State Space Models for Binary Response Data with Flexible Link Functions

Dipak K. Dey

University of Connecticut, USA,
Joint with Xun Jiang, University of Connecticut, USA,
Carlos Abanto Valle, Federal University of Rio de Janieiro, Brazil

Non-Gaussian Multivariate Statistical Models and their Applications,
Banff International Research Station,
May, 2013

Outline

1. Introduction
2. Binary State Space Models with Flexible Link Functions
3. Simulation Study
4. Case Study
5. Conclusion

Introduction

- Binary state space models (BSSM) using a flexible skewed link functions are introduced in this paper.
- A critical issue in modeling binary response data is the choice of the link function. Typically, link functions with fixed skewnewss (logit, cloglog, loglog) are adopted.
- However, commonly used logit, cloglog and loglog links are prone to link misspecification because of their fixed skewness.
- Flexible link functions are desired in modeling binary state space models to let the data tell how much skewness should be incorporated.

Introduction

- Binary state space models (BSSM) using a flexible skewed link functions are introduced in this paper.
- A critical issue in modeling binary response data is the choice of the link function. Typically, link functions with fixed skewnewss (logit, cloglog, loglog) are adopted.
- However, commonly used logit, cloglog and loglog links are prone to link misspecification because of their fixed skewness.
- Flexible link functions are desired in modeling binary state space models to let the data tell how much skewness should be incorporated.

Introduction

- Binary state space models (BSSM) using a flexible skewed link functions are introduced in this paper.
- A critical issue in modeling binary response data is the choice of the link function. Typically, link functions with fixed skewnewss (logit, cloglog, loglog) are adopted.
- However, commonly used logit, cloglog and loglog links are prone to link misspecification because of their fixed skewness.
- Flexible link functions are desired in modeling binary state space models to let the data tell how much skewness should be incorporated.

Introduction

- Binary state space models (BSSM) using a flexible skewed link functions are introduced in this paper.
- A critical issue in modeling binary response data is the choice of the link function. Typically, link functions with fixed skewnewss (logit, cloglog, loglog) are adopted.
- However, commonly used logit, cloglog and loglog links are prone to link misspecification because of their fixed skewness.
- Flexible link functions are desired in modeling binary state space models to let the data tell how much skewness should be incorporated.

Introduction

- Binary state space models (BSSM) using a flexible skewed link functions are introduced in this paper.
- A critical issue in modeling binary response data is the choice of the link function. Typically, link functions with fixed skewnewss (logit, cloglog, loglog) are adopted.
- However, commonly used logit, cloglog and loglog links are prone to link misspecification because of their fixed skewness.
- Flexible link functions are desired in modeling binary state space models to let the data tell how much skewness should be incorporated.

Literature Review

- There are lots of research done to introduce flexibility of skewness as well as tail behavior into the link functions.
- Stukel (1988) proposed a two-parameter class of generalized logistic models; Kim et al. (2008) used the skewed generalized t-link; And Bazan et al. (2010) adopted the skewed probit links and some variants with different parameterizations.
- Wang and Dey (2010), Wang and Dey (2011) and Jiang et al. (2012) introduced the GEV and SPLOGIT models for the binary cross sectional data, which includes many standard link functions as special cases.
- In state space model literature, Czado and Song (2008) introduced binary state space models with probit link as binary state space mixed models (BSSM). More recently, Abanto-Valle and Dey (2012) extended it to scale mixture of normal (SMN) links.

Literature Review

- There are lots of research done to introduce flexibility of skewness as well as tail behavior into the link functions.
- Stukel (1988) proposed a two-parameter class of generalized logistic models; Kim et al. (2008) used the skewed generalized t-link; And Bazan et al. (2010) adopted the skewed probit links and some variants with different parameterizations.
- Wang and Dey (2010), Wang and Dey (2011) and Jiang et al. (2012) introduced the GEV and SPLOGIT models for the binary cross sectional data, which includes many standard link functions as special cases.
- In state space model literature, Czado and Song (2008) introduced binary state space models with probit link as binary state space mixed models (BSSM). More recently, Abanto-Valle and Dey (2012) extended it to scale mixture of normal (SMN) links.

Literature Review

- There are lots of research done to introduce flexibility of skewness as well as tail behavior into the link functions.
- Stukel (1988) proposed a two-parameter class of generalized logistic models; Kim et al. (2008) used the skewed generalized t-link; And Bazan et al. (2010) adopted the skewed probit links and some variants with different parameterizations.
- Wang and Dey (2010), Wang and Dey (2011) and Jiang et al. (2012) introduced the GEV and SPLOGIT models for the binary cross sectional data, which includes many standard link functions as special cases.
- In state space model literature, Czado and Song (2008) introduced binary state space models with probit link as binary state space mixed models (BSSM). More recently, Abanto-Valle and Dey (2012) extended it to scale mixture of normal (SMN) links.

Literature Review

- There are lots of research done to introduce flexibility of skewness as well as tail behavior into the link functions.
- Stukel (1988) proposed a two-parameter class of generalized logistic models; Kim et al. (2008) used the skewed generalized t-link; And Bazan et al. (2010) adopted the skewed probit links and some variants with different parameterizations.
- Wang and Dey (2010), Wang and Dey (2011) and Jiang et al. (2012) introduced the GEV and SPLOGIT models for the binary cross sectional data, which includes many standard link functions as special cases.
- In state space model literature, Czado and Song (2008) introduced binary state space models with probit link as binary state space mixed models (BSSM). More recently, Abanto-Valle and Dey (2012) extended it to scale mixture of normal (SMN) links.

Literature Review

- There are lots of research done to introduce flexibility of skewness as well as tail behavior into the link functions.
- Stukel (1988) proposed a two-parameter class of generalized logistic models; Kim et al. (2008) used the skewed generalized t-link; And Bazan et al. (2010) adopted the skewed probit links and some variants with different parameterizations.
- Wang and Dey (2010), Wang and Dey (2011) and Jiang et al. (2012) introduced the GEV and SPLOGIT models for the binary cross sectional data, which includes many standard link functions as special cases.
- In state space model literature, Czado and Song (2008) introduced binary state space models with probit link as binary state space mixed models (BSSM). More recently, Abanto-Valle and Dey (2012) extended it to scale mixture of normal (SMN) links.

Binary State Space Models Setup

- Suppose we have a binary time series $\{Y_t, t = 1, ..., T\}$. The probability of success at each time t is π_t, which is further related with a time-varying covariates vector $\mathbf{x}_t = (x_{t1}, ..., x_{tk})'$ with coefficients β and a latent state variable $\boldsymbol{\theta}_t$.
- A Generalized linear state space model can be constructed in the following way

$$\begin{aligned} Y_t &\sim \mathcal{B}er(\pi_t) \qquad\qquad t = 1, \ldots, T \\ \pi_t &= F(\mathbf{x}_t'\beta + \mathbf{S}_t'\boldsymbol{\theta}_t) \\ \boldsymbol{\theta}_t &= \mathbf{H}_t\boldsymbol{\theta}_{t-1} + \boldsymbol{\eta}_t, \quad \boldsymbol{\eta}_t \sim \mathcal{N}_q(\mathbf{0}, \mathbf{W}_t). \end{aligned}$$

- The above process is defined as a first order Markov process, where $\mathbf{H}_t$ is the transition matrix, $\mathbf{W}_t$ is the covariance matrix of error term η_t. In the framework of generalized linear models (McCullagh and Nelder, 1989), F is a cumulative distribution function and F^{-1} is the corresponding link function.

Binary State Space Models Setup

- Suppose we have a binary time series $\{Y_t, t = 1, ..., T\}$. The probability of success at each time t is π_t, which is further related with a time-varying covariates vector $\mathbf{x}_t = (x_{t1}, ..., x_{tk})'$ with coefficients $\boldsymbol{\beta}$ and a latent state variable $\boldsymbol{\theta}_t$.
- A Generalized linear state space model can be constructed in the following way

$$
\begin{aligned}
Y_t &\sim \mathcal{B}er(\pi_t) \qquad\qquad t = 1, \ldots, T \\
\pi_t &= F(\mathbf{x}_t'\beta + \mathbf{S}_t'\boldsymbol{\theta}_t) \\
\boldsymbol{\theta}_t &= \mathbf{H}_t\boldsymbol{\theta}_{t-1} + \boldsymbol{\eta}_t, \quad \boldsymbol{\eta}_t \sim \mathcal{N}_q(\mathbf{0}, \mathbf{W}_t).
\end{aligned}
$$

- The above process is defined as a first order Markov process, where $\mathbf{H}_t$ is the transition matrix, $\mathbf{W}_t$ is the covariance matrix of error term η_t. In the framework of generalized linear models (McCullagh and Nelder, 1989), F is a cumulative distribution function and F^{-1} is the corresponding link function.

Binary State Space Models Setup

- Suppose we have a binary time series $\{Y_t, t = 1, ..., T\}$. The probability of success at each time t is π_t, which is further related with a time-varying covariates vector $\mathbf{x}_t = (x_{t1}, ..., x_{tk})'$ with coefficients $\boldsymbol{\beta}$ and a latent state variable $\boldsymbol{\theta}_t$.
- A Generalized linear state space model can be constructed in the following way

$$
\begin{aligned}
Y_t &\sim \mathcal{B}er(\pi_t) \qquad\qquad t = 1, \ldots, T \\
\pi_t &= F(\mathbf{x}_t'\beta + \mathbf{S}_t'\boldsymbol{\theta}_t) \\
\boldsymbol{\theta}_t &= \mathbf{H}_t\boldsymbol{\theta}_{t-1} + \boldsymbol{\eta}_t, \quad \boldsymbol{\eta}_t \sim \mathcal{N}_q(\mathbf{0}, \mathbf{W}_t).
\end{aligned}
$$

- The above process is defined as a first order Markov process, where $\mathbf{H}_t$ is the transition matrix, $\mathbf{W}_t$ is the covariance matrix of error term η_t. In the framework of generalized linear models (McCullagh and Nelder, 1989), F is a cumulative distribution function and F^{-1} is the corresponding link function.

Binary State Space Models Setup

- Suppose we have a binary time series $\{Y_t, t = 1, ..., T\}$. The probability of success at each time t is π_t, which is further related with a time-varying covariates vector $\mathbf{x}_t = (x_{t1}, ..., x_{tk})'$ with coefficients $\boldsymbol{\beta}$ and a latent state variable $\boldsymbol{\theta}_t$.
- A Generalized linear state space model can be constructed in the following way

$$
\begin{aligned}
Y_t &\sim \mathcal{B}er(\pi_t) \qquad\qquad t = 1, \ldots, T \\
\pi_t &= F(\mathbf{x}_t'\beta + \mathbf{S}_t'\boldsymbol{\theta}_t) \\
\boldsymbol{\theta}_t &= \mathbf{H}_t\boldsymbol{\theta}_{t-1} + \boldsymbol{\eta}_t, \quad \boldsymbol{\eta}_t \sim \mathcal{N}_q(\mathbf{0}, \mathbf{W}_t).
\end{aligned}
$$

- The above process is defined as a first order Markov process, where $\mathbf{H}_t$ is the transition matrix, $\mathbf{W}_t$ is the covariance matrix of error term η_t. In the framework of generalized linear models (McCullagh and Nelder, 1989), F is a cumulative distribution function and F^{-1} is the corresponding link function.

Binary State Space Models Setup

- In this paper, we compare the BSSM by assuming F^{-1} to be three standard link functions and three flexible link functions.

- The three standard links we consider here are logit, cloglog and loglog and we call the corresponding state space model BSSM-LOGIT, BSSM-CLOGLOG and BSSM-LOGLOG.

- We also consider three flexible link models which will be elaborated in the next couple of slides: Generalized Extreme Value model (Wang and Dey (2010, 2011)), Symmetric Power Logit model (Jiang et al. (2012)) and Scale Mixture of Nomral model (Abanto-Valle and Dey (2012)).

Binary State Space Models Setup

- In this paper, we compare the BSSM by assuming F^{-1} to be three standard link functions and three flexible link functions.
- The three standard links we consider here are logit, cloglog and loglog and we call the corresponding state space model BSSM-LOGIT, BSSM-CLOGLOG and BSSM-LOGLOG.
- We also consider three flexible link models which will be elaborated in the next couple of slides: Generalized Extreme Value model (Wang and Dey (2010, 2011)), Symmetric Power Logit model (Jiang et al. (2012)) and Scale Mixture of Nomral model (Abanto-Valle and Dey (2012)).

Binary State Space Models Setup

- In this paper, we compare the BSSM by assuming F^{-1} to be three standard link functions and three flexible link functions.
- The three standard links we consider here are logit, cloglog and loglog and we call the corresponding state space model BSSM-LOGIT, BSSM-CLOGLOG and BSSM-LOGLOG.
- We also consider three flexible link models which will be elaborated in the next couple of slides: Generalized Extreme Value model (Wang and Dey (2010, 2011)), Symmetric Power Logit model (Jiang et al. (2012)) and Scale Mixture of Nomral model (Abanto-Valle and Dey (2012)).

