Web-based Supplementary Materials for 'Linear and non-linear mixed-effects models for censored HIV viral loads using normal/independent distributions' by Victor H. Lachos, Dipankar Bandyopadhyay and Dipak K. Dey

Web Appendix A : Densities and conditional posteriors for specific NI distributions in the linear case

- *The multivariate Student–t distribution*, denoted by $t_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu)$, where ν is the degrees of freedom, U is distributed as $Gamma(\nu/2, \nu/2)$, with $\nu > 0$ and it includes the normal case when $\nu \to \infty$. The pdf of $\mathbf{y}$ takes the following form:

$$T(\mathbf{y}|\boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu) = \frac{\Gamma(\frac{p+\nu}{2})}{\Gamma(\frac{\nu}{2})\pi^{p/2}} \nu^{-p/2} |\boldsymbol{\Sigma}|^{-1/2} \left(1 + \frac{d}{\nu}\right)^{-(p+\nu)/2}, \mathbf{y} \in \mathbb{R}^p,$$

where $\Gamma(.)$ is the standard gamma function and $d = (\mathbf{y} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\mathbf{y} - \boldsymbol{\mu})$ is the Mahalanobis distance. The posteriors are:

$$\begin{aligned} u_i|\mathbf{y}, \mathbf{b}, \boldsymbol{\theta} &\sim Gamma((n_i + q + \nu)/2, \nu/2 + \lambda_i/2), \\ \pi(\nu|\mathbf{y}, \mathbf{b}, \mathbf{u}, \boldsymbol{\theta}_{(-\nu)}) &\propto \frac{(\nu/2)^{n\nu/2}}{(\Gamma(\nu/2))^n} \exp\left\{-\nu \left[(1/2)\sum_{i=1}^n (u_i - \log(u_i)) + \gamma\right]\right\} \mathbb{I}_{\{(2,\infty)\}}(\nu) \end{aligned}$$

and $\lambda_i = \frac{1}{\sigma^2} \left(\mathbf{y}_i - \mathbf{X}_i\boldsymbol{\beta} - \mathbf{Z}_i\mathbf{b}_i\right)^\top \left(\mathbf{y}_i - \mathbf{X}_i\boldsymbol{\beta} - \mathbf{Z}_i\mathbf{b}_i\right) + \mathbf{b}_i^\top \mathbf{D}^{-1}\mathbf{b}_i$. As $\pi(\nu|\cdots)$ do not have a standard form, we can generate a sample from this distribution using the Metropolis-Hastings path with a *log-normal* proposal density (see Liu, 1996).

- *The multivariate slash distribution*, denoted by $SL_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu)$, arises when the distribution of U is $Beta(\nu, 1)$, with $u \in (0, 1)$, $\nu > 0$ and it includes the normal case when $\nu \to \infty$. The pdf is given by $SL(\mathbf{y}|\boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu) = \nu \int_0^1 u^{\nu-1} \phi_p(\mathbf{y}; \boldsymbol{\mu}, u^{-1}\boldsymbol{\Sigma}) du, \quad \mathbf{y} \in \mathbb{R}^p$. The posteriors are:

$$\begin{aligned} u_i|\mathbf{y}, \mathbf{b}, \boldsymbol{\theta} &\sim TGamma((n_i + q + 2\nu)/2, \lambda_i/2; (0, 1)), \\ \nu|\mathbf{y}, \mathbf{b}, \mathbf{u}, \boldsymbol{\theta}_{(-\nu)} &\sim Gamma(n + a, b - \sum_{i=1}^n \log u_i), \end{aligned}$$

where $TGamma(a, b; (0,1))$ denotes the Gamma distribution with parameters a and b truncated in the interval $(0,1)$.

