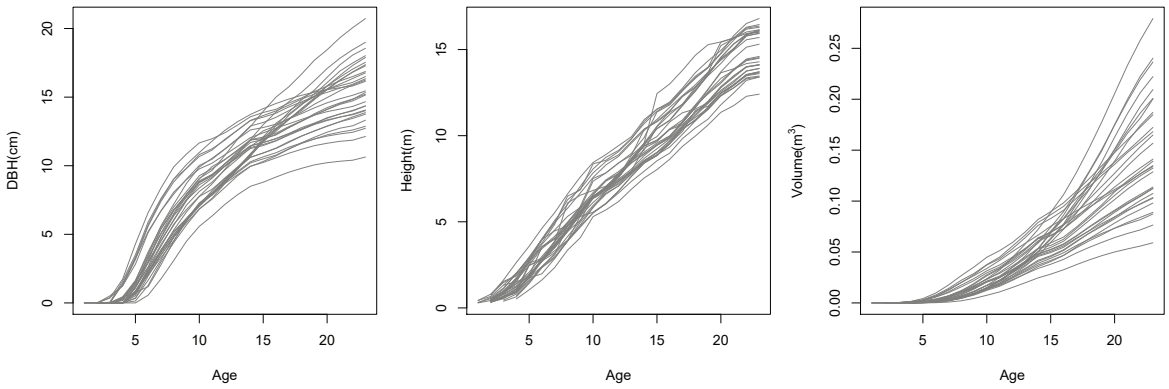


Figure 2. Observed individual growths of DBH, height and volume.

$$
\begin{aligned}
E\left[DBH_i(t)\right] &= \lambda_D(t)\exp\left(\beta_D \times Altitude_i\right) \\
E\left[Height_i(t)\right] &= \lambda_H(t)\exp\left(\beta_H \times Altitude_i\right) \\
E\left[Volume_i(t)\right] &= \lambda_V(t)\exp\left(\beta_V \times Altitude_i\right)
\end{aligned}
\qquad [11]
$$

where $DBH_i(t)$, $Height_i(t)$ and $Volume_i(t)$ are DBH, tree height and volume of sample i at age t; $\lambda_D(t)$, $\lambda_H(t)$ and $\lambda_V(t)$ correspond to the nuisance baseline growth of DBH, tree height and volume; β_D, β_H and β_V are effects of altitude on DBH, tree height and volume.

3. Results

Table 1 shows the maximum likelihood estimates of β_D without specifying the shape of the baseline growth $\lambda_D(t)$ for the growth data on DBH, tree height and volume. The effect of altitude on DBH growth was significantly negative, $\hat{\beta}_D = -0.177 (p < 0.001)$. Thus, DBH growth tends to decrease with increasing altitude and the relative ratio of DBH growth against one-unit difference in scaled altitude is estimated as: $\exp(-0.177) \approx 0.838$. This result is similar to those of Kamo et al. (2017). In their paper, they indicated two reasons for the negative effect of altitude: water and nutrients flow (Sands and Mulligan 1990), and strong winds in high altitude areas (King 1986). The results for both tree height and volume also show significantly negative effect of altitude on growth.

Table 1. Estimated regression coefficient for altitude.

Response	Estimate	Standard Error	p value
DBH	− 0.177	0.029	<0.001
Height	− 0.083	0.006	<0.001
Volume	− 0.250	0.020	<0.001

Table 2. Estimated dispersion and shape parameters.

Response	$\hat{\phi}$	$\hat{\nu}$
DBH	0.524	1.910
Height	0.023	43.268
Volume	0.265	3.767

Table 2 shows the estimates of dispersion parameter ϕ and shape parameter ν. The value of the shape parameter for tree height is large. Note that the gamma distribution can be approximated by normal distribution with mean μ and variance ϕ for large shape parameter ν (or small dispersion parameter ϕ). Therefore, from the growth data of tree height, it is also possible to fit a regression model with normal error in Kamo et al. (2017).