Binary State Space Models Setup

- In this paper, we compare the BSSM by assuming F^{-1} to be three standard link functions and three flexible link functions.
- The three standard links we consider here are logit, cloglog and loglog and we call the corresponding state space model BSSM-LOGIT, BSSM-CLOGLOG and BSSM-LOGLOG.
- We also consider three flexible link models which will be elaborated in the next couple of slides: Generalized Extreme Value model (Wang and Dey (2010, 2011)), Symmetric Power Logit model (Jiang et al. (2012)) and Scale Mixture of Nomral model (Abanto-Valle and Dey (2012)).

Generalized Extreme Value Link

- The GEV link models are based on the Generalized Extreme Value (GEV) distribution with distribution function as

$$G(x) = \exp\left[-\left\{1+\xi\frac{x-\mu}{\sigma}\right\}_{+}^{-\frac{1}{\xi}}\right],$$

  where $\mu \in R$ is the location parameter, $\sigma \in R^{+}$ is the scale parameter, $\xi \in R$ is the shape parameter and $x_{+} = \max(x, 0)$.

- A more detailed discussion on the extreme value distributions can be found in Coles (2001) and Smith (2003).

- Extreme value analysis finds wide application in many areas, including climatology, environmental science, risk management and biomedical data processing.

Generalized Extreme Value Link

- The GEV link models are based on the Generalized Extreme Value (GEV) distribution with distribution function as

$$G(x) = \exp\left[-\left\{1+\xi\frac{x-\mu}{\sigma}\right\}_{+}^{-\frac{1}{\xi}}\right],$$

where $\mu \in R$ is the location parameter, $\sigma \in R^{+}$ is the scale parameter, $\xi \in R$ is the shape parameter and $x_{+} = \max(x, 0)$.

- A more detailed discussion on the extreme value distributions can be found in Coles (2001) and Smith (2003).
- Extreme value analysis finds wide application in many areas, including climatology, environmental science, risk management and biomedical data processing.

Generalized Extreme Value Link

- The GEV link models are based on the Generalized Extreme Value (GEV) distribution with distribution function as

$$G(x) = \exp\left[-\left\{1+\xi\frac{x-\mu}{\sigma}\right\}_{+}^{-\frac{1}{\xi}}\right],$$

where $\mu \in R$ is the location parameter, $\sigma \in R^{+}$ is the scale parameter, $\xi \in R$ is the shape parameter and $x_{+} = \max(x, 0)$.

- A more detailed discussion on the extreme value distributions can be found in Coles (2001) and Smith (2003).

- Extreme value analysis finds wide application in many areas, including climatology, environmental science, risk management and biomedical data processing.

Generalized Extreme Value Link

- The GEV link models are based on the Generalized Extreme Value (GEV) distribution with distribution function as

$$G(x) = \exp\left[-\left\{1+\xi\frac{x-\mu}{\sigma}\right\}_{+}^{-\frac{1}{\xi}}\right],$$

where $\mu \in R$ is the location parameter, $\sigma \in R^{+}$ is the scale parameter, $\xi \in R$ is the shape parameter and $x_{+} = \max(x, 0)$.

- A more detailed discussion on the extreme value distributions can be found in Coles (2001) and Smith (2003).

- Extreme value analysis finds wide application in many areas, including climatology, environmental science, risk management and biomedical data processing.

Generalized Extreme Value Link

- The GEV distribution serves naturally as a flexible link function as the shape parameter ξ purely controls the tail behavior of the distribution (Wang and Dey, 2010, 2011).
- Figure 1 provides a comparison of pdf and cdf plots of the GEV class with different ξ to show the flexibility of such distributions.
- Here we adopt Arnold and Groeneveld (1995)'s measure of skewness measure, which varies between -1 and 1 with 0 indicates symmetry. The GEV distribution is negatively skewed for $X < -0.307$ and positively skewed for $X > -0.307$.

Generalized Extreme Value Link

- The GEV distribution serves naturally as a flexible link function as the shape parameter ξ purely controls the tail behavior of the distribution (Wang and Dey, 2010, 2011).

- Figure 1 provides a comparison of pdf and cdf plots of the GEV class with different ξ to show the flexibility of such distributions.

- Here we adopt Arnold and Groeneveld (1995)'s measure of skewness measure, which varies between -1 and 1 with 0 indicates symmetry. The GEV distribution is negatively skewed for $X < -0.307$ and positively skewed for $X > -0.307$.

Generalized Extreme Value Link

- The GEV distribution serves naturally as a flexible link function as the shape parameter ξ purely controls the tail behavior of the distribution (Wang and Dey, 2010, 2011).
- Figure 1 provides a comparison of pdf and cdf plots of the GEV class with different ξ to show the flexibility of such distributions.
- Here we adopt Arnold and Groeneveld (1995)'s measure of skewness measure, which varies between -1 and 1 with 0 indicates symmetry. The GEV distribution is negatively skewed for $X < -0.307$ and positively skewed for $X > -0.307$.

Generalized Extreme Value Link

- The GEV distribution serves naturally as a flexible link function as the shape parameter ξ purely controls the tail behavior of the distribution (Wang and Dey, 2010, 2011).

- Figure 1 provides a comparison of pdf and cdf plots of the GEV class with different ξ to show the flexibility of such distributions.

- Here we adopt Arnold and Groeneveld (1995)'s measure of skewness measure, which varies between -1 and 1 with 0 indicates symmetry. The GEV distribution is negatively skewed for $X < -0.307$ and positively skewed for $X > -0.307$.

GEV Plots

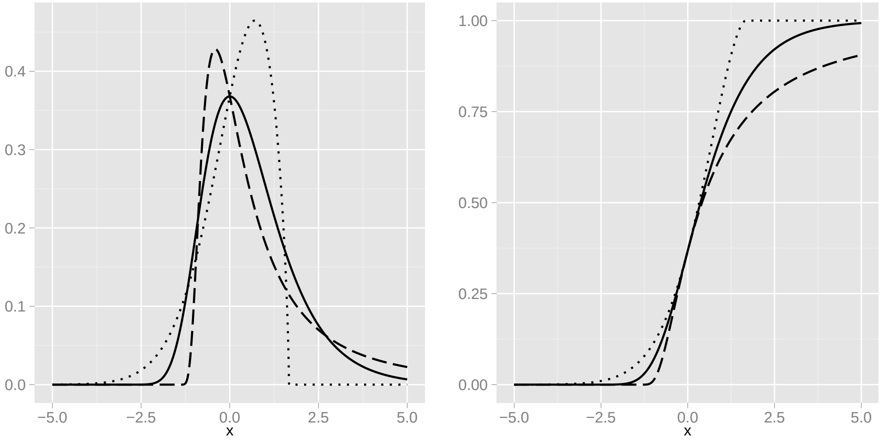


Figure : Left: pdf plot of GEV distribution. Right: cdf plot of GEV distribution. Solid line ($\xi = 0$), dashed line ($\xi = 0.6$), and dotted line ($\xi = -0.6$).

Symmetric Power Logit Link

- Suppose F_0 is a baseline link function for which the pdf is symmetric about zero, the symmetric power distribution based on F_0 is defined as

$$F(x, r) = F_0^r(\frac{x}{r})\mathbf{I}_{(0,1]}(r) + \left(1 - F_0^{\frac{1}{r}}(-rx)\right)\mathbf{I}_{(1,+\infty)}(r),$$

  where $\mathbf{I}$ is the indicator function.

- The intuition for symmetric power distribution is that $F_0^r(x)$ is a valid cdf and achieves flexible left skewness when $r < 1$, while the same property holds for its mirror reflection $1 - F_0^{\frac{1}{r}}(-x)$ with skewness being in the opposite direction.

- The proposed family includes F_0 as a special case with $r = 1$. The pdf of the symmetric power family with power parameter r is the mirror image of the pdf with power parameter $\frac{1}{r}$ since $F(x, r) = 1 - F(-x, \frac{1}{r})$.

Symmetric Power Logit Link

- Suppose F_0 is a baseline link function for which the pdf is symmetric about zero, the symmetric power distribution based on F_0 is defined as

$$F(x,r) = F_0^r(\frac{x}{r})\mathbf{I}_{(0,1]}(r) + \left(1 - F_0^{\frac{1}{r}}(-rx)\right)\mathbf{I}_{(1,+\infty)}(r),$$

where $\mathbf{I}$ is the indicator function.

- The intuition for symmetric power distribution is that $F_0^r(x)$ is a valid cdf and achieves flexible left skewness when $r < 1$, while the same property holds for its mirror reflection $1 - F_0^{\frac{1}{r}}(-x)$ with skewness being in the opposite direction.

- The proposed family includes F_0 as a special case with $r = 1$. The pdf of the symmetric power family with power parameter r is the mirror image of the pdf with power parameter $\frac{1}{r}$ since $F(x,r) = 1 - F(-x, \frac{1}{r})$.

Symmetric Power Logit Link

- Suppose F_0 is a baseline link function for which the pdf is symmetric about zero, the symmetric power distribution based on F_0 is defined as

$$F(x,r) = F_0^r(\frac{x}{r})\mathbf{I}_{(0,1]}(r) + \left(1 - F_0^{\frac{1}{r}}(-rx)\right)\mathbf{I}_{(1,+\infty)}(r),$$

  where $\mathbf{I}$ is the indicator function.

- The intuition for symmetric power distribution is that $F_0^r(x)$ is a valid cdf and achieves flexible left skewness when $r < 1$, while the same property holds for its mirror reflection $1 - F_0^{\frac{1}{r}}(-x)$ with skewness being in the opposite direction.

- The proposed family includes F_0 as a special case with $r = 1$. The pdf of the symmetric power family with power parameter r is the mirror image of the pdf with power parameter $\frac{1}{r}$ since $F(x,r) = 1 - F(-x, \frac{1}{r})$.

Symmetric Power Logit Link

- Suppose F_0 is a baseline link function for which the pdf is symmetric about zero, the symmetric power distribution based on F_0 is defined as

$$F(x,r) = F_0^r(\frac{x}{r})\mathbf{I}_{(0,1]}(r) + \left(1 - F_0^{\frac{1}{r}}(-rx)\right)\mathbf{I}_{(1,+\infty)}(r),$$

where $\mathbf{I}$ is the indicator function.

- The intuition for symmetric power distribution is that $F_0^r(x)$ is a valid cdf and achieves flexible left skewness when $r < 1$, while the same property holds for its mirror reflection $1 - F_0^{\frac{1}{r}}(-x)$ with skewness being in the opposite direction.

- The proposed family includes F_0 as a special case with $r = 1$. The pdf of the symmetric power family with power parameter r is the mirror image of the pdf with power parameter $\frac{1}{r}$ since $F(x,r) = 1 - F(-x, \frac{1}{r})$.

Symmetric Power Logit Link

- Considering F_0 as the cdf of a logistic distribution which will lead us to SPLOGIT link adopted in this paper. Clearly, the skewness of SPLOGIT distribution can be adjusted from its baseline to achieve more flexibility as r varies.

- Similar ideas of bringing a power parameter into the cdf to allow more flexibility can be found in Nagler (1994), Samejima (2000) and Gupta and Gupta (2008).

- However, the construction of SPLOGIT ensures the flexibility is achieved symmetrically with respect to $r = 1$ and thus can accomondate greater skewnewss for both directions by choosing appropriate tail.

Symmetric Power Logit Link

- Considering F_0 as the cdf of a logistic distribution which will lead us to SPLOGIT link adopted in this paper. Clearly, the skewness of SPLOGIT distribution can be adjusted from its baseline to achieve more flexibility as r varies.

- Similar ideas of bringing a power parameter into the cdf to allow more flexibility can be found in Nagler (1994), Samejima (2000) and Gupta and Gupta (2008).

- However, the construction of SPLOGIT ensures the flexibility is achieved symmetrically with respect to $r = 1$ and thus can accomondate greater skewnewss for both directions by choosing appropriate tail.

Symmetric Power Logit Link

- Considering F_0 as the cdf of a logistic distribution which will lead us to SPLOGIT link adopted in this paper. Clearly, the skewness of SPLOGIT distribution can be adjusted from its baseline to achieve more flexibility as r varies.
- Similar ideas of bringing a power parameter into the cdf to allow more flexibility can be found in Nagler (1994), Samejima (2000) and Gupta and Gupta (2008).
- However, the construction of SPLOGIT ensures the flexibility is achieved symmetrically with respect to $r = 1$ and thus can accomondate greater skewnewss for both directions by choosing appropriate tail.

Symmetric Power Logit Link

- Considering F_0 as the cdf of a logistic distribution which will lead us to SPLOGIT link adopted in this paper. Clearly, the skewness of SPLOGIT distribution can be adjusted from its baseline to achieve more flexibility as r varies.

