A Parametric Quantile Regression Model for Asymmetric Response Variables on the Real Line

Diego I. Gallardo [1] , **Marcelo Bourguignon** [2] , **Christian E. Galarza** [3] and **Héctor W. Gómez** [4,*]

[1] Departamento de Matemática, Facultad de Ingeniería, Universidad de Atacama, Copiapó 1530000, Chile; diego.gallardo@uda.cl
[2] Departamento de Estatística, Universidade Federal do Rio Grande do Norte, Natal, RN 59000-000, Brazil; ufrn00609@ufrn.edu.br
[3] Departamento de Matemáticas, Escuela Superior Politécnica del Litoral, ESPOL, Guayaquil 090150, Ecuador; chedgala@espol.edu.ec
[4] Departamento de Matemáticas, Facultad de Ciencias Básicas, Universidad de Antofagasta, Antofagasta 1240000, Chile
[*] Correspondence: hector.gomez@uantof.cl; Tel.: +56-55-2637-278

Abstract: In this paper, we introduce a novel parametric quantile regression model for asymmetric response variables, where the response variable follows a power skew-normal distribution. By considering a new convenient parametrization, these distribution results are very useful for modeling different quantiles of a response variable on the real line. The maximum likelihood method is employed to estimate the model parameters. Besides, we present a local influence study under different perturbation settings. Some numerical results of the estimators in finite samples are illustrated. In order to illustrate the potential for practice of our model, we apply it to a real dataset.

Keywords: asymmetric data; parametric inference; quantile regression model; skew-normal distributions

1. Introduction

Frequently, in real life, we find continuous data on the real line that are asymmetrical; these data cannot be modeled by known symmetric distributions as the normal, Student-*t*, Cauchy, Laplace, and logistic distributions. It is therefore more interesting to propose more flexible models that will be useful for modeling highly skewed data which arises in several areas.

In this context, the seminal work in Azzalini [1] introduces a skew-symmetric family of distributions, where this last is established by using a symmetric distribution as a kernel. When this last follows a normal distribution, it rises the well-know skew-normal (SN) distribution. The SN distribution has a skewness parameter which makes possible to have a reasonable model for a skewed distribution. Furthermore, the SN distributions include the normal distribution and possesses several properties which coincide or are similar to the ones of the normal distribution (Azzalini [1,2]). However, the SN distribution is limited in terms of flexibility, that is, for moderate values of the skewness parameter nearly all the mass accumulates either on the positive or negative real line, as determined by the sign of the skewness parameter. In such cases, the SN distribution closely resembles the half-normal density, with a nearly linear shape in the side with smaller mass (Arellano-Valle et al. [3]).

Another alternative to model skewed data is using the family of power-symmetric distributions (see Pewsey et al. [4]) of which the most widely used is the power-normal (PN) distribution. Some references where this family is discussed are Lehmann [5], Durrans [6], Gupta and Gupta [7],

Castillo et al. [8], among others. In a series of papers by Martínez-Flórez et al. ([9–13]) extensions and applications of the PN distribution can be found.

An unification of the SN and PN distributions was proposed by Martínez-Flórez et al. [11], namely the power skew-normal (PSN) which is a generalization of the SN and PN distributions. Even though sample information about the SN distribution has been widely studied, there is not the same scope for the PSN distribution, which being a generalization of the first one, has characteristics of interest such as: (i) the SN and PN distributions as particular cases, and (ii) the PSN distribution provides greater range for skewness and kurtosis coefficients compared with the SN distribution (see Table 1), being more flexible to model highly skewed data, which arises frequently in many practical situations. However, the expectation and variance of the PSN distribution cannot be expressed in closed form (have complicated forms), which makes these distributions unsuitable for regression modeling (Martínez-Flórez et al. [11]). Fortunately, the cumulative distribution function (cdf) of the PSN distribution has a simple form that depends on Owen's T function (to be defined in the next section). This facilitates the calculation of the quantile function (inverse of the cdf), allowing its utilization in the quantile regression (QR) framework. Quantile regression quantifies the association of the explanatory variables with a given quantile of a dependent variable. In this study, we propose a quantile linear regression model based on the PSN distribution, adopting a new parametrization of this model indexed by the quantile, precision and shape parameters. In particular, for this work, inference is conducted via maximum likelihood.

Table 1. Range for skewness and kurtosis coefficients for SN, PN and PSN models.

Coefficient	SN	PN	PSN
Skewness	$(-0.9953, 0.9953)$	$[-0.6115, 0.9007]$	$[-1.6476, 0.9953)$
Kurtosis	$[3, 3.8692)$	$[1.7170, 4.3556]$	$[1.4672, 5.4386]$

The rest of the paper proceeds as follows. In Section 2, we introduce a new parameterization of the PSN distribution that is indexed by the location, precision and shape parameters and its association with a quantile regression model. In addition, elements related to the maximum likelihood (ML) method are presented as well. Section 3 presents local influence measures under three different perturbation schemes, whereas in Section 4 a real data analysis is conducted in order to show the applicability of our proposed reparametrized PSN (RPSN) based QR model. Final section summarizes the contributions of the paper.

2. A PSN Distribution Parameterized by Its Quantile Parameter, and Its Associated Quantile Regression Model

In this section, we briefly study the PSN distribution based on Martínez-Flórez et al. [11]. We introduce a RPSN distribution which is characterized by its quantile, which allows us to use this distribution in the context of QR models.

The probability density function (pdf) of the PSN distribution is given by

$$f(y; \boldsymbol{\theta}) = \frac{\alpha}{\sigma} \phi_\lambda(z) \left[\Phi_\lambda(z)\right]^{\alpha-1},$$

where $\boldsymbol{\theta} = (\mu, \sigma, \lambda, \alpha)^\top$, $z = (y - \mu)/\sigma$ and $\phi_\lambda(\cdot)$ and $\Phi_\lambda(\cdot)$ denote the pdf and cdf of the (standard) skew normal model given by

$$\phi_\lambda(y) = 2\,\phi(y)\Phi(\lambda\,y) \qquad \text{and} \qquad \Phi_\lambda(y) = \int_{-\infty}^{y} \phi_\lambda(t)\mathrm{d}t = \Phi(y) - 2\,T(y, \lambda),$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ denote the pdf and cdf of the standard normal distribution and $T(\cdot, \cdot)$ is the Owen's T function defined as

$$T(y, \lambda) = \frac{1}{2\pi} \int_0^\lambda \frac{e^{-\frac{1}{2}y^2(1+t^2)}}{1+t^2} \, dt.$$

Moreover, the cdf of the PSN model is given by

$$F(y; \boldsymbol{\theta}) = [\Phi_\lambda(z)]^\alpha = [\Phi(z) - 2\,T(z, \lambda)]^\alpha.$$

Note that $\alpha = 1$ and $\lambda = 0$ corresponds to the very well known SN and PN models, respectively. The main advantage of the PSN model is that provides greater range for skewness and kurtosis coefficients compared with the SN and PN models. Table 1 shows the range for those coefficients.