- *The multivariate contaminated normal distribution*, denoted by $CN_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu, \rho)$, where $\nu, \rho \in (0,1)$. Here, U is a discrete random variable taking one of two states 1 or ρ, with probability function given by $h(u|\boldsymbol{\nu}) = \nu\mathbb{I}_{\{\rho\}}(u) + (1-\nu)\mathbb{I}_{\{1\}}(u)$, where $\boldsymbol{\nu} = (\nu, \rho)$ and $\mathbb{I}_{\{\tau\}}(u)$ is the indicator function of the set τ whose value equals 1 if $u \in \tau$ and 0 elsewhere. The associated density is $CN(\mathbf{y}|\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\nu}) = \nu\phi_p(\mathbf{y}; \boldsymbol{\mu}, \rho^{-1}\boldsymbol{\Sigma}) + (1-\nu)\phi_p(\mathbf{y}; \boldsymbol{\mu}, \boldsymbol{\Sigma})$. Parameter ν can be interpreted as the proportion of outliers while ρ is a scale factor and it includes the normal case when $\rho \to 1$. The posteriors are:

$$\mathrm{P}(u_i = \rho|\cdots) = 1 - \mathrm{P}(u_i = 1|\cdots) = \eta_i/(\eta_i + \zeta_i),$$

where

$$\eta_i = \nu\rho^{(n_i+q)/2} \exp\left\{-(1/2)\lambda_i\rho\right\} \text{ and } \zeta_i = (1-\nu)\exp\left\{-(1/2)\lambda_i\right\},$$

$$\nu|\cdots \sim \mathrm{Beta}(\nu_0 + m_\rho, n - m_\rho + \nu_1),$$

where

$$m_\rho = \frac{n - \sum_{i=1}^n u_i}{1-\rho}$$

is the cardinality of the set $\{i;\ u_i = \rho\}$. The full conditional for ρ is given by

$$\pi(\rho|\cdots) \propto \nu^{m_\rho}(1-\nu)^{n-m_\rho}\rho^{\rho_0-1}(1-\rho)^{\rho_1-1},$$

which does not have a closed form. An interesting Metropolis–Hastings method to update from ρ is described in Rosa et. al. (2003).

Web Appendix B : Proof of Theorem 1

Proof. For simplicity we omit the subindex i. Thus, for $\boldsymbol{\beta}$ fixed and based on first-order Taylor expansion of the function η around $\widetilde{\mathbf{b}}$, we have that

$$\mathbf{y} - \eta(\mathbf{A}\boldsymbol{\beta} + \mathbf{B}\mathbf{b}, \mathbf{X}) \approx \mathbf{y} - [\eta(\mathbf{A}\boldsymbol{\beta} + \mathbf{B}\widetilde{\mathbf{b}}, \mathbf{X}) + \widetilde{\mathbf{H}}\mathbf{b} - \widetilde{\mathbf{H}}\widetilde{\mathbf{b}}].$$

Then from (15) and (16), we have

$$\mathbf{y} - [\eta(\mathbf{A}\boldsymbol{\beta} + \widetilde{\mathbf{b}}, \mathbf{X}) + \widetilde{\mathbf{H}}\mathbf{b} - \widetilde{\mathbf{H}}\widetilde{\mathbf{b}}] \mid \mathbf{b} \dot{\sim} NI_n(\mathbf{0}, \sigma_e^2\mathbf{I}, H)$$

and the approximate conditional distribution of $\mathbf{y}$ is

$$\mathbf{y} \mid \mathbf{b} \dot{\sim} NI_n(\eta(\mathbf{A}\boldsymbol{\beta} + \mathbf{B}\widetilde{\mathbf{b}}, \mathbf{X}) - \widetilde{\mathbf{H}}\widetilde{\mathbf{b}} + \widetilde{\mathbf{H}}\mathbf{b}, \sigma_e^2\mathbf{I}, H), \tag{1}$$