- Similar ideas of bringing a power parameter into the cdf to allow more flexibility can be found in Nagler (1994), Samejima (2000) and Gupta and Gupta (2008).

- However, the construction of SPLOGIT ensures the flexibility is achieved symmetrically with respect to $r = 1$ and thus can accomondate greater skewnewss for both directions by choosing appropriate tail.

SPLOGIT Plots

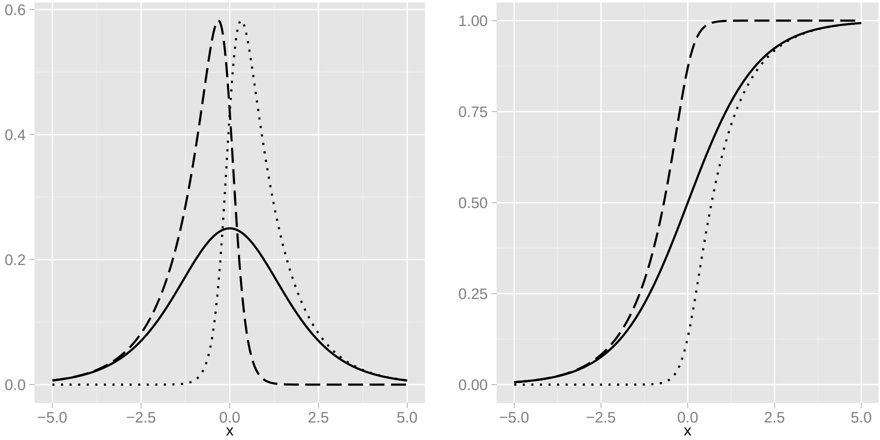


Figure : Left: pdf plot of SPLOGIT distribution. Right: cdf plot of SPLOGIT distribution. Solid line ($r = 1$), dashed line ($r = 0.2$), and dotted line ($r = 5$).

Scale Mixture of Normal Link

- A random variable Y is said to follow the scale mixture of normal (SMN) distribution if and only if it can be expressed as

$$Y = \mu + \kappa(\lambda)^{1/2} X,$$

where μ is the location parameter, $\kappa(.)$ is a positive weight function and λ is a positive mixing random variable.

- We follow Lange and Sinsheimer (1993) and let $\kappa(\lambda) = 1/\lambda$, then it is clear that $Y|\lambda \sim \mathcal{N}(\mu, \lambda^{-1}\sigma^2)$ and the pdf of Y given ν is given by

$$f_{SMN}(y|\mu, \sigma^2, \nu) = \int_{-\infty}^{\infty} \phi(y|\mu, \lambda^{-1}\sigma^2) dH(\lambda|\nu),$$

where $\phi(.|\mu, \sigma^2)$ denotes the pdf of $\mathcal{N}(\mu, \sigma^2)$ and $H(.|\nu)$ is the cdf of λ.

Scale Mixture of Normal Link

- A random variable Y is said to follow the scale mixture of normal (SMN) distribution if and only if it can be expressed as

$$Y = \mu + \kappa(\lambda)^{1/2} X,$$

where μ is the location parameter, $\kappa(.)$ is a positive weight function and λ is a positive mixing random variable.

- We follow Lange and Sinsheimer (1993) and let $\kappa(\lambda) = 1/\lambda$, then it is clear that $Y|\lambda \sim \mathcal{N}(\mu, \lambda^{-1}\sigma^2)$ and the pdf of Y given ν is given by

$$f_{SMN}(y|\mu, \sigma^2, \nu) = \int_{-\infty}^{\infty} \phi(y|\mu, \lambda^{-1}\sigma^2) dH(\lambda|\nu),$$

where $\phi(.|\mu, \sigma^2)$ denotes the pdf of $\mathcal{N}(\mu, \sigma^2)$ and $H(.|\nu)$ is the cdf of λ.

Scale Mixture of Normal Link

- A random variable Y is said to follow the scale mixture of normal (SMN) distribution if and only if it can be expressed as

$$Y = \mu + \kappa(\lambda)^{1/2} X,$$

where μ is the location parameter, $\kappa(.)$ is a positive weight function and λ is a positive mixing random variable.

- We follow Lange and Sinsheimer (1993) and let $\kappa(\lambda) = 1/\lambda$, then it is clear that $Y|\lambda \sim \mathcal{N}(\mu, \lambda^{-1}\sigma^2)$ and the pdf of Y given ν is given by

$$f_{SMN}(y|\mu, \sigma^2, \nu) = \int_{-\infty}^{\infty} \phi(y|\mu, \lambda^{-1}\sigma^2) dH(\lambda|\nu),$$

where $\phi(.|\mu, \sigma^2)$ denotes the pdf of $\mathcal{N}(\mu, \sigma^2)$ and $H(.|\nu)$ is the cdf of λ.

Scale Mixture of Normal Link

- Based on the pdf, the cdf of the SMN distribution is given by

$$\begin{aligned} F_{SMN}(y|\mu,\sigma^2,\nu) &= \int_{-\infty}^{y}\int_{-\infty}^{\infty}\phi(u|\mu,\lambda^{-1}\sigma^2)dH(\lambda|\nu)du \\ &= \int_{-\infty}^{\infty}\Phi(\frac{\lambda^{1/2}[y-\mu]}{\sigma})dH(\lambda|\nu), \end{aligned}$$

  where $\Phi(.)$ is the cdf of the standard normal distribution.

- Special cases of the SMN family include normal (by fixing $\lambda = 1$), Student-t (by setting $\lambda \sim \mathcal{G}(\frac{\nu}{2},\frac{\nu}{2})$) and Slash (by setting $\lambda \sim \mathcal{B}e(\nu,1)$), where $\mathcal{G}(.,.)$ and $\mathcal{B}e(.,.)$ denote the gamma and distribution respectively.

- Notice that unlike the GEV and SPLOGIT distributions, SMN distribution is always symmetric, which means the flexibility of the link only exists in tail behavior but not in skewness adjustment.

Scale Mixture of Normal Link

- Based on the pdf, the cdf of the SMN distribution is given by

$$F_{SMN}(y|\mu,\sigma^2,\nu) = \int_{-\infty}^{y}\int_{-\infty}^{\infty}\phi(u|\mu,\lambda^{-1}\sigma^2)dH(\lambda|\nu)du$$
$$= \int_{-\infty}^{\infty}\Phi(\frac{\lambda^{1/2}[y-\mu]}{\sigma})dH(\lambda|\nu),$$

where $\Phi(.)$ is the cdf of the standard normal distribution.

- Special cases of the SMN family include normal (by fixing $\lambda = 1$), Student-t (by setting $\lambda \sim \mathcal{G}(\frac{\nu}{2},\frac{\nu}{2})$) and Slash (by setting $\lambda \sim \mathcal{B}e(\nu,1)$), where $\mathcal{G}(.,.)$ and $\mathcal{B}e(.,.)$ denote the gamma and distribution respectively.
- Notice that unlike the GEV and SPLOGIT distributions, SMN distribution is always symmetric, which means the flexibility of the link only exists in tail behavior but not in skewness adjustment.

Scale Mixture of Normal Link

- Based on the pdf, the cdf of the SMN distribution is given by

$$
\begin{aligned}
F_{SMN}(y|\mu,\sigma^2,\nu) &= \int_{-\infty}^{y}\int_{-\infty}^{\infty}\phi(u|\mu,\lambda^{-1}\sigma^2)dH(\lambda|\nu)du \\
&= \int_{-\infty}^{\infty}\Phi(\frac{\lambda^{1/2}[y-\mu]}{\sigma})dH(\lambda|\nu),
\end{aligned}
$$

  where $\Phi(.)$ is the cdf of the standard normal distribution.

- Special cases of the SMN family include normal (by fixing $\lambda = 1$), Student-t (by setting $\lambda \sim \mathcal{G}(\frac{\nu}{2},\frac{\nu}{2})$) and Slash (by setting $\lambda \sim \mathcal{B}e(\nu,1)$), where $\mathcal{G}(.,.)$ and $\mathcal{B}e(.,.)$ denote the gamma and distribution respectively.

- Notice that unlike the GEV and SPLOGIT distributions, SMN distribution is always symmetric, which means the flexibility of the link only exists in tail behavior but not in skewness adjustment.

Scale Mixture of Normal Link

- Based on the pdf, the cdf of the SMN distribution is given by

$$F_{SMN}(y|\mu, \sigma^2, \nu) = \int_{-\infty}^{y} \int_{-\infty}^{\infty} \phi(u|\mu, \lambda^{-1}\sigma^2) dH(\lambda|\nu) du$$
$$= \int_{-\infty}^{\infty} \Phi(\frac{\lambda^{1/2}[y - \mu]}{\sigma}) dH(\lambda|\nu),$$

  where $\Phi(.)$ is the cdf of the standard normal distribution.

- Special cases of the SMN family include normal (by fixing $\lambda = 1$), Student-t (by setting $\lambda \sim \mathcal{G}(\frac{\nu}{2}, \frac{\nu}{2})$) and Slash (by setting $\lambda \sim \mathcal{B}e(\nu, 1)$), where $\mathcal{G}(., .)$ and $\mathcal{B}e(., .)$ denote the gamma and distribution respectively.
- Notice that unlike the GEV and SPLOGIT distributions, SMN distribution is always symmetric, which means the flexibility of the link only exists in tail behavior but not in skewness adjustment.

SMN Plots

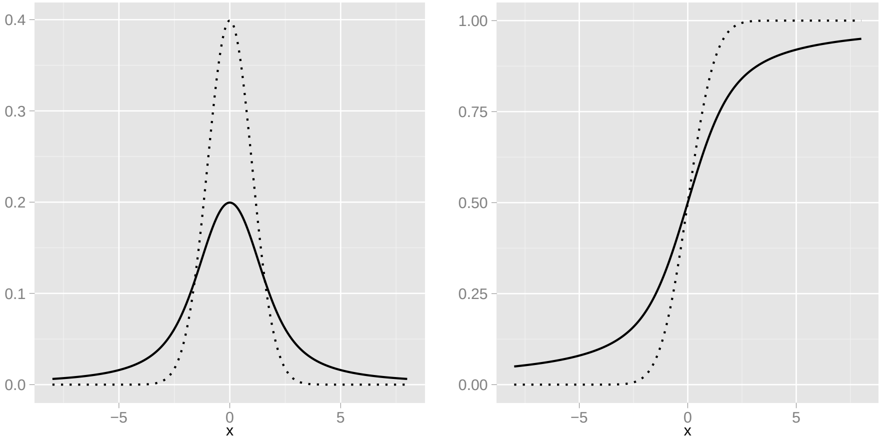


Figure : Comparisons of pdf and cdf of standard SLASH distribution ($\nu = 1$) and standard normal distribution. Solid line (SLASH) and dashed line (standard normal).

BSSM models with flexible links

- Let $\mathbf{Y}_{1:T} = (Y_1, \ldots, Y_T)'$, where $Y_t, t = 1, ..., T$, denote T independent binary random variables. Suppose $\mathbf{x}_t$ is a $k \times 1$ vector of covariates.
- We set the BSSM model as following:

$$\begin{aligned} Y_t &\sim \mathcal{B}er(\pi_t) \qquad\qquad t = 1, \ldots, T \\ \pi_t &= P(Y_t = 1 \mid \theta_t, \mathbf{x}_t, \boldsymbol{\beta}) = F(\mathbf{x}_t'\beta + \theta_t) \\ \theta_t &= \delta\theta_{t-1} + \tau\eta_t. \end{aligned}$$

- We choose F as coming from GEV, SPLOGIT and SMN distributions to construct BSSM with flexible link functions.

BSSM models with flexible links

- Let $\mathbf{Y}_{1:T} = (Y_1, \ldots, Y_T)'$, where $Y_t, t = 1, ..., T$, denote T independent binary random variables. Suppose $\mathbf{x}_t$ is a $k \times 1$ vector of covariates.
- We set the BSSM model as following:

$$\begin{aligned} Y_t &\sim \mathcal{B}er(\pi_t) \qquad\qquad t = 1, \ldots, T \\ \pi_t &= P(Y_t = 1 \mid \theta_t, \mathbf{x}_t, \beta) = F(\mathbf{x}_t'\beta + \theta_t) \\ \theta_t &= \delta\theta_{t-1} + \tau\eta_t. \end{aligned}$$

- We choose F as coming from GEV, SPLOGIT and SMN distributions to construct BSSM with flexible link functions.