The r-th moment of the distribution depends on the expected value of $[\Phi_\lambda^{-1}(Y)]^s$, $s = 1, \ldots, r$, where Y have beta distribution with shape parameters α and 1, respectively. For this reason, some interesting characteristic of the model, such as mean and variance, have cumbersome forms. On the other hand, quantiles of the model also need to be computed numerically since non-closed form are available for the distribution. For this reason, non-interpretation and useful reparametrizations can be performed for this model. Besides, as the Owen's T function satisfies $T(0, \lambda) = (2\pi)^{-1} \arctan(\lambda)$, we note that

$$F(\mu; \boldsymbol{\theta}) = \left[\frac{1}{2} - \frac{1}{\pi} \arctan(\lambda)\right]^\alpha.$$

For this reason, if we consider the restriction

$$\alpha = \alpha(\lambda, \tau) = \frac{\log(\tau)}{\log\left(\frac{1}{2} - \frac{1}{\pi} \arctan(\lambda)\right)}, \tag{1}$$

we have that $F(\mu; \boldsymbol{\theta}) = \tau$ with μ representing directly the τ-th quantile of the distribution. For a fixed τ and considering $\alpha(\lambda, \tau)$ as in (1), we have a flexible model for quantile regression. This parametrization has not been proposed in the statistical literature. Hence, we can rewrite the PSN distribution according to the parameters μ, σ and λ, whose cumulative distribution function is now given by

$$F(y; \mu, \sigma, \lambda) = [\Phi(z) - 2\,T(z, \lambda)]^{\frac{\log(\tau)}{\log\left(\frac{1}{2} - \frac{1}{\pi} \arctan(\lambda)\right)}},$$

where the quantile $\tau \in (0, 1)$ is assumed to be known. Hereafter, we use the notation $Y \sim \text{RPSN}(\mu, \sigma, \lambda)$ to indicate that Y is a random variable following a restricted PSN distribution with quantile parameter μ, precision parameter σ, and shape parameter λ. Figure 1 shows the density function for the RPSN model with location and scale parameters fixed at 0 and 1, respectively. Note that in all the curves, the zero represents the specified quantile τ. We also note that the curves are not necessarily symmetric for $\tau = 0.5$ (the median case).

Let $Y_1, \ldots, Y_n$ be the n independent random variables, where each $Y_i, i = 1, \ldots, n$, follows the PSN distribution with quantile parameter μ, precision parameter σ, and shape parameter λ. Suppose that, for a given $\tau \in (0, 1)$, the location, precision and shape parameters for the RPSN satisfy the following functional relations

$$\begin{aligned}
g_1(\mu_i(\tau)) &= \eta_{i1}(\tau) = \mathbf{x}_{i1}^\top \boldsymbol{\beta}_1(\tau), \\
g_2(\sigma_i(\tau)) &= \eta_{i2}(\tau) = \mathbf{x}_{i2}^\top \boldsymbol{\beta}_2(\tau) \quad \text{and} \\
g_3(\lambda_i(\tau)) &= \eta_{i3}(\tau) = \mathbf{x}_{i3}^\top \boldsymbol{\beta}_3(\tau),
\end{aligned} \tag{2}$$

where $\boldsymbol{\beta}_j(\tau) = (\beta_{j1}(\tau), \ldots, \beta_{jp_1}(\tau))^\top$, $j = 1, 2, 3$, are vectors of unknown regression coefficients which are assumed to be functionally independent, $\boldsymbol{\beta}_j(\tau) \in \mathbb{R}^{p_j}$, with $p_1 + p_2 + p_3 < n$, $\eta_{ji}(\tau)$ are the

linear predictors, and $\mathbf{x}_{ij} = (x_{ij1}, \ldots, x_{ijp_j})^\top$, are observations on p_1, p_2 and p_3 known regressors, for $i = 1, \ldots, n$. Moreover covariate matrices $\mathbf{X}_j = (\mathbf{x}_{1j}, \ldots, \mathbf{x}_{nj})^\top$ are assumed to have rank p_j, for $j = 1, 2, 3$. Link functions $g_1 : \mathbb{R} \to \mathbb{R}$, $g_2 : \mathbb{R} \to \mathbb{R}^+$ and $g_3 : \mathbb{R} \to \mathbb{R}$ in (2) must be strictly monotone and at least twice differentiable, and g_2 is also required to be a positive function. Such functions also satisfy that $\mu_i = g_1^{-1}(\mathbf{x}_{i1}^\top \boldsymbol{\beta}_1)$, $\sigma_i = g_2^{-1}(\mathbf{x}_{i2}^\top \boldsymbol{\beta}_2)$ and $\lambda_i = g_3^{-1}(\mathbf{x}_{i3}^\top \boldsymbol{\beta}_3)$, with $g_j^{-1}(\cdot)$ being the inverse function of $g_j(\cdot)$.
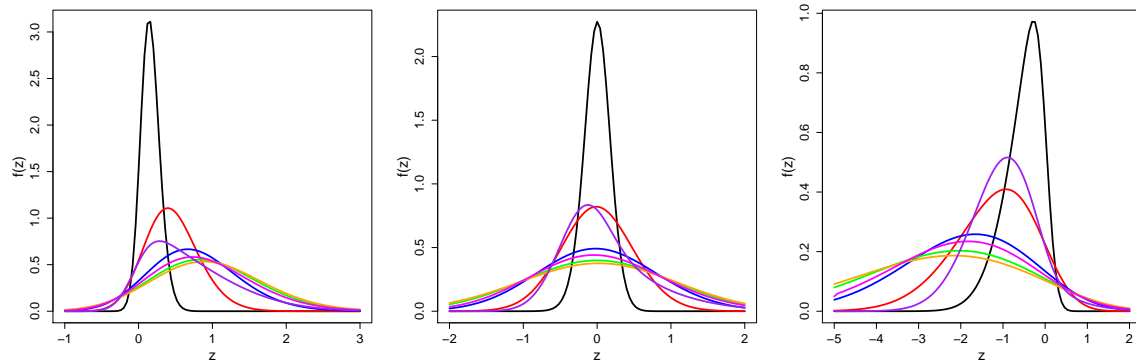

Figure 1. Pdf for the RPSN$(\mu = 0, \sigma = 1, \lambda)$ for different values of λ: $\tau = 0.1$ (left panel); $\tau = 0.5$ (center panel); $\tau = 0.9$ (right panel). Values for λ are: -5 (black line), -1.5 (red line), -0.5 (blue line), 0 (green line), 0.5 (orange line), 1.5 (magenta line) and 5 (purple line).