or equivalently

$$\mathbf{y} \mid \mathbf{b}, u \dot{\sim} \mathrm{N}_n(\eta(\mathbf{A}\boldsymbol{\beta} + \mathbf{B}\widetilde{\mathbf{b}}, \mathbf{X}) - \widetilde{\mathbf{H}}\widetilde{\mathbf{b}} + \widetilde{\mathbf{H}}\mathbf{b}, u^{-1}\sigma_e^2\mathbf{I}),$$

where $u \sim h(.|\boldsymbol{\nu})$. The rest of the proof follows by noting that

$$f(\mathbf{y} \,|\boldsymbol{\theta}, \widetilde{\mathbf{b}}) \approx \int_0^\infty \int_{\mathbb{R}^q} \phi_n(\mathbf{y}; \eta(\mathbf{A}\boldsymbol{\beta}+\mathbf{B}\widetilde{\mathbf{b}}, \mathbf{X})-\widetilde{\mathbf{H}}\widetilde{\mathbf{b}}+\widetilde{\mathbf{H}}\mathbf{b}, u^{-1}\sigma_e^2\mathbf{I})\phi_q(\mathbf{b}; \mathbf{0}, u^{-1}\mathbf{D})h(u|\boldsymbol{\nu})d\mathbf{b}du,$$

which can be easily accomplished by using successively Lemma 2 given in Arellano-Valle et. al., (2005). □

References

[1] Arellano-Valle, R. B., Bolfarine, H. and Lachos, V. H. (2005). Skew-normal linear mixed models. *Journal of Data Science,* **3**, 415–438.

[2] Liu, C. (1996). Bayesian robust multivariate linear regression with incomplete data. *Journal of the American Statistical Association* **91**, 1219-1227.

[3] Rosa, G. J. M., Padovani, C. R., Gianola, D. (2003). Robust linear mixed models with normal/independent distributions and Bayesian MCMC implementation. *Biometrical Journal* **45**, 573–590.

