Bayesian Skew-Normal and Skew-t Models of Birth Weight and Food Security

Carter Allen[1]; Brian Neelon, PhD[1]; Sara E. Benjamin-Neelon, PhD, JD, MPH[2]

[1]Department of Public Health Sciences, Medical University of South Carolina; [2]Bloomberg School of Public Health, Johns Hopkins University

Objectives

We examine the properties of skew-normal and skew-t models from both a Bayesian and frequentist perspective, and investigate the computational tools available for fitting these models. We apply skew-normal and skew-t models to data from the Nurture study, a cohort of mothers who gave birth between 2013 and 2016, where we seek to model the effect of food security during pregnancy on birth weight.

Introduction

In many applications of classical linear regression, the distribution of residuals exibits non- normal qualities such as skewness or heavy tails, making the assumption of normal error terms difficult to justify. The common statistical suggestion in these cases is to implement a transformation of the response variable, but this can result in a loss of interpretability. The skew-elliptical family is a broad class of probability distributions that contain the normal distribution as a special case and allow for flexible modeling when data exhibit skewness.

Definitions

Let ϕ and Φ be the standard normal pdf and cdf, respectively. Azzalini (1985) defined the density of a skew-normal random variable Z follows.

$$f(z;\lambda) = 2\phi(z)\Phi(\lambda z)$$

Similar to the construction of the familiar student's t random variable, Azzalini (2014) defines a skew-t random variable as the ratio of a skew normal and the square root of a χ^2 divided by its degrees of freedom. The resultant density is

$$t(x;\lambda,\nu) = 2t_0(x;\nu)T_0(\lambda x\sqrt{\frac{\nu+1}{\nu+x^2}};\nu+1)$$

where t_0 and T_0 are the density and mass functions of the student's t distribution, respectively. A linear regression model with skew error terms is a modification of classical regression with the modification of either assuming $\mathcal{SN}$ or $\mathcal{ST}$ random errors.

Motivation

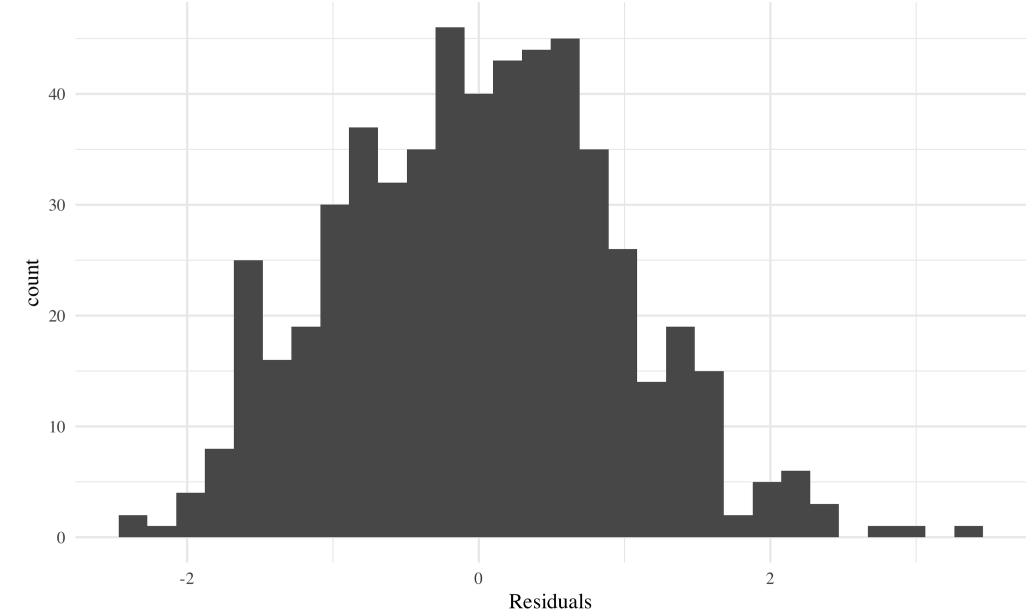

Figure 1: Residuals in Nurture study exhibiting skewness. Pearson's skewness = 0.17. Shapiro-Wilk p-value = 0.08.