The log-likelihood function for $\boldsymbol{\theta} = \boldsymbol{\theta}(\tau) = (\boldsymbol{\beta}_1(\tau), \boldsymbol{\beta}_2(\tau), \boldsymbol{\beta}_3(\tau))$ has the form $\ell(\boldsymbol{\theta}) = \sum_{i=1}^n \ell_i$, where

$$\ell_i = \ell(z_i, \mu_i, \sigma_i, \lambda_i) = \log \alpha(\lambda_i, \tau) - \log(\sigma_i) + \log\left[\phi_{\lambda_i}(z_i)\right] + [\alpha(\lambda_i, \tau) - 1] \log\left[\Phi_{\lambda_i}(z_i)\right]. \quad (3)$$

The $(p_1 + p_2 + p_3) \times 1$ score vector of the model is given by

$$\dot{\ell}(\boldsymbol{\theta}) = \begin{pmatrix} \frac{\partial \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\beta}_1} \\ \frac{\partial \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\beta}_2} \\ \frac{\partial \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\beta}_3} \end{pmatrix} = \begin{pmatrix} \mathbf{X}_1^\top \mathbf{W}_{\boldsymbol{\beta}_1}^{1/2} \dot{\ell}_\mu \\ \mathbf{X}_2^\top \mathbf{W}_{\boldsymbol{\beta}_2}^{1/2} \dot{\ell}_\sigma \\ \mathbf{X}_3^\top \mathbf{W}_{\boldsymbol{\beta}_3}^{1/2} \dot{\ell}_\lambda \end{pmatrix}, \quad (4)$$

where $\mathbf{W}_{\boldsymbol{\beta}_j} = \text{diag}(w_{\beta_{j1}}, \ldots, w_{\beta_{jn}})$, $w_{\beta_{1i}} = (\partial \mu_i / \partial \eta_{1i})^2$, $w_{\beta_{2i}} = (\partial \sigma_i / \partial \eta_{2i})^2$, $w_{\beta_{3i}} = (\partial \lambda_i / \partial \eta_{3i})^2$, $\dot{\ell}_\xi = (\dot{\ell}_{\xi_1}, \ldots, \dot{\ell}_{\xi_n})$, for $\xi \in \{\mu, \sigma, \lambda\}$, with $\dot{\ell}_{\xi_i} = \partial \ell(\mu_i, \sigma_i, \lambda_i) / \partial \xi_i$. Such elements are specified in the Appendix A.1 Section.

The Hessian for the model is

$$\mathbf{H}(\boldsymbol{\theta}) = \begin{pmatrix} \mathbf{H}_{\boldsymbol{\beta}_1 \boldsymbol{\beta}_1} & \mathbf{H}_{\boldsymbol{\beta}_1 \boldsymbol{\beta}_2} & \mathbf{H}_{\boldsymbol{\beta}_1 \boldsymbol{\beta}_3} \\ \cdot & \mathbf{H}_{\boldsymbol{\beta}_2 \boldsymbol{\beta}_2} & \mathbf{H}_{\boldsymbol{\beta}_2 \boldsymbol{\beta}_3} \\ \cdot & \cdot & \mathbf{H}_{\boldsymbol{\beta}_3 \boldsymbol{\beta}_3} \end{pmatrix}$$

$$= \begin{pmatrix} \mathbf{X}_1^\top \ddot{\ell}_{\mu\mu} \mathbf{W}_{\boldsymbol{\beta}_1} \mathbf{X}_1 & \mathbf{X}_1^\top \ddot{\ell}_{\mu\sigma} \mathbf{W}_{\boldsymbol{\beta}_1}^{1/2} \mathbf{W}_{\boldsymbol{\beta}_2}^{1/2} \mathbf{X}_2 & \mathbf{X}_1^\top \ddot{\ell}_{\mu\lambda} \mathbf{W}_{\boldsymbol{\beta}_1}^{1/2} \mathbf{W}_{\boldsymbol{\beta}_3}^{1/2} \mathbf{X}_3 \\ \cdot & \mathbf{X}_2^\top \ddot{\ell}_{\sigma\sigma} \mathbf{W}_{\boldsymbol{\beta}_2} \mathbf{X}_2 & \mathbf{X}_2^\top \ddot{\ell}_{\sigma\lambda} \mathbf{W}_{\boldsymbol{\beta}_2}^{1/2} \mathbf{W}_{\boldsymbol{\beta}_3}^{1/2} \mathbf{X}_3 \\ \cdot & \cdot & \mathbf{X}_3^\top \ddot{\ell}_{\lambda\lambda} \mathbf{W}_{\boldsymbol{\beta}_3} \mathbf{X}_3 \end{pmatrix}, \quad (5)$$

where $\ddot{\ell}_{\xi\xi'} = \text{diag}(\ddot{\ell}_{\xi_1\xi_1'}, \ldots, \ddot{\ell}_{\xi_n\xi_n'})$, for $\xi, \xi' \in \{\mu, \sigma, \lambda\}$, with $\ddot{\ell}_{\xi_i\xi_i'} = \partial^2 \ell_i / \partial \xi_i \partial \xi_i'$. Such elements are detailed in the Appendix A.1.

The ML estimators $\widehat{\boldsymbol{\beta}}_1(\tau)$, $\widehat{\boldsymbol{\beta}}_2(\tau)$ and $\widehat{\boldsymbol{\beta}}_3(\tau)$ of $\boldsymbol{\beta}_1(\tau)$, $\boldsymbol{\beta}_2(\tau)$ and $\boldsymbol{\beta}_3(\tau)$, respectively, can be obtained by solving simultaneously the nonlinear system of equations $\dot{\ell}(\boldsymbol{\theta}) = \mathbf{0}_{p_1+p_2+p_3}$, where $\mathbf{0}_r$ denotes a vector of zeros with dimension r. Unfortunately, it is not possible to obtain analytical

expressions for the ML estimators above, so numerical methods for solving nonlinear equations system are required.