Web Figure 1

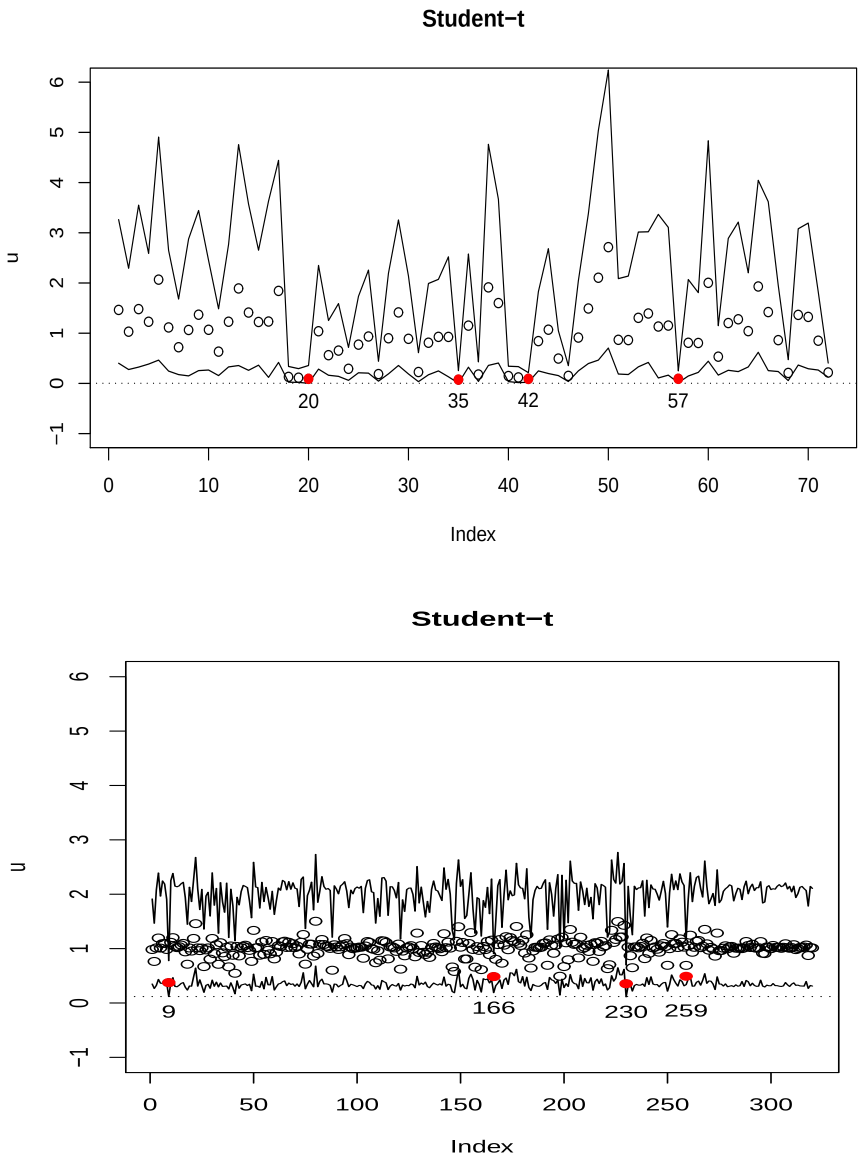

Figure 1: Estimated weight $\widehat{u}_i$ and 95% CI for the UTI data (Upper panel) and the AIEDRP data (Lower panel). Influential observations determined from Kullback-Leibler divergence measures are marked in bold red circles. The horizontal lines correspond to the lowest 2.5% CI. Both figures reveal that the Student–t model attributes small weight to the influential/outlying observations.