BSSM models with flexible links

- Let $\mathbf{Y}_{1:T} = (Y_1, \ldots, Y_T)'$, where $Y_t, t = 1, ..., T$, denote T independent binary random variables. Suppose $\mathbf{x}_t$ is a $k \times 1$ vector of covariates.
- We set the BSSM model as following:

$$\begin{aligned} Y_t &\sim \mathcal{B}er(\pi_t) \qquad\qquad t = 1, \ldots, T \\ \pi_t &= P(Y_t = 1 \mid \theta_t, \mathbf{x}_t, \boldsymbol{\beta}) = F(\mathbf{x}_t' \beta + \theta_t) \\ \theta_t &= \delta\theta_{t-1} + \tau\eta_t. \end{aligned}$$

- We choose F as coming from GEV, SPLOGIT and SMN distributions to construct BSSM with flexible link functions.

BSSM models with flexible links

- Let $\mathbf{Y}_{1:T} = (Y_1, \ldots, Y_T)'$, where $Y_t, t = 1, ..., T$, denote T independent binary random variables. Suppose $\mathbf{x}_t$ is a $k \times 1$ vector of covariates.
- We set the BSSM model as following:

$$
\begin{aligned}
Y_t &\sim \mathcal{B}er(\pi_t) \qquad\qquad t = 1, \ldots, T \\
\pi_t &= P(Y_t = 1 \mid \theta_t, \mathbf{x}_t, \boldsymbol{\beta}) = F(\mathbf{x}_t' \beta + \theta_t) \\
\theta_t &= \delta\theta_{t-1} + \tau\eta_t.
\end{aligned}
$$

- We choose F as coming from GEV, SPLOGIT and SMN distributions to construct BSSM with flexible link functions.

BSSM models with GEV link

- In the GEV case, we let $F(x) = 1 - G(-x)$, where $G(x)$ represents the cdf at x for the GEV distribution with $\mu = 0$ and $\sigma = 1$ and unknown shape parameter ξ.
- Notice that the GEV link specified here is the mirror reflection of GEV distribution discribed before, thus is positively skewed for $X < -0.307$ and negatively skewed for $X > -0.307$.
- When $\xi = 0$, the GEV model reduces to CLOGLOG model.

BSSM models with GEV link

- In the GEV case, we let $F(x) = 1 - G(-x)$, where $G(x)$ represents the cdf at x for the GEV distribution with $\mu = 0$ and $\sigma = 1$ and unknown shape parameter ξ.
- Notice that the GEV link specified here is the mirror reflection of GEV distribution discribed before, thus is positively skewed for $X < -0.307$ and negatively skewed for $X > -0.307$.
- When $\xi = 0$, the GEV model reduces to CLOGLOG model.

BSSM models with GEV link

- In the GEV case, we let $F(x) = 1 - G(-x)$, where $G(x)$ represents the cdf at x for the GEV distribution with $\mu = 0$ and $\sigma = 1$ and unknown shape parameter ξ.

- Notice that the GEV link specified here is the mirror reflection of GEV distribution discribed before, thus is positively skewed for $X < -0.307$ and negatively skewed for $X > -0.307$.

- When $\xi = 0$, the GEV model reduces to CLOGLOG model.

BSSM models with GEV link

- In the GEV case, we let $F(x) = 1 - G(-x)$, where $G(x)$ represents the cdf at x for the GEV distribution with $\mu = 0$ and $\sigma = 1$ and unknown shape parameter ξ.

- Notice that the GEV link specified here is the mirror reflection of GEV distribution discribed before, thus is positively skewed for $X < -0.307$ and negatively skewed for $X > -0.307$.

- When $\xi = 0$, the GEV model reduces to CLOGLOG model.

BSSM models with SPLOGIT link

- In the case of SPLOGIT link, we set $F(x) = F(x, r)$, the cdf of SPLOGIT distribution.
- The distribution associated with the link function is negatively skewed when $r < 1$ and positively skewed when $r > 0$, and reduce to LOGIT model when $r = 0$.
- Under Arnold and Groeneveld (1995)'s measure, as r tends to 0 and $+\infty$, the range of skewness provided by SPLOGIT family is unlimited, reaching -1 and 1 respectively.

BSSM models with SPLOGIT link

- In the case of SPLOGIT link, we set $F(x) = F(x, r)$, the cdf of SPLOGIT distribution.

- The distribution associated with the link function is negatively skewed when $r < 1$ and positively skewed when $r > 0$, and reduce to LOGIT model when $r = 0$.

- Under Arnold and Groeneveld (1995)'s measure, as r tends to 0 and $+\infty$, the range of skewness provided by SPLOGIT family is unlimited, reaching -1 and 1 respectively.

BSSM models with SPLOGIT link

- In the case of SPLOGIT link, we set $F(x) = F(x, r)$, the cdf of SPLOGIT distribution.
- The distribution associated with the link function is negatively skewed when $r < 1$ and positively skewed when $r > 0$, and reduce to LOGIT model when $r = 0$.
- Under Arnold and Groeneveld (1995)'s measure, as r tends to 0 and $+\infty$, the range of skewness provided by SPLOGIT family is unlimited, reaching -1 and 1 respectively.

BSSM models with SPLOGIT link

- In the case of SPLOGIT link, we set $F(x) = F(x, r)$, the cdf of SPLOGIT distribution.
- The distribution associated with the link function is negatively skewed when $r < 1$ and positively skewed when $r > 0$, and reduce to LOGIT model when $r = 0$.
- Under Arnold and Groeneveld (1995)'s measure, as r tends to 0 and $+\infty$, the range of skewness provided by SPLOGIT family is unlimited, reaching -1 and 1 respectively.

BSSM models with SMN link

- In SMN case, we set $F(x) = F_{SMN}(x)$ with mixing random variable λ.
- In this study, we assume $\lambda \sim \mathcal{B}e(\nu, 1)$ and thus only consider SLASH model as a representative of the SMN class.
- The SLASH model is always symmetric, thus only allows flexibility in tail behavior but not skewness.

BSSM models with SMN link

- In SMN case, we set $F(x) = F_{SMN}(x)$ with mixing random variable λ.

- In this study, we assume $\lambda \sim \mathcal{B}e(\nu, 1)$ and thus only consider SLASH model as a representative of the SMN class.

- The SLASH model is always symmetric, thus only allows flexibility in tail behavior but not skewness.

BSSM models with SMN link

- In SMN case, we set $F(x) = F_{SMN}(x)$ with mixing random variable λ.

- In this study, we assume $\lambda \sim \mathcal{B}e(\nu, 1)$ and thus only consider SLASH model as a representative of the SMN class.

- The SLASH model is always symmetric, thus only allows flexibility in tail behavior but not skewness.

BSSM models with SMN link

- In SMN case, we set $F(x) = F_{SMN}(x)$ with mixing random variable λ.
- In this study, we assume $\lambda \sim \mathcal{B}e(\nu, 1)$ and thus only consider SLASH model as a representative of the SMN class.
- The SLASH model is always symmetric, thus only allows flexibility in tail behavior but not skewness.

Inference Procedures

- Here we develop a Markov Chain Monte Carlo (MCMC) procedure to make inference about the BSSM model with flexible link functions under the Bayesian paradigm.
- The model depends on a parameter vector Ψ, where $\Psi = (\beta', \delta, \tau^2, \xi)'$ in GEV case, $\Psi = (\beta', \delta, \tau^2, r)'$ in SPLOGIT case and $\Psi = (\beta', \delta, \tau^2, \nu)'$ in SLASH model.
- Let $\boldsymbol{\theta}_{0:T} = (\theta_0, \theta_1, \ldots, \theta_T)'$ be the latent states. The Bayesian approach for estimating model parameters treats $\theta_{0:T}$ as latent parameters themselves and updates them in each step of MCMC.
- The joint posterior density of parameters and latent variables can be written as

$$p(\boldsymbol{\theta}_{0:T}, \Psi \mid \mathbf{y}_{1:T}) \propto p(\mathbf{Y}_{1:T} \mid \boldsymbol{\theta}_{0:T}, \Psi, \mathbf{y}_{1:T}) p(\boldsymbol{\theta}_{0:T} \mid \Psi) p(\Psi).$$

Inference Procedures

- Here we develop a Markov Chain Monte Carlo (MCMC) procedure to make inference about the BSSM model with flexible link functions under the Bayesian paradigm.
- The model depends on a parameter vector Ψ, where $\Psi = (\beta', \delta, \tau^2, \xi)'$ in GEV case, $\Psi = (\beta', \delta, \tau^2, r)'$ in SPLOGIT case and $\Psi = (\beta', \delta, \tau^2, \nu)'$ in SLASH model.
- Let $\boldsymbol{\theta}_{0:T} = (\theta_0, \theta_1, \dots, \theta_T)'$ be the latent states. The Bayesian approach for estimating model parameters treats $\theta_{0:T}$ as latent parameters themselves and updates them in each step of MCMC.
- The joint posterior density of parameters and latent variables can be written as

$$p(\boldsymbol{\theta}_{0:T}, \Psi \mid \mathbf{y}_{1:T}) \quad \propto \quad p(\mathbf{Y}_{1:T} \mid \boldsymbol{\theta}_{0:T}, \Psi, \mathbf{y}_{1:T}) p(\boldsymbol{\theta}_{0:T} \mid \Psi) p(\Psi).$$

Inference Procedures

- Here we develop a Markov Chain Monte Carlo (MCMC) procedure to make inference about the BSSM model with flexible link functions under the Bayesian paradigm.
- The model depends on a parameter vector Ψ, where $\Psi = (\boldsymbol{\beta}', \delta, \tau^2, \xi)'$ in GEV case, $\Psi = (\boldsymbol{\beta}', \delta, \tau^2, r)'$ in SPLOGIT case and $\Psi = (\boldsymbol{\beta}', \delta, \tau^2, \nu)'$ in SLASH model.
- Let $\boldsymbol{\theta}_{0:T} = (\theta_0, \theta_1, \ldots, \theta_T)'$ be the latent states. The Bayesian approach for estimating model parameters treats $\theta_{0:T}$ as latent parameters themselves and updates them in each step of MCMC.
- The joint posterior density of parameters and latent variables can be written as

$$p(\boldsymbol{\theta}_{0:T}, \Psi \mid \mathbf{y}_{1:T}) \quad \propto \quad p(\mathbf{Y}_{1:T} \mid \boldsymbol{\theta}_{0:T}, \Psi, \mathbf{y}_{1:T})p(\boldsymbol{\theta}_{0:T} \mid \Psi)p(\Psi).$$

Inference Procedures

- Here we develop a Markov Chain Monte Carlo (MCMC) procedure to make inference about the BSSM model with flexible link functions under the Bayesian paradigm.

- The model depends on a parameter vector Ψ, where $\Psi = (\boldsymbol{\beta}', \delta, \tau^2, \xi)'$ in GEV case, $\Psi = (\boldsymbol{\beta}', \delta, \tau^2, r)'$ in SPLOGIT case and $\Psi = (\boldsymbol{\beta}', \delta, \tau^2, \nu)'$ in SLASH model.

- Let $\boldsymbol{\theta}_{0:T} = (\theta_0, \theta_1, \ldots, \theta_T)'$ be the latent states. The Bayesian approach for estimating model parameters treats $\theta_{0:T}$ as latent parameters themselves and updates them in each step of MCMC.

- The joint posterior density of parameters and latent variables can be written as

$$p(\boldsymbol{\theta}_{0:T}, \Psi \mid \mathbf{y}_{1:T}) \quad \propto \quad p(\mathbf{Y}_{1:T} \mid \boldsymbol{\theta}_{0:T}, \Psi, \mathbf{y}_{1:T}) p(\boldsymbol{\theta}_{0:T} \mid \Psi) p(\Psi).$$

Inference Procedures

- Here we develop a Markov Chain Monte Carlo (MCMC) procedure to make inference about the BSSM model with flexible link functions under the Bayesian paradigm.
- The model depends on a parameter vector Ψ, where $\Psi = (\boldsymbol{\beta}', \delta, \tau^2, \xi)'$ in GEV case, $\Psi = (\boldsymbol{\beta}', \delta, \tau^2, r)'$ in SPLOGIT case and $\Psi = (\boldsymbol{\beta}', \delta, \tau^2, \nu)'$ in SLASH model.
- Let $\boldsymbol{\theta}_{0:T} = (\theta_0, \theta_1, \ldots, \theta_T)'$ be the latent states. The Bayesian approach for estimating model parameters treats $\theta_{0:T}$ as latent parameters themselves and updates them in each step of MCMC.
- The joint posterior density of parameters and latent variables can be written as

$$p(\boldsymbol{\theta}_{0:T}, \Psi \mid \mathbf{y}_{1:T}) \quad \propto \quad p(\mathbf{Y}_{1:T} \mid \boldsymbol{\theta}_{0:T}, \Psi, \mathbf{y}_{1:T}) p(\boldsymbol{\theta}_{0:T} \mid \Psi) p(\Psi).$$

Inference Procedures

- In the previous equation,

$$p(\theta_{0:T} \mid \Psi) \quad = \quad \phi(\theta_0 \mid 0, \frac{\tau^2}{1-\delta^2}) \prod_{t=1}^{T} \phi(\theta_t \mid \delta\theta_{t-1}, \tau^2),$$

where $\phi(x \mid \mu, \sigma^2)$ denotes the normal density with mean μ and variance σ^2 evaluated at x and $p(\Psi)$ indicates the prior distribution.