3. Local Influence

Global influence is related to case deletion, i.e, the effect of dropping a case from the dataset Cook [14]. The likelihood distance (LD) is defined as $\text{LD}(\omega) = 2[\ell(\widehat{\theta}) - \ell(\widehat{\theta}(\omega), \omega)]$, where $\widehat{\theta}(\omega)$ is the ML estimate of θ under a perturbed model related to $\omega = (\omega_1, \ldots, \omega_n)^\top$, a perturbation vector. Cook [14] studied the $\text{LD}(\omega)$ around the non-perturbed vector ω_0 such as $\widehat{\theta}(\omega_0) = \widehat{\theta}$. The normal curvature for $\widehat{\omega}$ at the direction of the orthonormal vector $||d||$ is defined as $C_d(\widehat{\theta}) = 2|d^\top \Delta_\omega^\top \ddot{\ell}_{\widehat{\theta}\widehat{\theta}} \Delta_\omega d|$, where $\ddot{\ell}_{\widehat{\theta}\widehat{\theta}}$ is the Hessian of $\ell(\theta)$ evaluated at $\theta = \widehat{\theta}$ and $\Delta_\omega = \partial^2 \ell(\theta, \omega) / \partial\theta\partial\omega^\top |_{\theta=\widehat{\theta}(\omega)}$ and both, Δ_ω and $\ddot{\ell}_{\widehat{\theta}\widehat{\theta}}$ are evaluated at $\widehat{\theta}(\omega)$. Hence, $C_{d_{\max}}$ is the largest eigenvalue of $\mathbf{B} = \Delta_{\omega_0}^\top \ddot{\ell}_{\widehat{\theta}\widehat{\theta}} \Delta_{\omega_0}$ and $d_{\max}$ the corresponding orthonormal eigenvector. The index $d_{\max}$ plot of the matrix $\mathbf{B}$ suggests how to perturb the model (or data) to obtain large changes in the estimates of θ.

For three common perturbation schemes we compute the matrix

$$\Delta_\omega = \frac{\partial^2 \ell(\theta, \omega)}{\partial\theta\partial\omega} = \left(\begin{array}{ccc} \Delta_{\omega,\beta_1}^\top & \Delta_{\omega,\beta_2}^\top & \Delta_{\omega,\beta_3}^\top \end{array} \right)^\top,$$

where $\Delta_{\omega,\beta_j}^\top = \frac{\partial^2 \ell(\theta, \omega)}{\partial\beta_1 \partial\omega}$.

3.1. Case Weights Perturbation

For this case, the perturbed log-likelihood function is defined as $\ell(\theta, \omega) = \sum_{i=1}^n \omega_i \ell_i$, where $\ell(z_i, \mu_i, \sigma_i, \lambda_i)$ is defined in (3) and $0 \leq \omega_i \leq 1$, for $i = 1, \ldots, n$. In this case, $\omega_0 = (1, \ldots, 1)$ and

$$\Delta_\omega = \left(\begin{array}{ccc} \mathbf{X}_1^\top \mathbf{W}_{\beta_1}^{1/2} \ell_\mu^\top & \mathbf{X}_2^\top \mathbf{W}_{\beta_2}^{1/2} \ell_\sigma^\top & \mathbf{X}_3^\top \mathbf{W}_{\beta_3}^{1/2} \ell_\lambda^\top \end{array} \right)^\top.$$

3.2. Case Response Perturbation

We consider now an additive perturbation on the ith response (say $y_i(\cdot)$) by making $y_i(\omega_i) = y_i + \omega_i S_{Y_i}$, where $\omega_i \in \mathbb{R}$ and S_{Y_i} is a scale factor. An usual consideration for such scale factor is $S_{Y_i} = S_Y$, with S_Y denoting the sample standard deviation of Y. Note that $\omega_0 = (0, \ldots, 0)$. Therefore, under the scheme of response perturbation, the log-likelihood function is given by $\ell(\theta, \omega) = \sum_{i=1}^n \ell(z_i(\omega_i), \mu_i, \sigma_i, \lambda_i)$, where $z_i(\omega_i) = (y_i(\omega_i) - \mu_i)/\sigma_i$ and

$$\Delta_\omega = S_Y \left(\begin{array}{ccc} \mathbf{X}_1^\top \mathbf{W}_{\beta_1}^{1/2} \ddot{\ell}_{\mu\mu}^\top & \mathbf{X}_2^\top \mathbf{W}_{\beta_2}^{1/2} \ddot{\ell}_{\mu\sigma}^\top & \mathbf{X}_3^\top \mathbf{W}_{\beta_3}^{1/2} \ddot{\ell}_{\mu\lambda}^\top \end{array} \right)^\top \Bigg|_{z_i = z_i(\omega_i)}.$$

3.3. Case Continuous Covariate Perturbation

Consider an additive perturbation on a particular continuous covariate including on the quantile parameter, namely x_t, for $t \in \{1, \ldots, p_1\}$, by making $x_{it}(\omega_i) = x_{it} + \omega_i S_{X_{it}}$, where $S_{X_{it}}$ is a scale factor. Again, a usual consideration is $S_{X_{it}} = S_{X_t}$, with S_{X_t} the sample standard deviation for X_t. Note that $\omega_0 = (0, \ldots, 0)$. Then, under the scheme of response perturbation, the log-likelihood function is given by $\ell(\theta, \omega) = \sum_{i=1}^n \ell(z_i, \mu_i(\omega_i), \sigma_i, \lambda_i)$, where $\mu_i(\omega_i) = g_1^{-1}(\mathbf{x}_{i1}^\top(\omega_i)\beta_1)$ and $\mathbf{x}_{i1}^\top(\omega_i) = \mathbf{x}_{i1}^\top + \omega_i S_{X_{it}} \mathcal{J}_t$, with $\mathcal{J}_t$ a vector of dimension p_1 with zeros, except in the t-th element where is a one. Finally

$$\Delta_\omega = S_{X_t} \left(\begin{array}{ccc} \text{diag}(\mathcal{J}_t)\mathbf{X}_1^\top \mathbf{W}_{\beta_1}^{1/2} \ddot{\ell}_{\mu\mu}^\top & \text{diag}(\mathcal{J}_t)\mathbf{X}_2^\top \mathbf{W}_{\beta_2}^{1/2} \ddot{\ell}_{\mu\sigma}^\top & \text{diag}(\mathcal{J}_t)\mathbf{X}_3^\top \mathbf{W}_{\beta_3}^{1/2} \ddot{\ell}_{\mu\lambda}^\top \end{array} \right)^\top \Bigg|_{z_i = z_i(\omega_i)}.$$

4. Real Data Analysis

In this section, we present an application to 202 Australian athletes from the Australian Institute of Sport. Such data were discussed in Cook and Weisberg [15]. In order to exemplify the proposed model, we consider the following quantile regression model: $\mathtt{bmi}_i(\tau) = \mu_i(\tau) + \sigma_i(\tau)\epsilon_i(\tau)$, where $\epsilon_i(\tau) \sim \mathrm{PSN}(0, 1, \lambda, \tau)$ and

$$\mu_i(\tau) = \beta_{10}(\tau) + \beta_{11}(\tau)\mathtt{lbm}_i + \beta_{12}(\tau)\mathtt{sex}_i$$
$$\sigma_i(\tau) = \beta_{20}(\tau) + \beta_{21}(\tau)\mathtt{lbm}_i.$$

Here, the response $\mathtt{bmi}$ represents the body mass index, while the covariates $\mathtt{lbm}$ and $\mathtt{sex}$ represent the lean body mass and sex of the athletes, respectively. Note that λ is not modeled by covariates and $\mathtt{sex}$ was not included in the scale parameter because in preliminary analysis we found the coefficient related to such term was not significant (to any $\tau \in (0, 1)$). This same problem was illustrated in Galarza et al. [16] with a class of skew distributions (SKD), but considering a regression scheme only in the quantile parameter. For comparison purpose, we considered the skewed normal (SKN) and skewed Student-t (SKT) models, that are models belonging to the SKD class. Additionally, we also considered the Gamma-Sinh Cauchy (GSC) model, including covariates only in the quantile parameter. Table 2 shows the Akaike Information Criterion (AIC, Akaike; [17]) for the referred models. Note that, except for $\tau = 0.25$, the RPSN-QR model attached the minimum AIC for the considered quantiles.