Web Figure 2

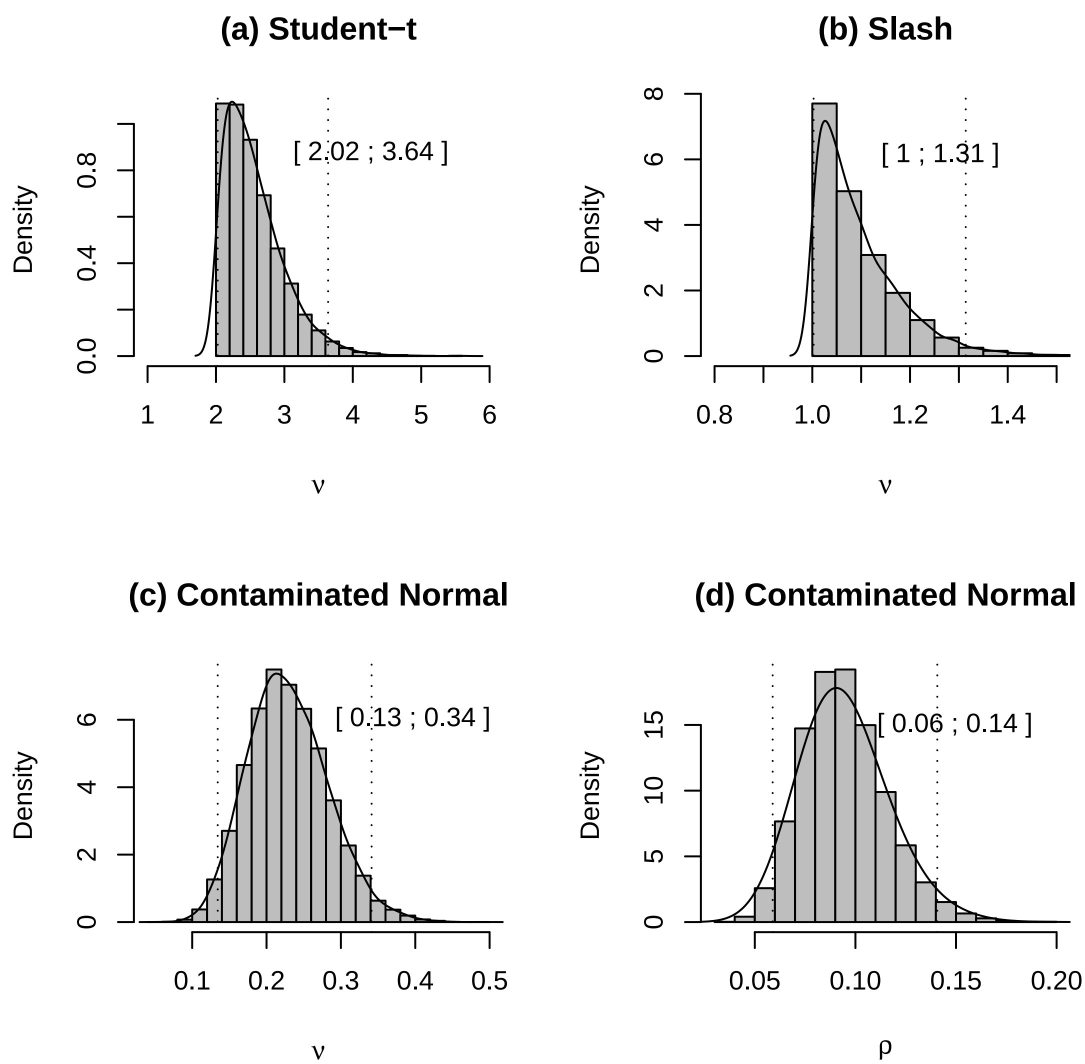

Figure 2: Marginal posterior densities of $\boldsymbol{\nu}$ for members of the NI class, (a) Student–t, (b) slash and (c) contaminated normal (CN) distributions for the UTI data. Panel (d) plots densities for the parameter ρ for the CN sub-class. The vertical lines and values between [;] represents 95% Credible Intervals (CI).