Full Conditionals

Let $Y_i = \mathbf{x}_i\beta + \psi z_i + \sigma\epsilon$ with $Z \sim N_+(0,1)$, and $\epsilon \sim N(0,1)$. Define $\mathbf{X}^* = [\mathbf{X}|\mathbf{z}]$, and $\beta^* = [\beta_0, \beta_1, ..., \beta_p, \psi]$. Then,

$$\beta^*|Y,\mathbf{X}^*,\tau \sim N_{p+2}\left(\frac{(T_0\beta_0 + \tau\mathbf{X}^{*T}Y)}{\tau\mathbf{X}^{*T}\mathbf{X}^* + T_0}, \tau\mathbf{X}^{*T}\mathbf{X}^* + T_0\right)$$

$$\tau|\beta^*,\mathbf{X}^*,Y \sim \Gamma(n/2+\alpha, \frac{1}{2}(Y-\mathbf{X}^*\beta^*)^T(Y-\mathbf{X}^*\beta^*)+b)$$

$$z_i|y_i,\mathbf{x}_i,\beta,\tau \sim N_+\left(\frac{\psi\tau(y_i-\mathbf{x}_i\beta)}{\tau\psi^2+1}, \frac{1}{\tau\psi^2+1}\right)$$

Important Results

Through simulation studies, we validated Bayesian Gibbs samplers for skew-normal and skew-t error regression models. We used these models to find a significant increase in birth weight associated with very low food security during pregnancy using data from the Nurture study.

Modeling Approaches

1. *Maximum Likelihood Estimation*: For a random sample $Y_1, Y_2, ..., Y_n \overset{iid}{\sim} \mathcal{SN}(\xi, \omega^2, \alpha)$, where $\xi = \mathbf{x}^T\beta$ for a collection of predictors $x_1, x_2, ..., x_n$ and a vector of unknown regression coefficients $\beta \in \mathbb{R}^p$. Azzalini (2014) describes procedures for obtaining MLE estimates of β and α, our primary parameters of interest. Azzalini's `R` package `sn` contains a function `selm` for fitting regression models with $\mathcal{SN}$ or $\mathcal{ST}$ random errors.
2. *Bayesian Gibbs Sampler*: We introduce the following stochastic representation of the skew normal distribution

$$Y_i = \mathbf{x}_i\beta + \psi z_i + \sigma\epsilon$$

where $z_i \sim N_+(0,1)$ and $\epsilon \sim N(0,1)$. The marginal density of Y_i integrating over z_i and ϵ is $\mathcal{SN}(\mathbf{x}_i\beta, \omega^2)$. We use this stochastic representation to obtain full conditionals for all parameters.