Using equation [10] with the maximum likelihood estimates of $\beta_D, \beta_H, \beta_V$ in Table 1, Figure 3 shows the estimated baseline growth for DBH, tree height and volume. Figure 4 shows the 95% confidence intervals for baseline growth constructed using the profile likelihood-based approach in Venzon and Moolgavkar (1988). The scaled altitude was used in fitting the regression model [11]. The estimated baseline growth corresponds to the growth behavior for a given sample with average altitude.

4. Discussion

In this paper we introduced a nuisance baseline into a gamma regression model and then developed a method of estimating regression coefficients without modeling the shape of baseline growth trend. The proposed method is particularly useful when the primary interest is centered on the effects of covariates. If the longitudinal baseline trend is also of interest, we can estimate it after the effects of covariates have been evaluated. It is sometimes difficult to determine an appropriate function to model the longitudinal growth trend. Under such a circumstance, it is necessary to obtain estimates of the baseline trend without making assumptions on their shapes.

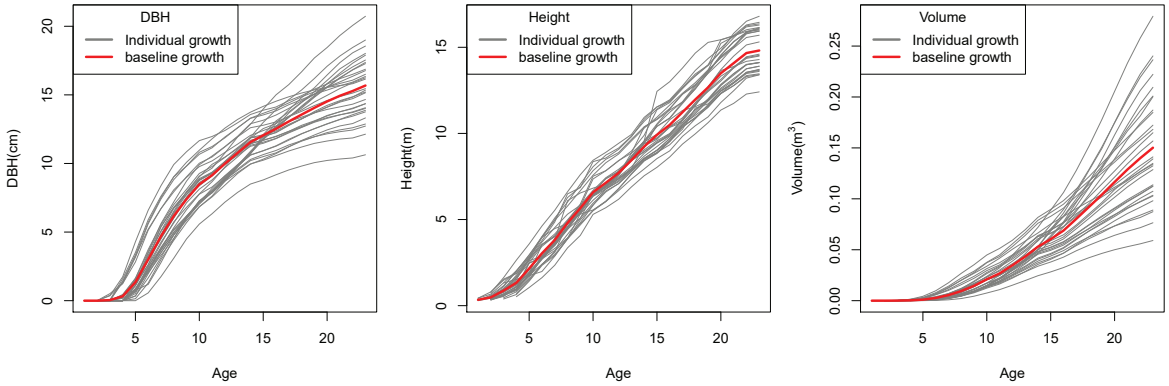


Figure 3. Estimated baseline growth of DBH, height and volume.

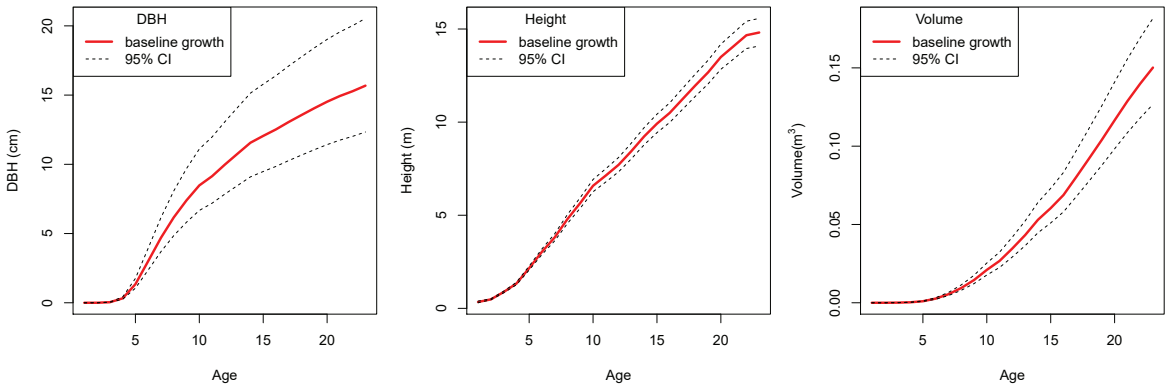


Figure 4. Confidence interval for baseline growth of DBH, height and volume.