- In GEV, SPLOGIT and SLASH models the three prior distributions of Ψ can be written as

$$\begin{aligned} p_{GEV}(\Psi) &= p(\beta)p(\delta)p(\tau^2)p(\xi), \\ p_{SPLOGIT}(\Psi) &= p(\beta)p(\delta)p(\tau^2)p(r), \\ p_{SLASH}(\Psi) &= p(\beta)p(\delta)p(\tau^2)\prod_{t=1}^{T} p(\lambda_t|\nu)p(\nu). \end{aligned}$$

Inference Procedures

- In the previous equation,

$$p(\boldsymbol{\theta}_{0:T} \mid \Psi) = \phi(\theta_0 \mid 0, \frac{\tau^2}{1-\delta^2}) \prod_{t=1}^{T} \phi(\theta_t \mid \delta\theta_{t-1}, \tau^2),$$

where $\phi(x \mid \mu, \sigma^2)$ denotes the normal density with mean μ and variance σ^2 evaluated at x and $p(\Psi)$ indicates the prior distribution.

- In GEV, SPLOGIT and SLASH models the three prior distributions of Ψ can be written as

$$\begin{aligned} p_{GEV}(\Psi) &= p(\boldsymbol{\beta})p(\delta)p(\tau^2)p(\xi), \\ p_{SPLOGIT}(\Psi) &= p(\boldsymbol{\beta})p(\delta)p(\tau^2)p(r), \\ p_{SLASH}(\Psi) &= p(\boldsymbol{\beta})p(\delta)p(\tau^2)\prod_{t=1}^{T} p(\lambda_t|\nu)p(\nu). \end{aligned}$$

Inference Procedures

- In the previous equation,

$$p(\boldsymbol{\theta}_{0:T} \mid \Psi) \quad = \quad \phi(\theta_0 \mid 0, \frac{\tau^2}{1-\delta^2}) \prod_{t=1}^{T} \phi(\theta_t \mid \delta\theta_{t-1}, \tau^2),$$

where $\phi(x \mid \mu, \sigma^2)$ denotes the normal density with mean μ and variance σ^2 evaluated at x and $p(\Psi)$ indicates the prior distribution.

- In GEV, SPLOGIT and SLASH models the three prior distributions of Ψ can be written as

$$\begin{aligned} p_{GEV}(\Psi) &= p(\boldsymbol{\beta})p(\delta)p(\tau^2)p(\xi), \\ p_{SPLOGIT}(\Psi) &= p(\boldsymbol{\beta})p(\delta)p(\tau^2)p(r), \\ p_{SLASH}(\Psi) &= p(\boldsymbol{\beta})p(\delta)p(\tau^2)\prod_{t=1}^{T} p(\lambda_t|\nu)p(\nu). \end{aligned}$$

Inference Procedures

- The prior distributions for individual parameters are set as: $\beta \sim \mathcal{N}_k(\mathbf{0}, 500^2 I)$, $\delta \sim \mathcal{N}_{(-1,1)}(0.95, 100)$, $\tau^2 \sim \mathcal{IG}(0.25, 0.01)$, $\xi \sim \mathcal{U}(-0.6, 0.6)$, $r \sim \mathcal{G}(1, 1)$, $\lambda_t \sim \mathcal{B}e(\nu, 1), t = 1, ..., T$ and $\nu \sim \mathcal{G}(5, 0.8)$.
- We adopt the JAGS software to implement these modes. The single-move sampler in JAGS produces higher correlated posterior samples. Such dependency can be compensated by running a longer Markov chain.
- There are growing number of advanced sampling schemes available these days. However, the gain in efficiency in using complex sampling schemes to some extent is outweighed by the ease of implementation in JAGS, therefore being unattractive to general practitioners.

Inference Procedures

- The prior distributions for individual parameters are set as: $\boldsymbol{\beta} \sim \mathcal{N}_k(\mathbf{0}, 500^2 I)$, $\delta \sim \mathcal{N}_{(-1,1)}(0.95, 100)$, $\tau^2 \sim \mathcal{IG}(0.25, 0.01)$, $\xi \sim \mathcal{U}(-0.6, 0.6)$, $r \sim \mathcal{G}(1, 1)$, $\lambda_t \sim \mathcal{B}e(\nu, 1), t = 1, ..., T$ and $\nu \sim \mathcal{G}(5, 0.8)$.

- We adopt the JAGS software to implement these modes. The single-move sampler in JAGS produces higher correlated posterior samples. Such dependency can be compensated by running a longer Markov chain.

- There are growing number of advanced sampling schemes available these days. However, the gain in efficiency in using complex sampling schemes to some extent is outweighed by the ease of implementation in JAGS, therefore being unattractive to general practitioners.

Inference Procedures

- The prior distributions for individual parameters are set as: $\boldsymbol{\beta} \sim \mathcal{N}_k(\mathbf{0}, 500^2 I)$, $\delta \sim \mathcal{N}_{(-1,1)}(0.95, 100)$, $\tau^2 \sim \mathcal{IG}(0.25, 0.01)$, $\xi \sim \mathcal{U}(-0.6, 0.6)$, $r \sim \mathcal{G}(1, 1)$, $\lambda_t \sim \mathcal{B}e(\nu, 1), t = 1, ..., T$ and $\nu \sim \mathcal{G}(5, 0.8)$.
- We adopt the JAGS software to implement these modes. The single-move sampler in JAGS produces higher correlated posterior samples. Such dependency can be compensated by running a longer Markov chain.
- There are growing number of advanced sampling schemes available these days. However, the gain in efficiency in using complex sampling schemes to some extent is outweighed by the ease of implementation in JAGS, therefore being unattractive to general practitioners.

Inference Procedures

- The prior distributions for individual parameters are set as: $\boldsymbol{\beta} \sim \mathcal{N}_k(\mathbf{0}, 500^2 I)$, $\delta \sim \mathcal{N}_{(-1,1)}(0.95, 100)$, $\tau^2 \sim \mathcal{IG}(0.25, 0.01)$, $\xi \sim \mathcal{U}(-0.6, 0.6)$, $r \sim \mathcal{G}(1, 1)$, $\lambda_t \sim \mathcal{B}e(\nu, 1), t = 1, ..., T$ and $\nu \sim \mathcal{G}(5, 0.8)$.

- We adopt the JAGS software to implement these modes. The single-move sampler in JAGS produces higher correlated posterior samples. Such dependency can be compensated by running a longer Markov chain.

- There are growing number of advanced sampling schemes available these days. However, the gain in efficiency in using complex sampling schemes to some extent is outweighed by the ease of implementation in JAGS, therefore being unattractive to general practitioners.

Simulation Study

- We conduct a simulation study to investigate the robustness of the BSSM-GEV, BSSM-SPLOGIT and BSSM-SLASH models against link misspecification when the data are generated from different standard models.
- We generate our data from BSSM setup with F set to be the cdf corresponding to LOGIT, CLOGLOG and LOGLOG links.
- Under Arnold and Groeneveld (1995)s measure, the skewness associated with the three links are 0, -0.264 and 0.264.
- For each of the true model we independently generate 200 datasets of sample sizes $T = 800$. For each dateset, we generate one covariates $\mathbf{x}_t, t = 1, ..., T$ from independent standard normal distributions.

Simulation Study

- We conduct a simulation study to investigate the robustness of the BSSM-GEV, BSSM-SPLOGIT and BSSM-SLASH models against link misspecification when the data are generated from different standard models.
- We generate our data from BSSM setup with F set to be the cdf corresponding to LOGIT, CLOGLOG and LOGLOG links.
- Under Arnold and Groeneveld (1995)s measure, the skewness associated with the three links are 0, -0.264 and 0.264.
- For each of the true model we independently generate 200 datasets of sample sizes $T = 800$. For each dateset, we generate one covariates $\mathbf{x}_t, t = 1, ..., T$ from independent standard normal distributions.

Simulation Study

- We conduct a simulation study to investigate the robustness of the BSSM-GEV, BSSM-SPLOGIT and BSSM-SLASH models against link misspecification when the data are generated from different standard models.
- We generate our data from BSSM setup with F set to be the cdf corresponding to LOGIT, CLOGLOG and LOGLOG links.
- Under Arnold and Groeneveld (1995)s measure, the skewness associated with the three links are 0, -0.264 and 0.264.
- For each of the true model we independently generate 200 datasets of sample sizes $T = 800$. For each dateset, we generate one covariates $\mathbf{x}_t$, $t = 1, ..., T$ from independent standard normal distributions.

Simulation Study

- We conduct a simulation study to investigate the robustness of the BSSM-GEV, BSSM-SPLOGIT and BSSM-SLASH models against link misspecification when the data are generated from different standard models.
- We generate our data from BSSM setup with F set to be the cdf corresponding to LOGIT, CLOGLOG and LOGLOG links.
- Under Arnold and Groeneveld (1995)s measure, the skewness associated with the three links are 0, -0.264 and 0.264.
- For each of the true model we independently generate 200 datasets of sample sizes $T = 800$. For each dateset, we generate one covariates $\mathbf{x}_t, t = 1, ..., T$ from independent standard normal distributions.

Simulation Study

- We conduct a simulation study to investigate the robustness of the BSSM-GEV, BSSM-SPLOGIT and BSSM-SLASH models against link misspecification when the data are generated from different standard models.
- We generate our data from BSSM setup with F set to be the cdf corresponding to LOGIT, CLOGLOG and LOGLOG links.
- Under Arnold and Groeneveld (1995)s measure, the skewness associated with the three links are 0, -0.264 and 0.264.
- For each of the true model we independently generate 200 datasets of sample sizes $T = 800$. For each dateset, we generate one covariates $\mathbf{x}_t, t = 1, ..., T$ from independent standard normal distributions.

Simulation Study

- The true value regression coefficients are set to be $\beta = (\beta_0, \beta_1)' = (0, 1)'$. We set other parameters $\delta = 0.95$ and $\tau = 0.2$.
- We fit the BSSM-LOGIT, BSSM-CLOGLOG, BSSM-LOGLOG, BSSM-GEV, BSSM-SPLOGIT and BSSM-SLASH models to each set of the generated data and compare the outcomes.
- For 200 replicates, the average posterior mean and posterior standard deviation of the parameters under different combination of true and fitted models are summarized.

Simulation Study

- The true value regression coefficients are set to be $\boldsymbol{\beta} = (\beta_0, \beta_1)' = (0, 1)'$. We set other parameters $\delta = 0.95$ and $\tau = 0.2$.
- We fit the BSSM-LOGIT, BSSM-CLOGLOG, BSSM-LOGLOG, BSSM-GEV, BSSM-SPLOGIT and BSSM-SLASH models to each set of the generated data and compare the outcomes.
- For 200 replicates, the average posterior mean and posterior standard deviation of the parameters under different combination of true and fitted models are summarized.

Simulation Study

- The true value regression coefficients are set to be $\boldsymbol{\beta} = (\beta_0, \beta_1)' = (0, 1)'$. We set other parameters $\delta = 0.95$ and $\tau = 0.2$.
- We fit the BSSM-LOGIT, BSSM-CLOGLOG, BSSM-LOGLOG, BSSM-GEV, BSSM-SPLOGIT and BSSM-SLASH models to each set of the generated data and compare the outcomes.
- For 200 replicates, the average posterior mean and posterior standard deviation of the parameters under different combination of true and fitted models are summarized.

Simulation Study

- The true value regression coefficients are set to be $\boldsymbol{\beta} = (\beta_0, \beta_1)' = (0, 1)'$. We set other parameters $\delta = 0.95$ and $\tau = 0.2$.
- We fit the BSSM-LOGIT, BSSM-CLOGLOG, BSSM-LOGLOG, BSSM-GEV, BSSM-SPLOGIT and BSSM-SLASH models to each set of the generated data and compare the outcomes.
- For 200 replicates, the average posterior mean and posterior standard deviation of the parameters under different combination of true and fitted models are summarized.

Simulation Study

True model: BSSMM-LOGIT												
	Fitted model											
	LOGIT		CLOGLOG		LOGLOG		GEV		SPLOGIT		SLASH	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
β_0	**0.036**	**0.174**	-0.475	0.122	0.432	0.120	-0.415	0.144	-0.040	0.406	0.024	0.131
β_1	**1.036**	**0.109**	0.703	0.076	0.703	0.080	0.672	0.091	0.822	0.155	0.733	0.115
δ	**0.902**	**0.078**	0.878	0.088	0.874	0.099	0.878	0.089	0.889	0.092	0.891	0.083
τ^2	**0.132**	**0.159**	0.090	0.121	0.107	0.157	0.079	0.098	0.074	0.075	0.101	0.083
$\xi/r/\nu$							-0.141	0.204	1.045	0.511	4.313	2.044

- The fit under LOGIT model recovers true model parameters.
- $\xi = -0.141$ in GEV and $r = 1.045$ in SPLOGIT captures the fact that the true model is symmetric.
- The fact that SPLOGIT fit approximates true model parameters nicely is reasonable since LOGIT model is a special case of SPLOGIT model.