Web Figure 3

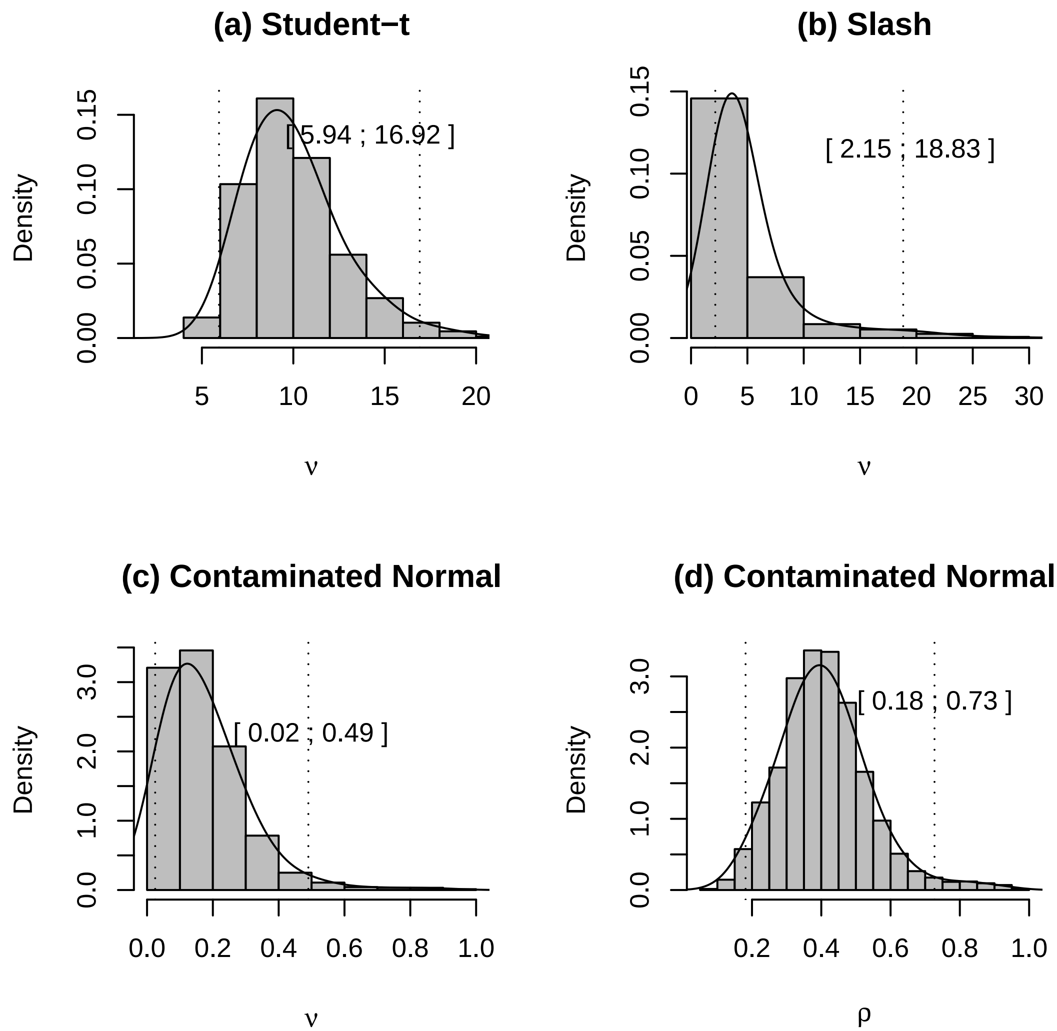

Figure 3: Marginal posterior densities of $\boldsymbol{\nu}$ for members of the NI class, (a) Student–t, (b) slash and (c) contaminated normal (CN) distributions for the AIEDRP data. Panel (d) plots densities for the parameter ρ for the CN sub-class. The vertical lines and values between [;] represents 95% Credible Intervals (CI).