One of the major applications of growth behavior prediction is in the selection of appropriate growth and yield model for deciding the timing of thinning and final harvest in forest management. The proposed model regards the baseline trend as nuisance parameters. Therefore, the future prediction of growth cannot be obtained only by this proposed method. However, we consider it useful to use the estimated baseline trend to fit a nonlinear function. This is because in the baseline trend, data is adjusted for the impact of environmental factors. If the information of covariates on environmental factors becomes "rich" (i.e., caused by global warming), the baseline trend is expected to contribute to the improvement of the accuracy of prediction by nonlinear functions.

We introduced the nuisance baseline trend under the framework of Gamma regression model and developed the method for estimating only the regression coefficients for the covariate of interest without assuming the shape of the nuisance baseline trend. We adopted the gamma distribution to handle random component in the regression model for non-negative data on DBH, tree height and volume. For positive data, use of other distributions, such as Lognormal, and Box-Cox transformations should be considered as alternatives. The proposed estimating method was constructed for the balanced data, in which all observed time points are common among samples with no missing observation. In the practical situations, forest growth data may contain missing observations and censoring due to death of measuring trees. Therefore, extension of the model for such unbalanced data, is a topic for future discussion.

Acknowledgement

The authors are grateful to the referee for his/her valuable comments and suggestions which improved an earlier version of the manuscript. Part of this research was funded by a Grant-in-Aid for Scientific Research from the Ministry of Education, Science, Sports, and Culture of Japan (No.17H00806, 17K09125, 20K06135).

References

Akaike, H. (1973) *Information theory and an extension of the maximum likelihood principle*, In: Csaki, F., Petrov, B.N. (eds.) Proceedings of the 2nd International Symposium on Information Theory; Budapest: Akademiai Kiado: 267-281.

Cox, D.R. (1972) Regression models and life-tables (with discussion), *J. R. Stat. Soc. Ser. B Methodol.* 34: 187-220.

Cox, D.R. (1975) Partial likelihood, *Biometrika* 62: 269-276.

Faraway, J.J. (2005) *Extending the linear model with R: Generalized linear, mixed effects and nonparametric regression models*, Chapman & Hall/CRC, Boca Raton London New York, 312p.

Gompertz, B. (1825) On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies, *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* 182:513-85.

Harris, T.E., Ross, F.S. (1955) *Fundamentals of a method for evaluating rail net capacities*, Research Memorandum, Rand Corporation.

Imada, M., Kunisaki, T., Mizoue, N., Teraoka Y. (1997) Optimum planting density for Japanese Oak (*Quercus mongolica* var. *grosseserrata*) based on spacing experiment with systematic design, *J. For. Res.* 2: 89-93.

Johnson, N.L., Kotz, S., Balakrishnan, N. (1994) *Continuous univariate distributions, Volume 1*, 2nd edition, John Wiley and Sons, New York, 784p.

Kamo, K. (2019) A new approach to classify growth patterns based on growth function selection and k-means method, *FORMATH* 18.003.

Kamo, K., Yoshimoto, A. (2013) Comparative analysis on selecting growth function based on three different information criteria for the purpose of carbon estimation, *For. Sci. Technol.* 9: 65-71.

Kamo, K., Tonda, T., Satoh, K. (2017) Growth analysis using nuisance baseline, *FORMATH* 16: 12-21.

King, D.A. (1986) Tree form, height growth, and susceptibility to wind damage in acer saccharum, *Ecology* 67: 980-990.

Kollo, T., von Rosen, D. (2005) *Advanced multivariate statistics with matrices*, Springer, Netherlands, 506p.

Mallows, C.L. (1973) Some comments of Cp., *Technometrics* 15: 661-675.

McCullagh, P., Nelder, J.A. (1989) *Generalized linear models, 2nd edition*, Chapman and Hall/CRC, Boca Raton London New York, 532p.

Nelder, J.A. (1962) New kinds of systematic designs for spacing experiments, *Biometrics* 18: 283-307.

Richards, F.J. (1958) A flexible growth function to empirical use, *J. Exp. Bot.* 10: 290-300.