Simulation Study

True model: BSSMM-LOGIT												
	Fitted model											
	LOGIT		CLOGLOG		LOGLOG		GEV		SPLOGIT		SLASH	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
β_0	**0.036**	**0.174**	-0.475	0.122	0.432	0.120	-0.415	0.144	-0.040	0.406	0.024	0.131
β_1	**1.036**	**0.109**	0.703	0.076	0.703	0.080	0.672	0.091	0.822	0.155	0.733	0.115
δ	**0.902**	**0.078**	0.878	0.088	0.874	0.099	0.878	0.089	0.889	0.092	0.891	0.083
τ^2	**0.132**	**0.159**	0.090	0.121	0.107	0.157	0.079	0.098	0.074	0.075	0.101	0.083
$\xi/r/\nu$							-0.141	0.204	1.045	0.511	4.313	2.044

- The fit under LOGIT model recovers true model parameters.
- $\xi = -0.141$ in GEV and $r = 1.045$ in SPLOGIT captures the fact that the true model is symmetric.
- The fact that SPLOGIT fit approximates true model parameters nicely is reasonable since LOGIT model is a special case of SPLOGIT model.

Simulation Study

True model: BSSMM-LOGIT												
	Fitted model											
	LOGIT		CLOGLOG		LOGLOG		GEV		SPLOGIT		SLASH	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
β_0	**0.036**	**0.174**	-0.475	0.122	0.432	0.120	-0.415	0.144	-0.040	0.406	0.024	0.131
β_1	**1.036**	**0.109**	0.703	0.076	0.703	0.080	0.672	0.091	0.822	0.155	0.733	0.115
δ	**0.902**	**0.078**	0.878	0.088	0.874	0.099	0.878	0.089	0.889	0.092	0.891	0.083
τ^2	**0.132**	**0.159**	0.090	0.121	0.107	0.157	0.079	0.098	0.074	0.075	0.101	0.083
$\xi/r/\nu$							-0.141	0.204	1.045	0.511	4.313	2.044

- The fit under LOGIT model recovers true model parameters.
- $\xi = -0.141$ in GEV and $r = 1.045$ in SPLOGIT captures the fact that the true model is symmetric.
- The fact that SPLOGIT fit approximates true model parameters nicely is reasonable since LOGIT model is a special case of SPLOGIT model.

Simulation Study

True model: BSSMM-LOGIT												
	Fitted model											
	LOGIT		CLOGLOG		LOGLOG		GEV		SPLOGIT		SLASH	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
β_0	**0.036**	**0.174**	-0.475	0.122	0.432	0.120	-0.415	0.144	-0.040	0.406	0.024	0.131
β_1	**1.036**	**0.109**	0.703	0.076	0.703	0.080	0.672	0.091	0.822	0.155	0.733	0.115
δ	**0.902**	**0.078**	0.878	0.088	0.874	0.099	0.878	0.089	0.889	0.092	0.891	0.083
τ^2	**0.132**	**0.159**	0.090	0.121	0.107	0.157	0.079	0.098	0.074	0.075	0.101	0.083
$\xi/r/\nu$							-0.141	0.204	1.045	0.511	4.313	2.044

- The fit under LOGIT model recovers true model parameters.
- $\xi = -0.141$ in GEV and $r = 1.045$ in SPLOGIT captures the fact that the true model is symmetric.
- The fact that SPLOGIT fit approximates true model parameters nicely is reasonable since LOGIT model is a special case of SPLOGIT model.

Simulation Study

	True model: BSSMM-CLOGLOG											
	Fitted model											
	LOGIT		CLOGLOG		LOGLOG		GEV		SPLOGIT		SLASH	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
β_0	0.819	0.262	**0.046**	**0.168**	1.043	0.191	-0.014	0.180	0.023	0.402	0.516	0.183
β_1	1.626	0.149	**1.041**	**0.099**	1.157	0.136	1.068	0.124	1.108	0.243	1.109	0.133
δ	0.929	0.036	**0.929**	**0.035**	0.894	0.060	0.926	0.036	0.926	0.036	0.927	0.039
τ^2	0.194	0.140	**0.089**	**0.074**	0.219	0.248	0.072	0.074	0.095	0.077	0.093	0.078
$\xi/r/\nu$							0.073	0.193	0.560	0.291	4.815	2.108

- The fit under CLOGLOG model recovers true model parameters.
- $\xi = 0.073$ in GEV and $r = 0.560$ in SPLOGIT captures the fact that the true model is left skewed.
- The fact that GEV fit approximates true model parameters nicely is reasonable since CLOGLOG model is a special case of GEV model.

Simulation Study

True model: BSSMM-CLOGLOG												
	Fitted model											
	LOGIT		CLOGLOG		LOGLOG		GEV		SPLOGIT		SLASH	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
β_0	0.819	0.262	**0.046**	**0.168**	1.043	0.191	-0.014	0.180	0.023	0.402	0.516	0.183
β_1	1.626	0.149	**1.041**	**0.099**	1.157	0.136	1.068	0.124	1.108	0.243	1.109	0.133
δ	0.929	0.036	**0.929**	**0.035**	0.894	0.060	0.926	0.036	0.926	0.036	0.927	0.039
τ^2	0.194	0.140	**0.089**	**0.074**	0.219	0.248	0.072	0.074	0.095	0.077	0.093	0.078
$\xi/r/\nu$							0.073	0.193	0.560	0.291	4.815	2.108

- The fit under CLOGLOG model recovers true model parameters.
- $\xi = 0.073$ in GEV and $r = 0.560$ in SPLOGIT captures the fact that the true model is left skewed.
- The fact that GEV fit approximates true model parameters nicely is reasonable since CLOGLOG model is a special case of GEV model.

Simulation Study

True model: BSSMM-CLOGLOG												
	Fitted model											
	LOGIT		CLOGLOG		LOGLOG		GEV		SPLOGIT		SLASH	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
β_0	0.819	0.262	**0.046**	**0.168**	1.043	0.191	-0.014	0.180	0.023	0.402	0.516	0.183
β_1	1.626	0.149	**1.041**	**0.099**	1.157	0.136	1.068	0.124	1.108	0.243	1.109	0.133
δ	0.929	0.036	**0.929**	**0.035**	0.894	0.060	0.926	0.036	0.926	0.036	0.927	0.039
τ^2	0.194	0.140	**0.089**	**0.074**	0.219	0.248	0.072	0.074	0.095	0.077	0.093	0.078
$\xi/r/\nu$							0.073	0.193	0.560	0.291	4.815	2.108

- The fit under CLOGLOG model recovers true model parameters.
- $\xi = 0.073$ in GEV and $r = 0.560$ in SPLOGIT captures the fact that the true model is left skewed.
- The fact that GEV fit approximates true model parameters nicely is reasonable since CLOGLOG model is a special case of GEV model.

Simulation Study

True model: BSSMM-CLOGLOG												
	Fitted model											
	LOGIT		CLOGLOG		LOGLOG		GEV		SPLOGIT		SLASH	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
β_0	0.819	0.262	**0.046**	**0.168**	1.043	0.191	-0.014	0.180	0.023	0.402	0.516	0.183
β_1	1.626	0.149	**1.041**	**0.099**	1.157	0.136	1.068	0.124	1.108	0.243	1.109	0.133
δ	0.929	0.036	**0.929**	**0.035**	0.894	0.060	0.926	0.036	0.926	0.036	0.927	0.039
τ^2	0.194	0.140	**0.089**	**0.074**	0.219	0.248	0.072	0.074	0.095	0.077	0.093	0.078
$\xi/r/\nu$							0.073	0.193	0.560	0.291	4.815	2.108

- The fit under CLOGLOG model recovers true model parameters.
- $\xi = 0.073$ in GEV and $r = 0.560$ in SPLOGIT captures the fact that the true model is left skewed.
- The fact that GEV fit approximates true model parameters nicely is reasonable since CLOGLOG model is a special case of GEV model.

Simulation Study

True model: BSSMM-LOGLOG												
	Fitted model											
	LOGIT		CLOGLOG		LOGLOG		GEV		SPLOGIT		SLASH	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
β_0	-0.664	0.258	-1.082	0.189	**-0.052**	**0.167**	-0.750	0.178	-0.107	0.414	-0.470	0.184
β_1	1.590	0.145	1.126	0.129	**1.023**	**0.094**	0.872	0.099	1.160	0.230	1.072	0.125
δ	0.931	0.037	0.902	0.058	**0.931**	**0.036**	0.933	0.034	0.931	0.036	0.931	0.037
τ^2	0.171	0.124	0.210	0.247	**0.071**	**0.052**	0.052	0.038	0.097	0.076	0.079	0.056
$\xi/r/\nu$							-0.413	0.107	1.956	0.874	4.680	2.049

- The fit under LOGLOG model recovers true model parameters.
- $\xi = -0.413$ in GEV and $r = 1.956$ in SPLOGIT captures the fact that the true model is right skewed.
- The mixing parameter λ in SLASH model performs similarly in 3 cases in terms of parameter estimate.

Simulation Study

True model: BSSMM-LOGLOG												
	Fitted model											
	LOGIT		CLOGLOG		LOGLOG		GEV		SPLOGIT		SLASH	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
β_0	-0.664	0.258	-1.082	0.189	**-0.052**	**0.167**	-0.750	0.178	-0.107	0.414	-0.470	0.184
β_1	1.590	0.145	1.126	0.129	**1.023**	**0.094**	0.872	0.099	1.160	0.230	1.072	0.125
δ	0.931	0.037	0.902	0.058	**0.931**	**0.036**	0.933	0.034	0.931	0.036	0.931	0.037
τ^2	0.171	0.124	0.210	0.247	**0.071**	**0.052**	0.052	0.038	0.097	0.076	0.079	0.056
$\xi/r/\nu$							-0.413	0.107	1.956	0.874	4.680	2.049

- The fit under LOGLOG model recovers true model parameters.
- $\xi = -0.413$ in GEV and $r = 1.956$ in SPLOGIT captures the fact that the true model is right skewed.
- The mixing parameter λ in SLASH model performs similarly in 3 cases in terms of parameter estimate.

Simulation Study

True model: BSSMM-LOGLOG												
	Fitted model											
	LOGIT		CLOGLOG		LOGLOG		GEV		SPLOGIT		SLASH	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
β_0	-0.664	0.258	-1.082	0.189	**-0.052**	**0.167**	-0.750	0.178	-0.107	0.414	-0.470	0.184
β_1	1.590	0.145	1.126	0.129	**1.023**	**0.094**	0.872	0.099	1.160	0.230	1.072	0.125
δ	0.931	0.037	0.902	0.058	**0.931**	**0.036**	0.933	0.034	0.931	0.036	0.931	0.037
τ^2	0.171	0.124	0.210	0.247	**0.071**	**0.052**	0.052	0.038	0.097	0.076	0.079	0.056
$\xi/r/\nu$							-0.413	0.107	1.956	0.874	4.680	2.049

- The fit under LOGLOG model recovers true model parameters.
- $\xi = -0.413$ in GEV and $r = 1.956$ in SPLOGIT captures the fact that the true model is right skewed.
- The mixing parameter λ in SLASH model performs similarly in 3 cases in terms of parameter estimate.

Simulation Study

True model: BSSMM-LOGLOG												
	Fitted model											
	LOGIT		CLOGLOG		LOGLOG		GEV		SPLOGIT		SLASH	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
β_0	-0.664	0.258	-1.082	0.189	**-0.052**	**0.167**	-0.750	0.178	-0.107	0.414	-0.470	0.184
β_1	1.590	0.145	1.126	0.129	**1.023**	**0.094**	0.872	0.099	1.160	0.230	1.072	0.125
δ	0.931	0.037	0.902	0.058	**0.931**	**0.036**	0.933	0.034	0.931	0.036	0.931	0.037
τ^2	0.171	0.124	0.210	0.247	**0.071**	**0.052**	0.052	0.038	0.097	0.076	0.079	0.056
$\xi/r/\nu$							-0.413	0.107	1.956	0.874	4.680	2.049

- The fit under LOGLOG model recovers true model parameters.
- $\xi = -0.413$ in GEV and $r = 1.956$ in SPLOGIT captures the fact that the true model is right skewed.
- The mixing parameter λ in SLASH model performs similarly in 3 cases in terms of parameter estimate.

Simulation Study

Table : Percentage of best performance and average DIC under different model fitting among 200 replicates.