Gibbs Sampler Simulation

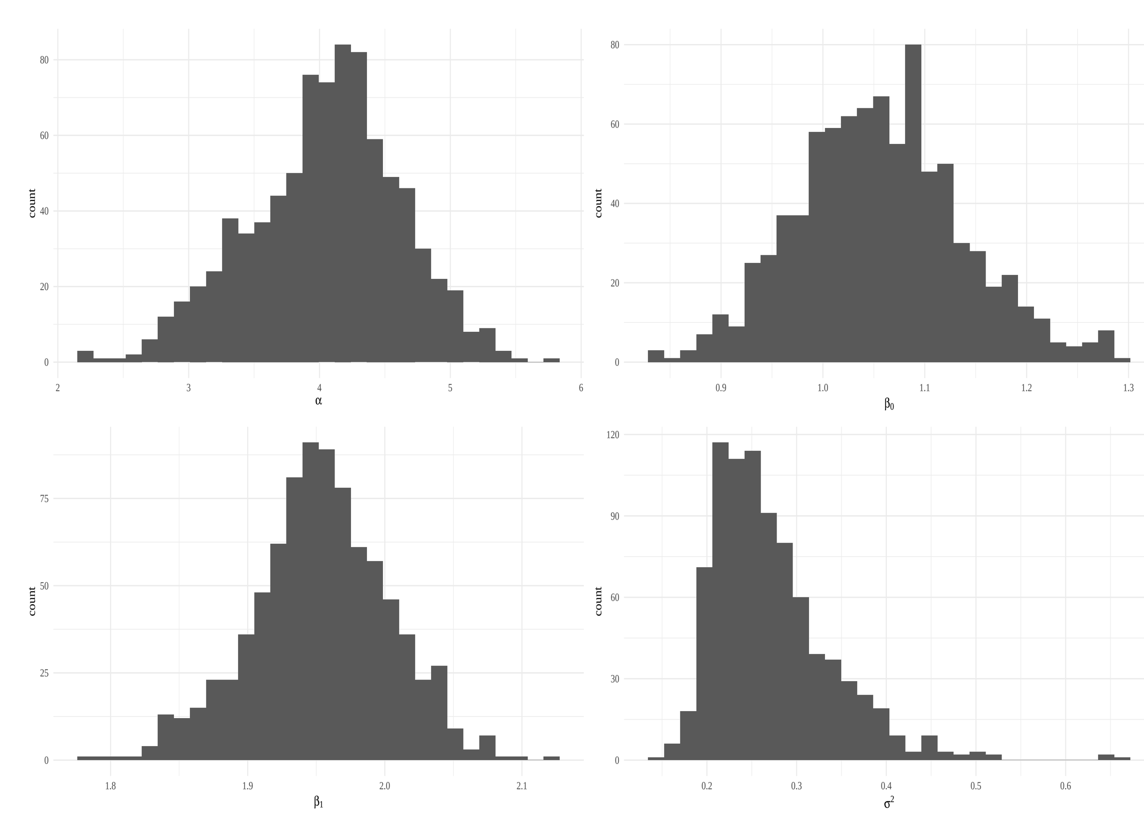

Figure 2: Posterior distributions of α, β_0, β_1, σ^2 in simulation.

Param.	True	MLE	Gibbs	SLR
α	4.00	0.3928	4.016	–
β_0	1.00	1.013	1.009	2.566
β_1	2.00	2.006	2.01	2.081
σ^2	0.235	0.257	0.257	1.753