Sands, R., Mulligan, R.S. (1990) Water and nutrient dynamics and tree growth, *For. Ecol. Manag.* 30: 91-111.

Satoh, K., Tonda, T. (2016) Estimating regression coefficients for balanced growth curve model when time trend of baseline is not specified, *Am. J. Math. Manag. Sci.* 35: 183-193

Stone, M. (1974) Cross-validatory choice and assessment of statistical predictions, *J. R. Stat. Soc. Ser. B Stat. Methodol.* 36: 111-147.

Stone, M. (1977) An asymptotic equivalence of choice of model by cross-validation and Akaike's criterion, *J. R. Stat. Soc. Ser. B Stat. Methodol.* 39: 44-47.

Tjørve, K.M.C., Tjørve, E. (2017) The use of Gompertz models in growth analyses, and new Gompertz-model approach: An addition to the Unified-Richards family, *PLoS One* 12:e0178691.

Venzon, D.J., Moolgavkar, S.H. (1988) A method for computing profile-likelihood based confidence intervals, *J. R. Stat. Soc. Ser. C Appl. Stat.* 37: 87-94.

Yoshimoto, A., Kamo, K., Yanagihara, H. (2012) *Environmental data analysis by R*, Asakura Syoten (in Japanese), 202p.

Zeide, B. (1993) Analysis of growth equation, *For. Sci.* 39: 594-616.

Appendix 1 (The derivation for log-likelihood function on $\boldsymbol{\beta}$ and ν)

We here assume that $y_i(t)$ is described by the following structure:

$$y_i(t) \sim \text{Gamma}\left(\mu_i(t), \nu\right), \quad \mu_i(t) = \lambda(t) e^{\boldsymbol{a}_i'\boldsymbol{\beta}}. \tag{12}$$

Then, for the observations $y_i(t)$ $(i = 1, ..., n; t = t_1, ..., t_m)$, the log-likelihood function is given by:

$$\begin{aligned} \log L &= \sum_{i=1}^{n} \sum_{t=t_1}^{t_m} \left\{ -\log \Gamma(\nu) + \nu \log \frac{\nu}{\mu_i(t)} + (\nu - 1) \log y_i(t) - \frac{\nu y_i(t)}{\mu_i(t)} \right\} \\ &= f(\nu) - \nu \sum_{i=1}^{n} \sum_{t=t_1}^{t_m} \left\{ \log \mu_i(t) + \frac{y_i(t)}{\mu_i(t)} \right\} \end{aligned} \tag{13}$$

where $f(\nu) = -nm \log \Gamma(\nu) + nm\nu \log \nu + (\nu - 1) \sum_{i=1}^{n} \sum_{t=t_1}^{t_m} \log y_i(t)$. For m measurement time points: $t_1, t_2, ..., t_m$, let $\lambda_1, \lambda_2, ..., \lambda_m$ be the nuisance parameters defined by $\lambda_j = \lambda(t_j), j = 1, 2, ..., m$. Then $\mu_i(t)$ can be expressed by:

$$\mu_i(t) = \sum_{j=1}^{m} \lambda_j \delta_j(t) e^{\boldsymbol{a}_i'\boldsymbol{\beta}} \tag{14}$$

where $\delta_j(t)$ denotes the indicator function given by:

$$\delta_j(t) = \begin{cases} 1, t = t_j \\ 0, t \neq t_j \end{cases}. \tag{15}$$