Web Figure 4

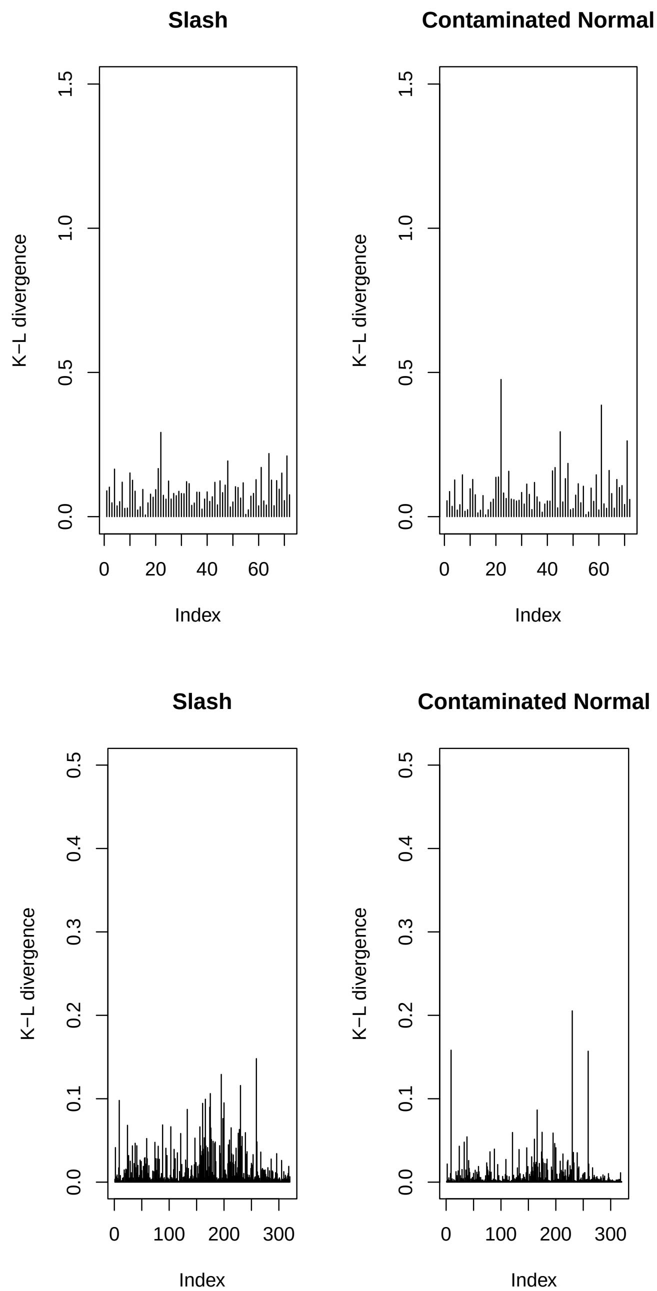

Figure 4: Index plots of $K(P, P_{(-i)})$. The upper 2 plots are for the UTI data while the lower 2 plots are for the AIEDRP data.

Web Figure 5

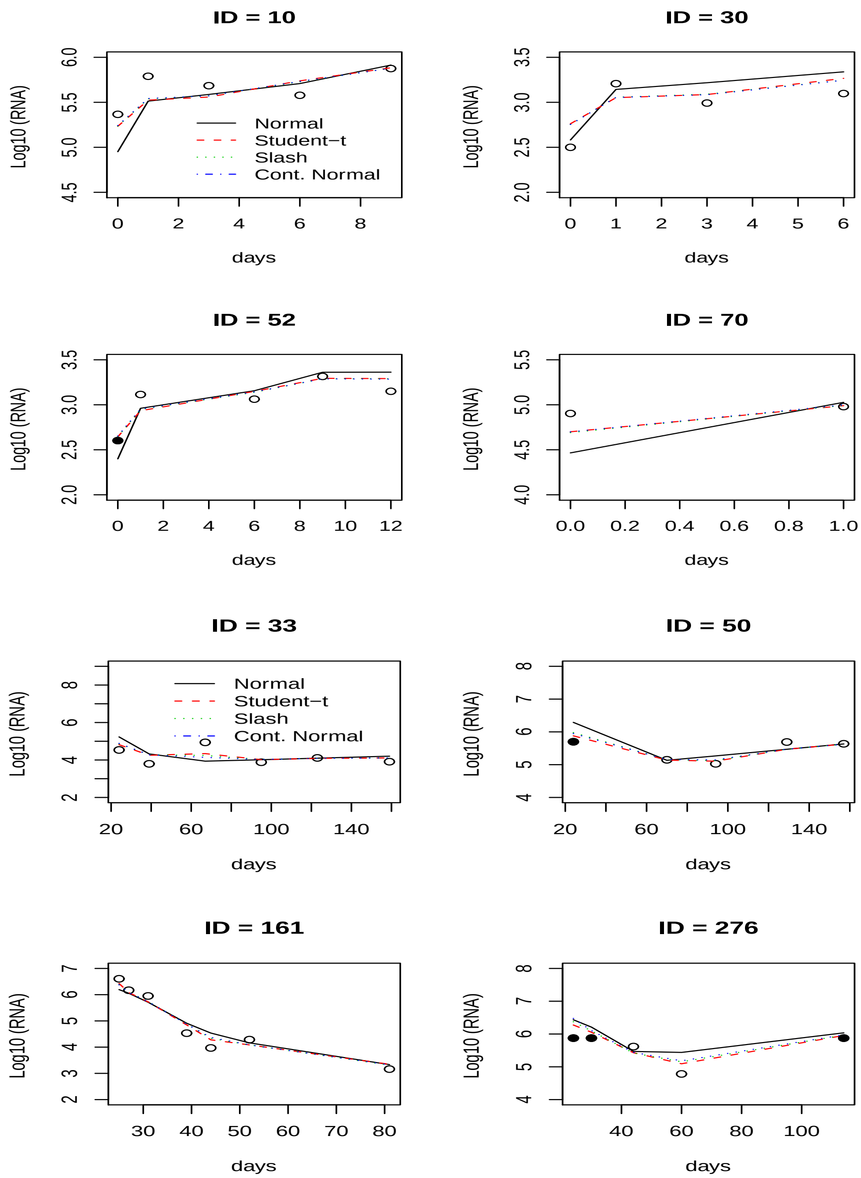

Figure 5: Estimated viral load trajectories (in $\log_{10}$ scale) for randomly selected subjects using the NI class for the UTI data (upper four plots) and AIEDRP data (lower four plots). The observed data are indicated by blank circles, while the censored ones by filled circles.