Table 2. AIC criterion for different models parameterized in terms of the quantile.

τ	SKN	SKT	GSC	RPSN (σ Constant)	RPSN (Modeling σ)
0.10	1097.74	817.77	803.08	808.64	801.37
0.25	1084.46	803.90	801.96	811.08	803.11
0.50	1095.56	810.99	854.38	815.79	806.76
0.75	1151.40	854.57	861.37	824.56	814.01
0.90	1220.96	914.43	865.16	838.78	825.95

Tables 2 displays the MLEs with corresponding standard errors (SE) for the fitted proposed model for each $\tau = 0.10, 0.50$ and 0.90. Note that we have a positive relationship between the response variable ($\mathtt{bmi}$) and $\mathtt{lbm}$ in all quantiles. We also observe that the quantile intercepts increases as τ increases. Regarding the parameter λ, the greater τ, the greater the estimate of λ.

Table 3. Estimates and SE for parameters in athletes dataset in RPSN-quantile regression (QR) model for different values of τ.

| Parameter | $\tau = 0.10$ | | | $\tau = 0.50$ | | | $\tau = 0.90$ | | |
	Est.	SE	p-Value	Est.	SE	p-Value	Est.	SE	p-Value
$\beta_{10}(\tau)$	6.4642	1.1552	-	6.7727	1.0867	-	6.1798	1.0859	-
$\beta_{11}(\tau)$	2.3077	0.3742	<0.0001	2.5008	0.3728	<0.0001	2.9324	0.3695	<0.0001
$\beta_{12}(\tau)$	0.2037	0.0157	<0.0001	0.2299	0.0147	<0.0001	0.2728	0.0151	<0.0001
$\beta_{20}(\tau)$	0.7633	0.7952	-	0.2252	0.3996	-	−0.6261	0.2700	-
$\beta_{21}(\tau)$	0.0096	0.0036	0.0040	0.0108	0.0037	0.0017	0.0125	0.0035	0.0002
$\beta_{30}(\tau)$	−1.1940	1.4984	-	−0.8381	0.5984	-	−0.4916	0.2588	-

Figure 2 shows point estimates and 95% confidence intervals (CIs) for model parameters under the RPSN-QR model for different quantiles. It can be seen that as τ increases the coefficient of lean body mass and the coefficient of gender become larger. Moreover, $\mathtt{bmi}$ and $\mathtt{lbm}$ are significant in explaining all the quantile modeled in μ_i. Figure 3 presented the estimated quantiles $0.10, 0.25, 0.50, 0.75$ and 0.90 for the $\mathtt{bmi}$ in terms of $\mathtt{lbm}$ and the $\mathtt{sex}$ of the athlete.
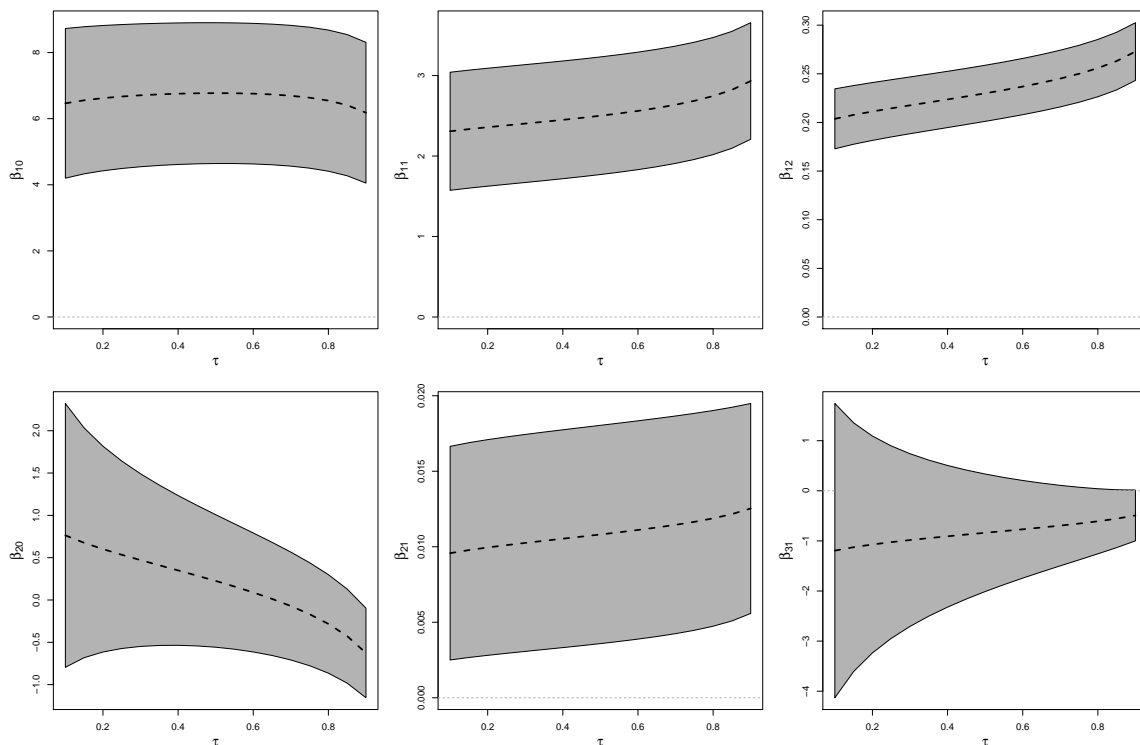
Figure 2. Athletes dataset: Point estimates (center line) and 95% confidence intervals (CIs) for model parameters under RPSN-QR model.
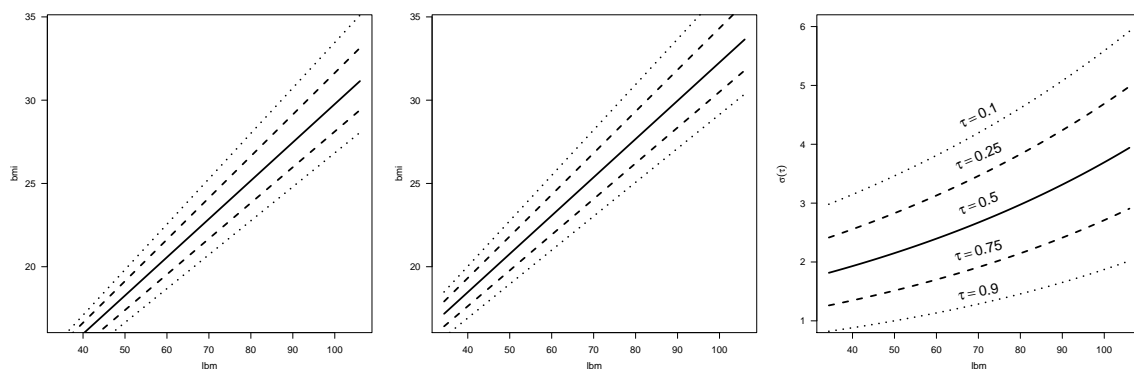