Substituting $\mu_i(t)$ into the equation [13] leads to:

$$\begin{aligned} \log L &= f(\nu) - \nu \sum_{i=1}^{n} \sum_{t=t_1}^{t_m} \left\{ \log \left\{ \sum_{j=1}^{m} \lambda_j \delta_j(t) e^{\boldsymbol{a}_i'\boldsymbol{\beta}} \right\} + \left\{ \sum_{j=1}^{m} \lambda_j \delta_j(t) e^{\boldsymbol{a}_i'\boldsymbol{\beta}} \right\}^{-1} y_i(t) \right\} \\ &= f(\nu) - \nu \sum_{i=1}^{n} \sum_{t=t_1}^{t_m} \left\{ \sum_{j=1}^{m} \log \left(\lambda_j e^{\boldsymbol{a}_i'\boldsymbol{\beta}} \right) \delta_j(t) + \sum_{j=1}^{m} \frac{y_i(t)}{\lambda_j e^{\boldsymbol{a}_i'\boldsymbol{\beta}}} \delta_j(t) \right\} \\ &= f(\nu) - \nu \sum_{j=1}^{m} \sum_{t=t_1}^{t_m} \sum_{i=1}^{n} \delta_j(t) \left\{ \log \lambda_j + \boldsymbol{a}_i'\boldsymbol{\beta} + \frac{y_i(t)}{\lambda_j e^{\boldsymbol{a}_i'\boldsymbol{\beta}}} \right\} \\ &= f(\nu) - \nu \left\{ n \sum_{j=1}^{m} \log \lambda_j + m \sum_{i=1}^{n} \boldsymbol{a}_i'\boldsymbol{\beta} + \sum_{j=1}^{m} \frac{1}{\lambda_j} \sum_{i=1}^{n} \frac{y_i(t_j)}{e^{\boldsymbol{a}_i'\boldsymbol{\beta}}} \right\}. \end{aligned} \tag{16}$$

By differentiating the log-likelihood function [16] with λ_j, the following equation is obtained.

$$\frac{\partial}{\partial \lambda_j} \log L = -\nu \left\{ \frac{n}{\lambda_j} - \frac{1}{\lambda_j^2} \sum_{i=1}^{n} \frac{y_i\left(t_j\right)}{e^{\boldsymbol{a}_i'\boldsymbol{\beta}}} \right\} \tag{17}$$

Solving $\partial \log L / \partial \lambda_j = 0$ with λ_j leads to:

$$\lambda_j = \frac{1}{n} \sum_{i=1}^{n} \frac{y_i(t_j)}{e^{\boldsymbol{a}_i'\boldsymbol{\beta}}}. \tag{18}$$

Substituting this into the equation [16], the log-likelihood function only on $\boldsymbol{\beta}$ and ν is obtained:

$$[19] \qquad \log L = g(\nu) - \nu n \sum_{j=1}^{m} \log \left\{ \frac{1}{n} \sum_{i=1}^{n} \frac{y_i\left(t_j\right)}{e^{\boldsymbol{a}_i'\boldsymbol{\beta}}} \right\} - \nu m \sum_{i=1}^{n} \boldsymbol{a}_i'\boldsymbol{\beta}$$

where $g(\nu) = f(\nu) - nm\nu$.

Appendix 2 (The derivation for Fisher information matrix)

Differentiating the equation [13] with β_j and β_k leads to:

$$[20] \qquad \frac{\partial \log L}{\partial \beta_j} = \nu \sum_{i=1}^{n} \sum_{t=t_1}^{t_m} \left\{ \frac{y_i(t) a_{ij}}{\mu_i(t)} - a_{ij} \right\}, \quad \frac{\partial^2 \log L}{\partial \beta_j \partial \beta_k} = \nu \sum_{i=1}^{n} \sum_{t=t_1}^{t_m} \frac{y_i(t) a_{ij} a_{ik}}{\mu_i(t)}.$$

Then the (j,k) element of Fisher information matrix can be obtained as:

$$[21] \qquad -E\left[\frac{\partial^2 \log L}{\partial \beta_j \partial \beta_k}\right] = \nu m \sum_{i=1}^{n} a_{ij} a_{ik}.$$

Let $\boldsymbol{A}$ be the design matrix of explanatory variables given by $\boldsymbol{A} = \left[\begin{array}{cccc} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \ldots & \boldsymbol{a}_n \end{array} \right]'$, the Fisher information matrix can be obtained as:

$$[22] \qquad \boldsymbol{I}(\boldsymbol{\beta}) = \nu m \boldsymbol{A}'\boldsymbol{A}.$$