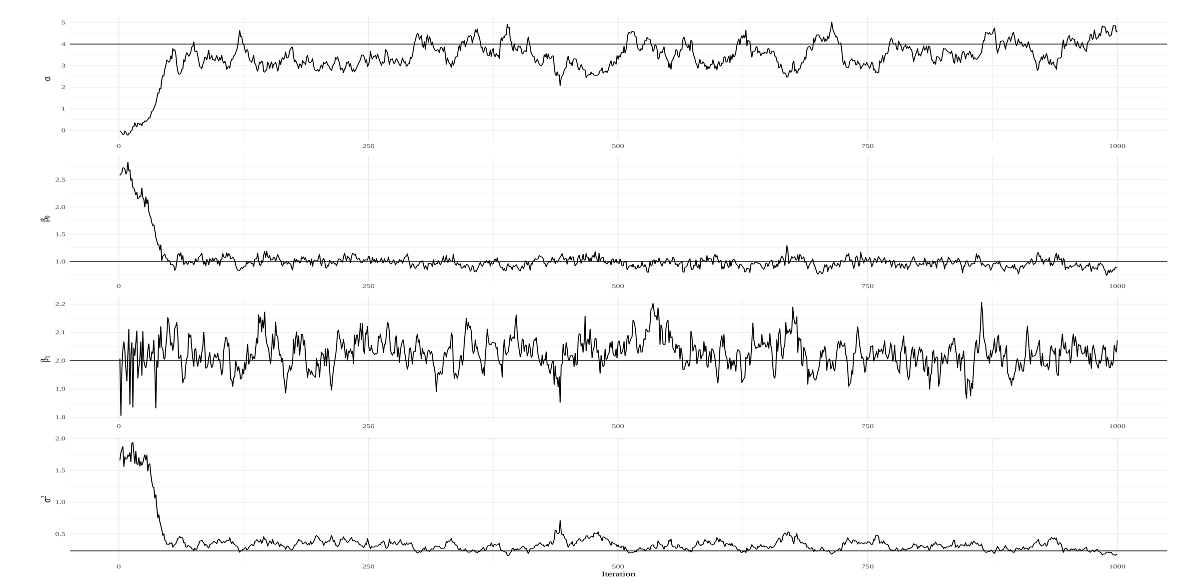

Figure 3: Trace plots of parameter estimates in simulation

Modeling Results

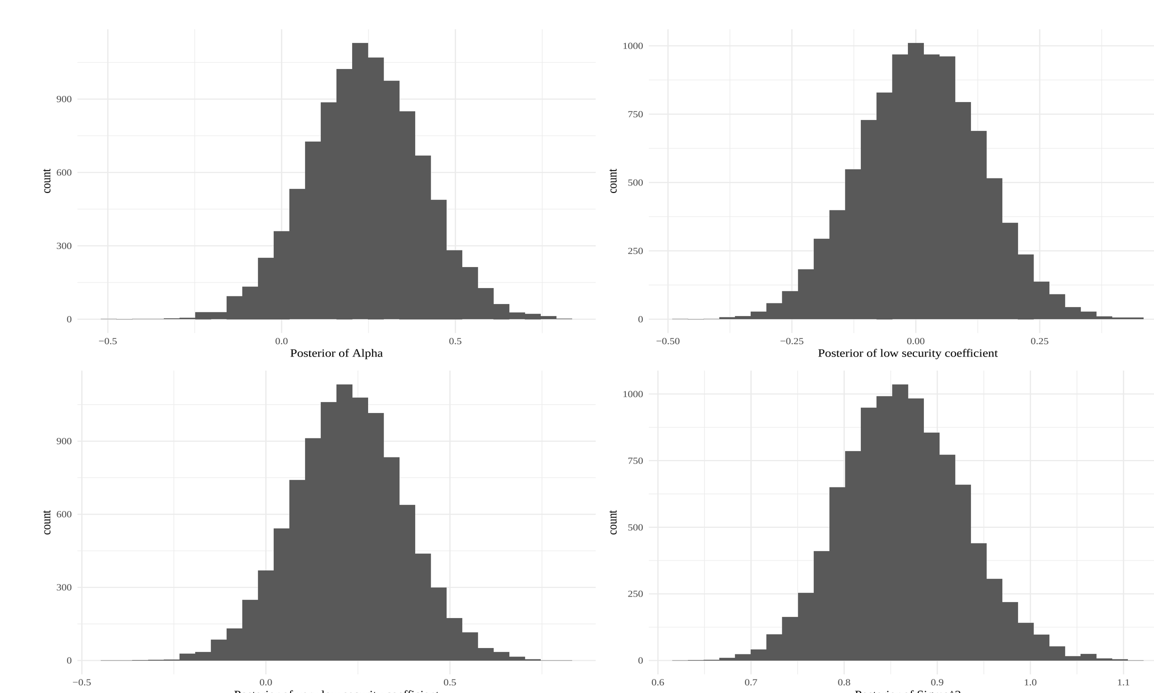

Figure 4: Posterior distributions of model parameters

Param.	Est.	95% CI
α	0.239	(-0.169,0.624)
β_{low}	0.009	(-0.225,0.245)
$\beta_{v.low}$	0.230	(0.020,0.579)

References

(1) Azzalini, S. (1985). A class of distributions which includes the normal ones. SJS; **(2) Fruhwirth-Schnatter, S and Pyne, S. (2010)**. Bayesian inference for finite mixtures of univariate ... Biostatistics; **(3) Benjamin-Neelon SE, Ostbye T, Bennett GG, et al.** Cohort profile for the Nurture Observational Study ... BMJ Open 2017; **(4) Neelon, B. (2015)** Bayesian two-part spatial models

Funding: *This work is supported by a grant from the NIH (**R01DK094841**).*

Further Resources

https://carter-allen.github.io/SN