Figure 3. Data analysis: Fitted RPSN-QR model lines for the response (left panel for males, center panel for females) and scale parameter (right panel) over the grid $\tau = \{0.10, 0.25, 0.50, 0.75, 0.90\}$.

We also present in Table 4 the *p*-value to validate the normality hypothesis based on the Kolmogorov–Smirnov (KS; Kolmogorov, [18]) for the quantile residuals (Dunn and Smyth, [19]) using different quantile τ of such residuals. In all cases, the KS test did not reject the null hypothesis of normality. Therefore, the RPSN is appropriated to model all the quantile in this problem.

We also performed a local influence analysis. Figure 4 shows such analysis under the three perturbation schemes discussed in Section 3 for $\tau = 0.5$. The Appendix A.2 shows the analysis for other quantiles $\tau = 0.1, 0.25, 0.75$ and 0.9. Note that observations 75, 162 and 178 are detected as potentially influent for all the mentioned quantiles and the observation 53 appears for the quantile 0.9.

To check the impact on the inference of possible influential cases, we consider the relative change (RC), which is computed by removing the possible influential cases for each parameter and its SE as

$$RC_{\theta_{j(i)}} = 100\% \times \left| \frac{\widehat{\theta}_j - \widehat{\theta}_{j(i)}}{\widehat{\theta}_j} \right| \quad \text{and} \quad RC_{SE(\theta_{j(i)})} = 100\% \times \left| \frac{SE\left(\widehat{\theta}_j\right) - SE\left(\widehat{\theta}_{j(i)}\right)}{SE\left(\widehat{\theta}_j\right)} \right|,$$

where θ_j is any component of the vector $\boldsymbol{\theta} = \boldsymbol{\theta}(\tau)$, where $\widehat{\theta}_{j(i)}$ and $SE\left(\widehat{\theta}_{j(i)}\right)$ denote the ML estimate of θ_j and its corresponding SE, respectively, after dropping the i-th observation. Table 5 shows such RC for the non-intercept regression coefficients when observations 53, 75, 162 and 178 are removed. Note that the RC is greater for the estimated parameters than its estimated SE. However, the significance of $\beta_{11}(\tau)$ and $\beta_{12}(\tau)$ is maintained whereas $\beta_{21}(\tau)$ is not significant with a 5%. More combinations of dropped observations are presented in the Appendix A.2.

Table 4. *p*-values for normality K-S test for residuals under our RPSN-QR model for the athletes dataset for different quantilesτ's.

τ	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
p-value	0.995	0.996	0.991	0.976	0.951	0.914	0.864	0.853	0.837

τ	0.55	0.60	0.65	0.70	0.75	0.80	0.85	0.90	
p-value	0.765	0.810	0.777	0.683	0.604	0.524	0.394	0.191	

Table 5. Relative changes (RC) (in %) in ML estimates and their corresponding SE's for the indicated parameter and respective *p*-values for the athletes dataset when observations 53, 75, 162 and 178 are dropped.

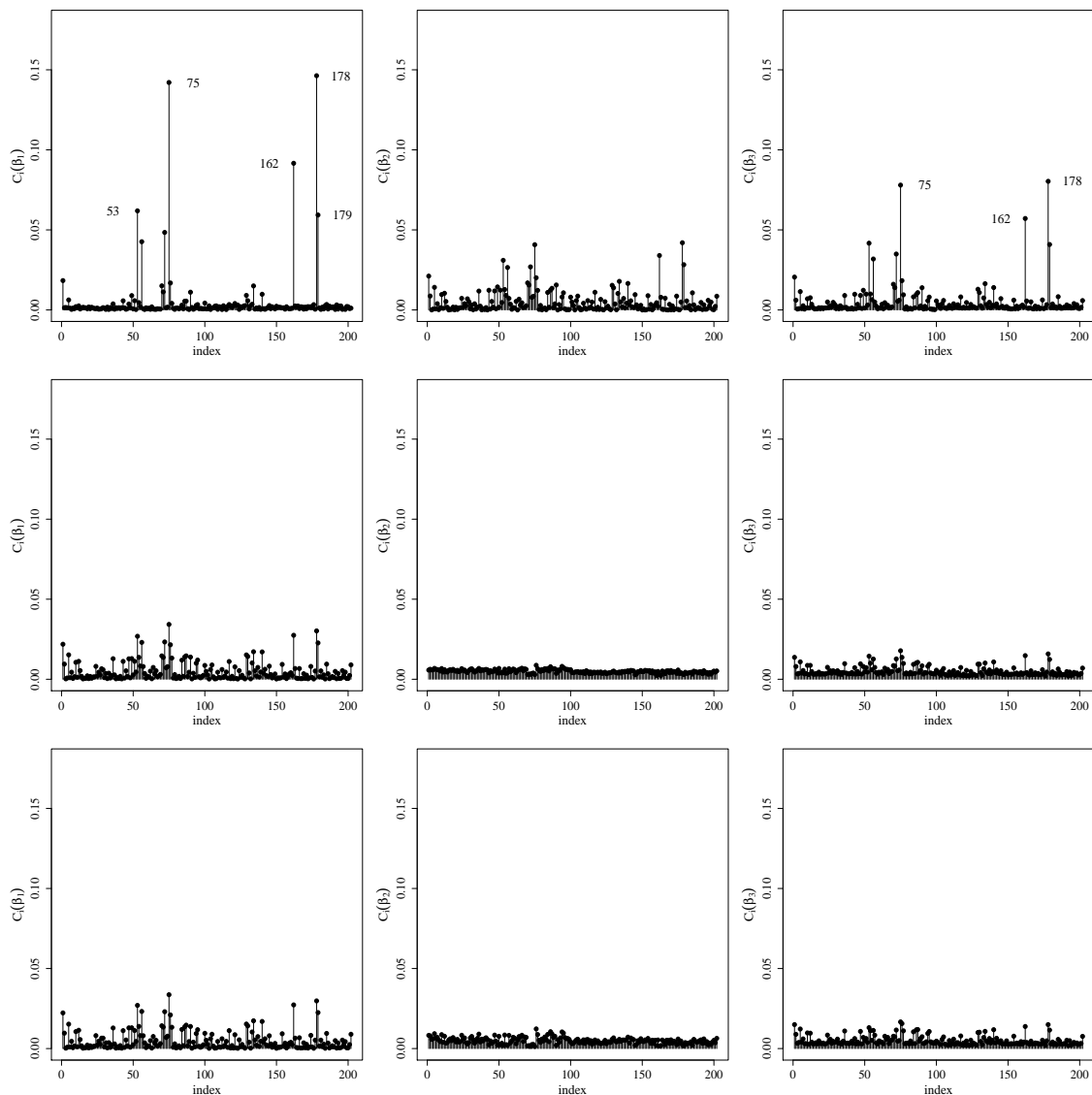| Dropped Cases		Parameter			τ		
			0.10	0.25	0.50	0.75	0.90
53, 75,	RC		7.20	9.06	12.56	17.82	24.05
162 and 178	RCSE	$\beta_{11}(\tau)$	0.74	0.74	0.69	0.63	0.52
	p-value		<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
	RC		1.98	5.63	10.93	17.22	23.38
	RCSE	$\beta_{12}(\tau)$	0.47	0.26	0.20	0.34	0.42
	p-value		<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
	RC		34.55	36.83	40.05	43.61	46.96
	RCSE	$\beta_{21}(\tau)$	0.23	0.53	0.69	0.36	0.66
	p-value		0.0806	0.0762	0.0727	0.0697	0.0670