	True model: BSSMM-LOGIT				
	CLOGLOG	LOGLOG	GEV	SPLOGIT	SLASH
% best	16.0%	20.5%	24.0%	1.5%	38.0%
avg DIC	934.9	935.0	933.4	937.1	933.5
	True model: BSSMM-CLOGLOG				
	LOGIT	LOGLOG	GEV	SPLOGIT	SLASH
% best	0.0%	14.5%	68.5%	3.5%	13.5%
avg DIC	770.8	768.9	755.5	762.7	765.5
	True model: BSSMM-LOGLOG				
	LOGIT	CLOGLOG	GEV	SPLOGIT	SLASH
% best	0.5%	14.5%	70.5%	1.0%	13.5%
agv DIC	778.6	778.1	768.5	774.4	773.8

- GEV model is extremely robust against link misspecifications.
- SPLOGIT model performs better than other standard links in terms average DIC but is usually worse than GEV.
- SLASH model performs best when the true model is LOGIT (symmetric).

Simulation Study

Table : Percentage of best performance and average DIC under different model fitting among 200 replicates.

True model: BSSMM-LOGIT					
	CLOGLOG	LOGLOG	GEV	SPLOGIT	SLASH
% best	16.0%	20.5%	24.0%	1.5%	38.0%
avg DIC	934.9	935.0	933.4	937.1	933.5
True model: BSSMM-CLOGLOG					
	LOGIT	LOGLOG	GEV	SPLOGIT	SLASH
% best	0.0%	14.5%	68.5%	3.5%	13.5%
avg DIC	770.8	768.9	755.5	762.7	765.5
True model: BSSMM-LOGLOG					
	LOGIT	CLOGLOG	GEV	SPLOGIT	SLASH
% best	0.5%	14.5%	70.5%	1.0%	13.5%
agv DIC	778.6	778.1	768.5	774.4	773.8

- GEV model is extremely robust against link misspecifications.
- SPLOGIT model performs better than other standard links in terms average DIC but is usually worse than GEV.
- SLASH model performs best when the true model is LOGIT (symmetric).

Simulation Study

Table : Percentage of best performance and average DIC under different model fitting among 200 replicates.

	True model: BSSMM-LOGIT				
	CLOGLOG	LOGLOG	GEV	SPLOGIT	SLASH
% best	16.0%	20.5%	24.0%	1.5%	38.0%
avg DIC	934.9	935.0	933.4	937.1	933.5
	True model: BSSMM-CLOGLOG				
	LOGIT	LOGLOG	GEV	SPLOGIT	SLASH
% best	0.0%	14.5%	68.5%	3.5%	13.5%
avg DIC	770.8	768.9	755.5	762.7	765.5
	True model: BSSMM-LOGLOG				
	LOGIT	CLOGLOG	GEV	SPLOGIT	SLASH
% best	0.5%	14.5%	70.5%	1.0%	13.5%
agv DIC	778.6	778.1	768.5	774.4	773.8

- GEV model is extremely robust against link misspecifications.
- SPLOGIT model performs better than other standard links in terms average DIC but is usually worse than GEV.
- SLASH model performs best when the true model is LOGIT (symmetric).

Simulation Study

Table : Percentage of best performance and average DIC under different model fitting among 200 replicates.

	True model: BSSMM-LOGIT				
	CLOGLOG	LOGLOG	GEV	SPLOGIT	SLASH
% best	16.0%	20.5%	24.0%	1.5%	38.0%
avg DIC	934.9	935.0	933.4	937.1	933.5
	True model: BSSMM-CLOGLOG				
	LOGIT	LOGLOG	GEV	SPLOGIT	SLASH
% best	0.0%	14.5%	68.5%	3.5%	13.5%
avg DIC	770.8	768.9	755.5	762.7	765.5
	True model: BSSMM-LOGLOG				
	LOGIT	CLOGLOG	GEV	SPLOGIT	SLASH
% best	0.5%	14.5%	70.5%	1.0%	13.5%
agv DIC	778.6	778.1	768.5	774.4	773.8

- GEV model is extremely robust against link misspecifications.
- SPLOGIT model performs better than other standard links in terms average DIC but is usually worse than GEV.
- SLASH model performs best when the true model is LOGIT (symmetric).

Case Study: DBS on attention reaction time

- To illustrate the technique applied to binary responses, we consider responses from a monkey performing the attention paradigm described in Smith et al. (2009).
- In order to receive a reward, the monkey must release the bar within a brief time window cued by the change in target color, which requires sustained attention of the animal.
- The behavioral data set for this experiment are composed of a time series of binary observations with a 1 corresponding to reward being delivered and a 0 corresponding to reward not being delivered at each trial, respectively.
- The goal of the experiment is to determine whether deep brain stimulation (DBS) allows the animal to recover its pre-fatigue level of performance. In this experiment, the monkey performed 1250 trials.
- The DBS was applied during 4 periods across trials 300-364, 498-598, 700-799 and 1000-1099.

Case Study: DBS on attention reaction time

- To illustrate the technique applied to binary responses, we consider responses from a monkey performing the attention paradigm described in Smith et al. (2009).
- In order to receive a reward, the monkey must release the bar within a brief time window cued by the change in target color, which requires sustained attention of the animal.
- The behavioral data set for this experiment are composed of a time series of binary observations with a 1 corresponding to reward being delivered and a 0 corresponding to reward not being delivered at each trial, respectively.
- The goal of the experiment is to determine whether deep brain stimulation (DBS) allows the animal to recover its pre-fatigue level of performance. In this experiment, the monkey performed 1250 trials.
- The DBS was applied during 4 periods across trials 300-364, 498-598, 700-799 and 1000-1099.

Case Study: DBS on attention reaction time

- To illustrate the technique applied to binary responses, we consider responses from a monkey performing the attention paradigm described in Smith et al. (2009).
- In order to receive a reward, the monkey must release the bar within a brief time window cued by the change in target color, which requires sustained attention of the animal.
- The behavioral data set for this experiment are composed of a time series of binary observations with a 1 corresponding to reward being delivered and a 0 corresponding to reward not being delivered at each trial, respectively.
- The goal of the experiment is to determine whether deep brain stimulation (DBS) allows the animal to recover its pre-fatigue level of performance. In this experiment, the monkey performed 1250 trials.
- The DBS was applied during 4 periods across trials 300-364, 498-598, 700-799 and 1000-1099.

Case Study: DBS on attention reaction time

- To illustrate the technique applied to binary responses, we consider responses from a monkey performing the attention paradigm described in Smith et al. (2009).
- In order to receive a reward, the monkey must release the bar within a brief time window cued by the change in target color, which requires sustained attention of the animal.
- The behavioral data set for this experiment are composed of a time series of binary observations with a 1 corresponding to reward being delivered and a 0 corresponding to reward not being delivered at each trial, respectively.
- The goal of the experiment is to determine whether deep brain stimulation (DBS) allows the animal to recover its pre-fatigue level of performance. In this experiment, the monkey performed 1250 trials.
- The DBS was applied during 4 periods across trials 300-364, 498-598, 700-799 and 1000-1099.

Case Study: DBS on attention reaction time

- To illustrate the technique applied to binary responses, we consider responses from a monkey performing the attention paradigm described in Smith et al. (2009).
- In order to receive a reward, the monkey must release the bar within a brief time window cued by the change in target color, which requires sustained attention of the animal.
- The behavioral data set for this experiment are composed of a time series of binary observations with a 1 corresponding to reward being delivered and a 0 corresponding to reward not being delivered at each trial, respectively.
- The goal of the experiment is to determine whether deep brain stimulation (DBS) allows the animal to recover its pre-fatigue level of performance. In this experiment, the monkey performed 1250 trials.
- The DBS was applied during 4 periods across trials 300-364, 498-598, 700-799 and 1000-1099.

Case Study: DBS on attention reaction time

- To illustrate the technique applied to binary responses, we consider responses from a monkey performing the attention paradigm described in Smith et al. (2009).
- In order to receive a reward, the monkey must release the bar within a brief time window cued by the change in target color, which requires sustained attention of the animal.
- The behavioral data set for this experiment are composed of a time series of binary observations with a 1 corresponding to reward being delivered and a 0 corresponding to reward not being delivered at each trial, respectively.
- The goal of the experiment is to determine whether deep brain stimulation (DBS) allows the animal to recover its pre-fatigue level of performance. In this experiment, the monkey performed 1250 trials.
- The DBS was applied during 4 periods across trials 300-364, 498-598, 700-799 and 1000-1099.

Case Study: DBS on attention reaction time

- We divide the results into periods when stimulation is applied ("ON") and not applied ("OFF"). There are 240 correct responses out of 367 trials during the "ON' periods and 501 correct responses from 883 trials during the "OFF" periods.

- For this data set we fit the Binary state space model with three standard link functions (LOGIT, CLOGLOG and LOGLOG), as well as three flexible link functions (SLASH, SPLOGIT and GEV).

- For each case, we conducted the MCMC simulation for 50000 iterations, and the first 10000 draws were discarded as a burn-in period. In order to reduce the autocorrelation between successive values of the simulated chain, we store every 5th values of the chain as the posterior samples.

Case Study: DBS on attention reaction time

- We divide the results into periods when stimulation is applied ("ON") and not applied ("OFF"). There are 240 correct responses out of 367 trials during the "ON' periods and 501 correct responses from 883 trials during the "OFF" periods.
- For this data set we fit the Binary state space model with three standard link functions (LOGIT, CLOGLOG and LOGLOG), as well as three flexible link functions (SLASH, SPLOGIT and GEV).
- For each case, we conducted the MCMC simulation for 50000 iterations, and the first 10000 draws were discarded as a burn-in period. In order to reduce the autocorrelation between successive values of the simulated chain, we store every 5th values of the chain as the posterior samples.

Case Study: DBS on attention reaction time

- We divide the results into periods when stimulation is applied ("ON") and not applied ("OFF"). There are 240 correct responses out of 367 trials during the "ON' periods and 501 correct responses from 883 trials during the "OFF" periods.

- For this data set we fit the Binary state space model with three standard link functions (LOGIT, CLOGLOG and LOGLOG), as well as three flexible link functions (SLASH, SPLOGIT and GEV).

- For each case, we conducted the MCMC simulation for 50000 iterations, and the first 10000 draws were discarded as a burn-in period. In order to reduce the autocorrelation between successive values of the simulated chain, we store every 5th values of the chain as the posterior samples.

Case Study: DBS on attention reaction time

- We divide the results into periods when stimulation is applied ("ON") and not applied ("OFF"). There are 240 correct responses out of 367 trials during the "ON' periods and 501 correct responses from 883 trials during the "OFF" periods.
- For this data set we fit the Binary state space model with three standard link functions (LOGIT, CLOGLOG and LOGLOG), as well as three flexible link functions (SLASH, SPLOGIT and GEV).
- For each case, we conducted the MCMC simulation for 50000 iterations, and the first 10000 draws were discarded as a burn-in period. In order to reduce the autocorrelation between successive values of the simulated chain, we store every 5th values of the chain as the posterior samples.

Case Study: DBS on attention reaction time

Table : Estimation results for monkey performance data set. First row: Posterior mean. Second row: Posterior 95% HPD interval in parentheses.

	Model					
	LOGIT	CLOGLOG	LOGLOG	SLASH	SPLOGIT	GEV
	0.9950	0.9967	0.9958	0.9947	0.9968	0.9941
δ	(0.9888,0.9999)	(0.9911,0.9999)	(0.9906,0.9999)	(0.9883,0.9999)	(0.9929,0.9999)	(0.9869,0.9999)
	0.0277	0.0113	0.0125	0.0126	0.0085	0.0074
τ^2	(0.0109,0.0488)	(0.0039,0.0195)	(0.0040,0.0232)	(0.0049,0.0230)	(0.0033,0.0148)	(0.0027,0.0137)
	–	–	–	5.0188	4.1812	-0.5483
$\nu/r/\xi$	–	–	–	(1.8556, 8.1409)	(1.9350,7.0834)	(-0.5998, -0.4478)

- For all the models considered here, the posterior means of δ are close to 1, showing higher persistence of the autoregressive parameter for states variables and thus in binary time series.

- $r = 4.1812$ and $\xi = -0.5483$ both indicate the data favors positively skewed link functions, which corresponds to LOGLOG among the standard links we considered here.

Case Study: DBS on attention reaction time

Table : Estimation results for monkey performance data set. First row: Posterior mean. Second row: Posterior 95% HPD interval in parentheses.