Web Figure 6

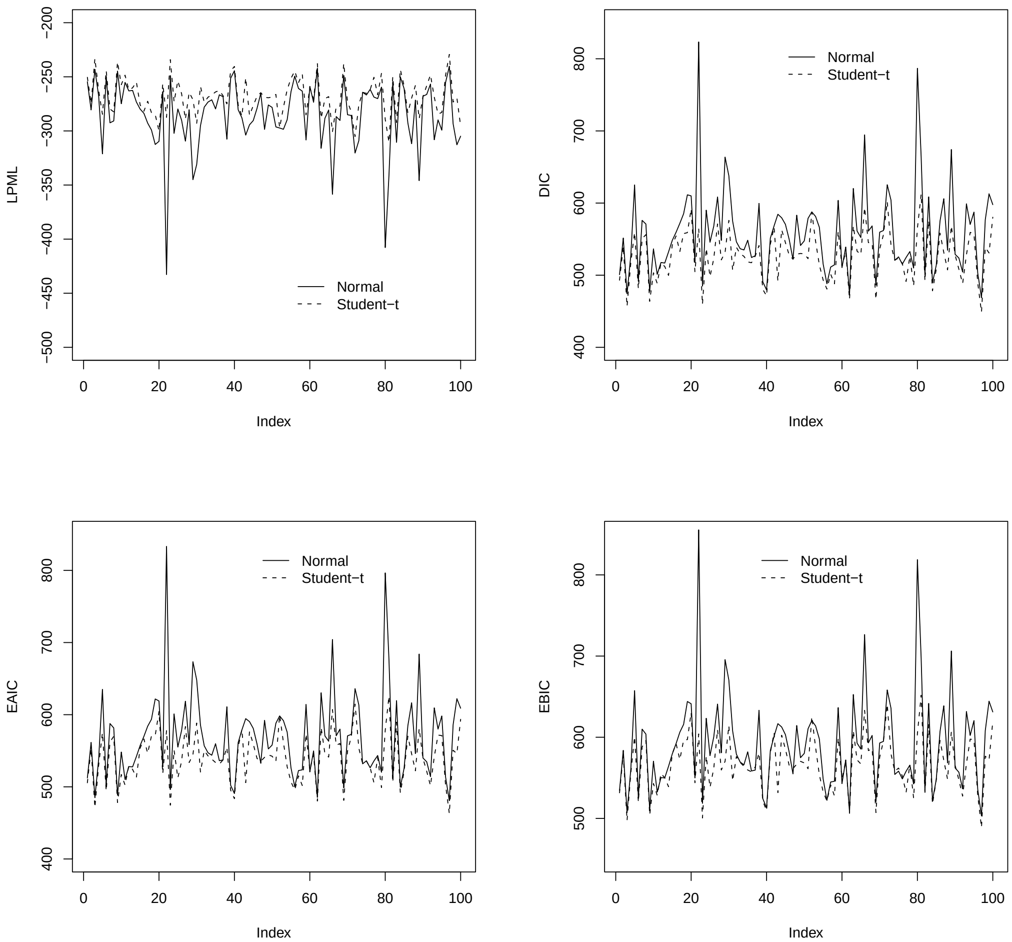

Figure 6: LPML, DIC, EAIC, EBIC values from 100 samples generated under Normal and Student–t models where proportion of censoring is 40%.