Figure 4. Index plots for $C_i(\widehat{\beta}_1)$ (left), $C_i(\widehat{\beta}_2)$ (center) and $C_i(\widehat{\beta}_3)$ (right) under the weight perturbation (upper), response perturbation (center) and covariate perturbation (lower) schemes for RPSN model for $\tau = 0.5$.

5. Concluding Remarks

Extending the quantile regression methods to include asymmetric response variables on the real line is promising area of research. In this paper, we have introduced a novel flexible parametric quantile regression model for asymmetric response variables, which can be very useful in modeling response variables on the real line at different quantiles. The proposed quantile regression model was built based on PSN distribution using a new parameterization of this distribution that is indexed by quantile, precision and shape parameters, in which a function of any quantile of the response variable is given by a linear predictor that is defined by regression parameters and explanatory variables. We consider a frequentist approach to estimate the model parameters, and the maximum likelihood inference is employed to estimate the model parameters. An application using a real dataset was presented and discussed. Results of the application showed that the model is adequate; it elaborately showed which covariates influence the response at different levels of quantiles. Finally, there are many possible extensions of the current work, for instance, mixtures of RPSN regression models in order to accommodate multimodality, a semi-parametric component to include a functional covariate to model

nonlinearity of the response, and measurement errors, among others. An in-depth investigation of these topics is beyond the scope of this work, and will be considered elsewhere.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Appendix A.1. Details for Score and Hessian

For the score vector in Equation (4), the elements of the form $\frac{\partial \ell_i}{\partial \xi_i}$, with $\xi \in \{\mu, \sigma, \lambda\}$ are given by

$$\frac{\partial \ell_i}{\partial \mu_i} = -\frac{1}{\sigma_i} \left\{ \lambda_i m_0(\lambda_i z_i) - z_i + [\alpha(\lambda_i, \tau) - 1] m_{\lambda_i}(z_i) \right\},$$

$$\frac{\partial \ell_i}{\partial \sigma_i} = -\frac{z_i}{\sigma_i} \left\{ \lambda_i m_0(\lambda_i z_i) - z_i + [\alpha(\lambda_i, \tau) - 1] m_{\lambda_i}(z_i) \right\} - \frac{1}{\sigma_i},$$

$$\frac{\partial \ell_i}{\partial \lambda_i} = \log(\tau) \left[\alpha(\lambda_i, \tau)(\lambda_i^2 + 1) \left(\frac{\pi}{2} - \arctan(\lambda_i) \right) \right]^{-1} \left(1 + \log \Phi_{\lambda_i}(z_i) \right)$$

$$+ z_i m_0(\lambda_i z_i) - [\alpha(\lambda_i, \tau) - 1] \frac{m_0(\lambda_i z_i)}{(1 + \lambda_i^2)},$$

where $m_\lambda(z) = \phi_\lambda(z) / \Phi_\lambda(z)$.

For the Hessian in Equation (5), the elements of the form $\partial^2 \ell_i / \partial \xi_i \partial \xi_i'$, with $\xi, \xi' \in \{\mu, \sigma, \lambda\}$ are given by

$$\frac{\partial^2 \ell_i}{\partial \mu_i^2} = \frac{1}{\sigma_i^2} \left\{ \lambda_i^2 m_0'(\lambda_i z_i) - 1 + [\alpha(\lambda_i, \tau) - 1] m_{\lambda_i}'(z_i) \right\}$$

$$\frac{\partial^2 \ell_i}{\partial \mu_i \partial \sigma_i} = \frac{z_i}{\sigma_i} \left\{ \lambda_i^2 m_0'(\lambda_i z_i) - 1 + [\alpha(\lambda_i, \tau) - 1] m_{\lambda_i}'(z_i) \right\}$$

$$\frac{\partial^2 \ell_i}{\partial \mu_i \partial \lambda_i} = -\frac{1}{\sigma_i} \left\{ \frac{\log(\tau) m_\lambda(z_i)}{(1 + \lambda_i^2) \left(\frac{\pi}{2} - \arctan(\lambda_i) \right) \left(1 + \log \Phi_{\lambda_i}(z_i) \right)} + m(\lambda_i z_i) + \lambda_i z_i m_0'(\lambda_i z_i) \right.$$

$$\left. - \frac{[\alpha(\lambda_i, \tau) - 1] \lambda_i m_0'(\lambda_i z_i)}{(1 + \lambda_i^2)} \right\}$$

$$\frac{\partial^2 \ell_i}{\partial \sigma_i^2} = \frac{z_i}{\sigma_i^2} \left\{ \lambda_i^2 m_0'(\lambda_i z_i) - 1 + [\alpha(\lambda_i, \tau) - 1] m_{\lambda_i}'(z_i) \right\} + \frac{1}{\sigma_i^2}$$