	Model					
	LOGIT	CLOGLOG	LOGLOG	SLASH	SPLOGIT	GEV
	0.9950	0.9967	0.9958	0.9947	0.9968	0.9941
δ	(0.9888,0.9999)	(0.9911,0.9999)	(0.9906,0.9999)	(0.9883,0.9999)	(0.9929,0.9999)	(0.9869,0.9999)
	0.0277	0.0113	0.0125	0.0126	0.0085	0.0074
τ^2	(0.0109,0.0488)	(0.0039,0.0195)	(0.0040,0.0232)	(0.0049,0.0230)	(0.0033,0.0148)	(0.0027,0.0137)
	–	–	–	5.0188	4.1812	-0.5483
$\nu/r/\xi$	–	–	–	(1.8556, 8.1409)	(1.9350,7.0834)	(-0.5998, -0.4478)

- For all the models considered here, the posterior means of δ are close to 1, showing higher persistence of the autoregressive parameter for states variables and thus in binary time series.

- $r = 4.1812$ and $\xi = -0.5483$ both indicate the data favors positively skewed link functions, which corresponds to LOGLOG among the standard links we considered here.

Case Study: DBS on attention reaction time

Table : Estimation results for monkey performance data set. First row: Posterior mean. Second row: Posterior 95% HPD interval in parentheses.

	Model					
	LOGIT	CLOGLOG	LOGLOG	SLASH	SPLOGIT	GEV
	0.9950	0.9967	0.9958	0.9947	0.9968	0.9941
δ	(0.9888,0.9999)	(0.9911,0.9999)	(0.9906,0.9999)	(0.9883,0.9999)	(0.9929,0.9999)	(0.9869,0.9999)
	0.0277	0.0113	0.0125	0.0126	0.0085	0.0074
τ^2	(0.0109,0.0488)	(0.0039,0.0195)	(0.0040,0.0232)	(0.0049,0.0230)	(0.0033,0.0148)	(0.0027,0.0137)
	–	–	–	5.0188	4.1812	-0.5483
$\nu/r/\xi$	–	–	–	(1.8556, 8.1409)	(1.9350,7.0834)	(-0.5998, -0.4478)

- For all the models considered here, the posterior means of δ are close to 1, showing higher persistence of the autoregressive parameter for states variables and thus in binary time series.

- $r = 4.1812$ and $\xi = -0.5483$ both indicate the data favors positively skewed link functions, which corresponds to LOGLOG among the standard links we considered here.

Case Study: DBS on attention reaction time

Table : Monkey performance data set. DIC: deviance information criterion.

Model	DIC	Rank
BSSM-LOGIT	1422.0	5
BSSM-CLOGLOG	1433.1	6
BSSM-LOGLOG	1412.5	1
BSSM-SLASH	1420.6	4
BSSM-SPLOGIT	1413.1	2
BSSM-GEV	1415.2	3

- The DIC measure selects the BSSM-LOGLOG as the best model for the monkey performance data set, whereas two of the flexible models BSSM-SPLOGIT and BSSM-GEV are close as well.
- This confirms our observation that the data supports positively skewed link function, namely the BSSM-LOGLOG standard link, as well as BSSM-GEV and BSSM-SPLOGIT with positive skewed parameters.

Case Study: DBS on attention reaction time

Table : Monkey performance data set. DIC: deviance information criterion.

Model	DIC	Rank
BSSM-LOGIT	1422.0	5
BSSM-CLOGLOG	1433.1	6
BSSM-LOGLOG	1412.5	1
BSSM-SLASH	1420.6	4
BSSM-SPLOGIT	1413.1	2
BSSM-GEV	1415.2	3

- The DIC measure selects the BSSM-LOGLOG as the best model for the monkey performance data set, whereas two of the flexible models BSSM-SPLOGIT and BSSM-GEV are close as well.
- This confirms our observation that the data supports positively skewed link function, namely the BSSM-LOGLOG standard link, as well as BSSM-GEV and BSSM-SPLOGIT with positive skewed parameters.

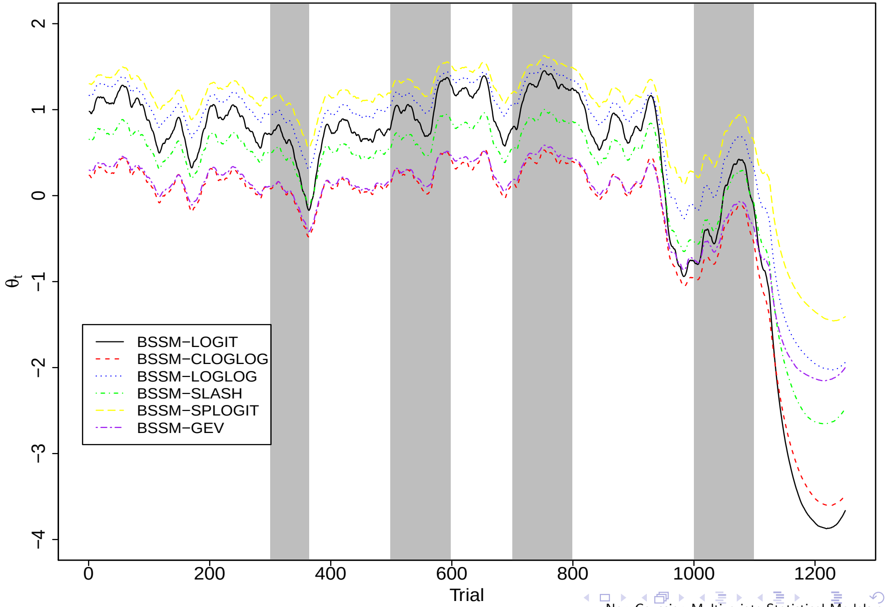
Case Study: DBS on attention reaction time

Table : Monkey performance data set. DIC: deviance information criterion.

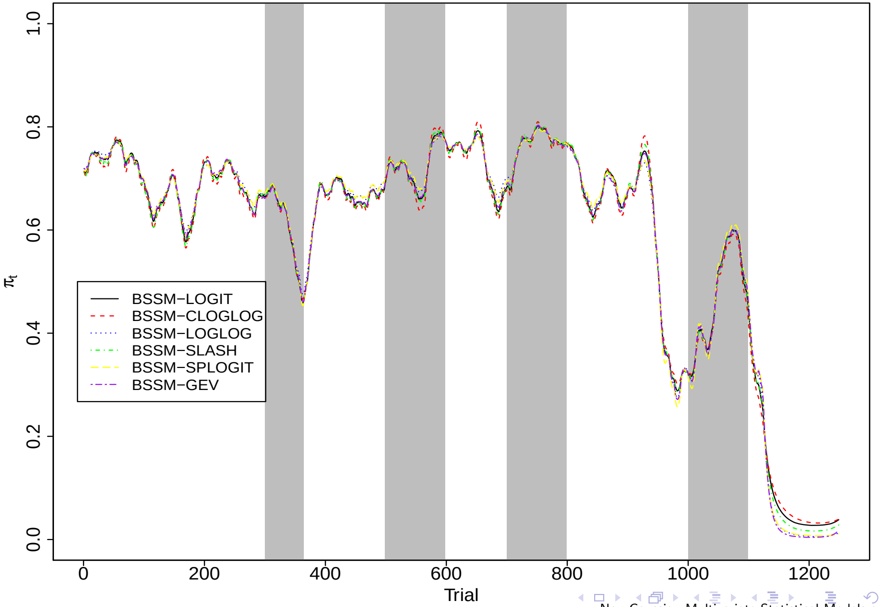
Model	DIC	Rank
BSSM-LOGIT	1422.0	5
BSSM-CLOGLOG	1433.1	6
BSSM-LOGLOG	1412.5	1
BSSM-SLASH	1420.6	4
BSSM-SPLOGIT	1413.1	2
BSSM-GEV	1415.2	3

- The DIC measure selects the BSSM-LOGLOG as the best model for the monkey performance data set, whereas two of the flexible models BSSM-SPLOGIT and BSSM-GEV are close as well.
- This confirms our observation that the data supports positively skewed link function, namely the BSSM-LOGLOG standard link, as well as BSSM-GEV and BSSM-SPLOGIT with positive skewed parameters.

Case Study: Posterior smoothed mean of θ_t



Case Study: Posterior smoothed mean of π_t



Conclusion

- We have proposed three flexible classes of state space mixed models for longitudinal bi- nary data using GEV, SPLOGIT and SMN distributions as extensions of Czado and Song (2008) and Abanto-Valle and Dey (2012).
- The flexibility in links is important to avoid link misspecification. The parameters controling the skewness and tail behavior are estimated along with model fitting.
- An attractive aspect of the model is that can be easily implemented, under a Bayesian perspective, via MCMC by using JAGS.
- Several extensions of the models are possible. First, we focus on binary observations here, but the setup can be extended to binomial and ordinal data.
- Also, other skewed links such as the skew normal or the skew Student-t can be compared with our proposed models. However, in such case, it is necessary to develop efficient sampler for the states variables.

Conclusion

- We have proposed three flexible classes of state space mixed models for longitudinal bi- nary data using GEV, SPLOGIT and SMN distributions as extensions of Czado and Song (2008) and Abanto-Valle and Dey (2012).
- The flexibility in links is important to avoid link misspecification. The parameters controling the skewness and tail behavior are estimated along with model fitting.
- An attractive aspect of the model is that can be easily implemented, under a Bayesian perspective, via MCMC by using JAGS.
- Several extensions of the models are possible. First, we focus on binary observations here, but the setup can be extended to binomial and ordinal data.
- Also, other skewed links such as the skew normal or the skew Student-t can be compared with our proposed models. However, in such case, it is necessary to develop efficient sampler for the states variables.

Conclusion

- We have proposed three flexible classes of state space mixed models for longitudinal bi- nary data using GEV, SPLOGIT and SMN distributions as extensions of Czado and Song (2008) and Abanto-Valle and Dey (2012).
- The flexibility in links is important to avoid link misspecification. The parameters controling the skewness and tail behavior are estimated along with model fitting.
- An attractive aspect of the model is that can be easily implemented, under a Bayesian perspective, via MCMC by using JAGS.
- Several extensions of the models are possible. First, we focus on binary observations here, but the setup can be extended to binomial and ordinal data.
- Also, other skewed links such as the skew normal or the skew Student-t can be compared with our proposed models. However, in such case, it is necessary to develop efficient sampler for the states variables.

Conclusion

- We have proposed three flexible classes of state space mixed models for longitudinal bi- nary data using GEV, SPLOGIT and SMN distributions as extensions of Czado and Song (2008) and Abanto-Valle and Dey (2012).
- The flexibility in links is important to avoid link misspecification. The parameters controling the skewness and tail behavior are estimated along with model fitting.
- An attractive aspect of the model is that can be easily implemented, under a Bayesian perspective, via MCMC by using JAGS.
- Several extensions of the models are possible. First, we focus on binary observations here, but the setup can be extended to binomial and ordinal data.
- Also, other skewed links such as the skew normal or the skew Student-t can be compared with our proposed models. However, in such case, it is necessary to develop efficient sampler for the states variables.

Conclusion

- We have proposed three flexible classes of state space mixed models for longitudinal bi- nary data using GEV, SPLOGIT and SMN distributions as extensions of Czado and Song (2008) and Abanto-Valle and Dey (2012).
- The flexibility in links is important to avoid link misspecification. The parameters controling the skewness and tail behavior are estimated along with model fitting.
- An attractive aspect of the model is that can be easily implemented, under a Bayesian perspective, via MCMC by using JAGS.
- Several extensions of the models are possible. First, we focus on binary observations here, but the setup can be extended to binomial and ordinal data.
- Also, other skewed links such as the skew normal or the skew Student-t can be compared with our proposed models. However, in such case, it is necessary to develop efficient sampler for the states variables.

Conclusion

- We have proposed three flexible classes of state space mixed models for longitudinal bi- nary data using GEV, SPLOGIT and SMN distributions as extensions of Czado and Song (2008) and Abanto-Valle and Dey (2012).
- The flexibility in links is important to avoid link misspecification. The parameters controling the skewness and tail behavior are estimated along with model fitting.
- An attractive aspect of the model is that can be easily implemented, under a Bayesian perspective, via MCMC by using JAGS.
- Several extensions of the models are possible. First, we focus on binary observations here, but the setup can be extended to binomial and ordinal data.
- Also, other skewed links such as the skew normal or the skew Student-t can be compared with our proposed models. However, in such case, it is necessary to develop efficient sampler for the states variables.

Selected References

- Abanto-Valle C, Dey D. 2012. State space mixed models for binary responses with scale mixture of normal distributions links. Technical Report 22, Department of Statistics, University of Connecticut.
- Czado C, Santner T. 1992. The effect of link misspecification on binary regression inference. Journal of Statistical Planning and Inference 33: 213231.
- Jiang X, Dey D, Prunier R, Wilson A, Holsinger K. 2012. A new class of flexible link function with application to spatially correlated species co-occurrence in cape floristic region. Technical report, Department of Statistics, University of Connecticut.
- Wang X, Dey DK. 2010. Generalized extreme value regression for binary response data: An application to B2B electronic payments system adoption. The Annals of Applied Statistics 4: 20002023.