$$\frac{\partial^2 \ell_i}{\partial \sigma_i \partial \lambda_i} = -\frac{z_i}{\sigma_i} \left\{ \frac{\log(\tau) m_\lambda(z_i)}{(1 + \lambda_i^2) \left(\frac{\pi}{2} - \arctan(\lambda_i) \right) \left(1 + \log \Phi_{\lambda_i}(z_i) \right)} + m(\lambda_i z_i) + \lambda_i z_i m_0'(\lambda_i z_i) \right.$$

$$\left. - \frac{[\alpha(\lambda_i, \tau) - 1] \lambda_i m_0'(\lambda_i z_i)}{(1 + \lambda_i^2)} \right\}$$

$$\frac{\partial^2 \ell_i}{\partial \lambda_i^2} = \frac{\log^2(\tau) \left[(1 - \pi \lambda + 2\lambda_i \arctan(\lambda_i)) \log \left(\frac{1}{2} - \frac{1}{\pi} \arctan(\lambda_i) \right) + 1 \right] \left(1 + \log \Phi_{\lambda_i}(z_i) \right)}{\left[\alpha(\lambda_i, \tau)(1 + \lambda_i^2) \left(\frac{\pi}{2} - \arctan(\lambda_i) \right) \right]^2}$$

$$- \frac{\log(\tau) m_0(\lambda_i z_i) m_{\lambda_i}(z_i)}{\alpha(\lambda_i, \tau)(1 + \lambda_i^2)^2 \left(\frac{\pi}{2} - \arctan(\lambda_i) \right) \left(1 + \log \Phi_{\lambda_i}(z_i) \right)} + z_i^2 m_0'(\lambda_i z_i)$$

$$- \frac{m(\lambda_i z_i)}{\log(\tau) \alpha^2(\lambda_i, \tau)(1 + \lambda_i^2)^2 \left(\frac{\pi}{2} - \arctan(\lambda_i) \right)}$$

$$- \frac{[\alpha(\lambda_i, \tau) - 1]}{(1 + \lambda_i^2)} \left\{ z_i m'(\lambda_i z_i) - 2 \frac{\lambda_i m(\lambda_i z_i)}{(1 + \lambda_i^2)} \right\},$$

where $m_\lambda'(z) = \lambda m_0(\lambda z) m_\lambda(z) - z m_\lambda(z) - m_\lambda^2(z)$.

Appendix A.2. Local Influence

In this section, we present additional information for the local influence analysis in the Athletes dataset discussed in Section 5.
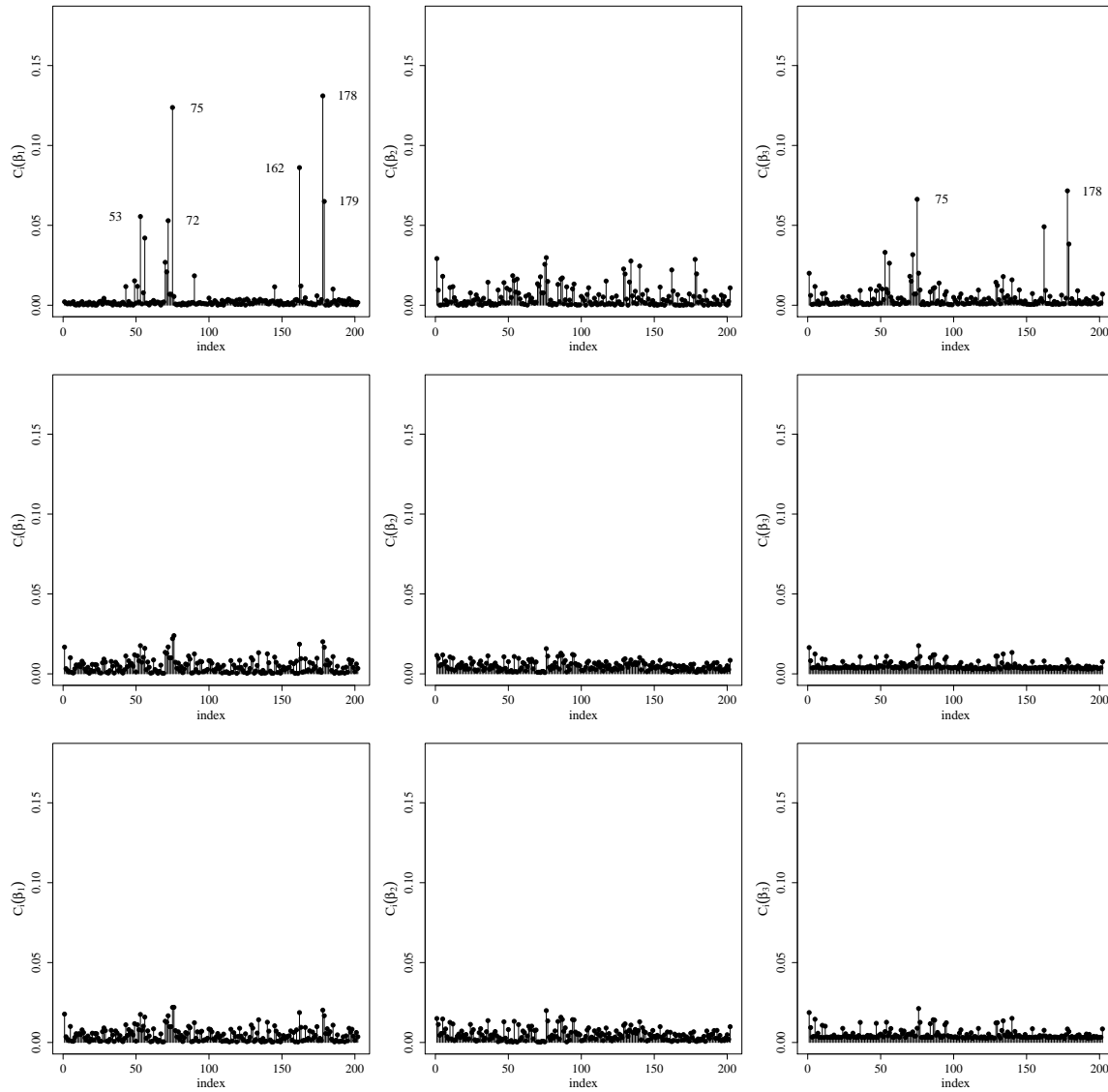

Figure A1. Index plots for $C_i(\widehat{\beta}_1)$ (left), $C_i(\widehat{\beta}_2)$ (center) and $C_i(\widehat{\beta}_3)$ (right) under the weight perturbation (upper), response perturbation (center) and covariate perturbation (lower) schemes for RPSN model for $\tau = 0.1$.
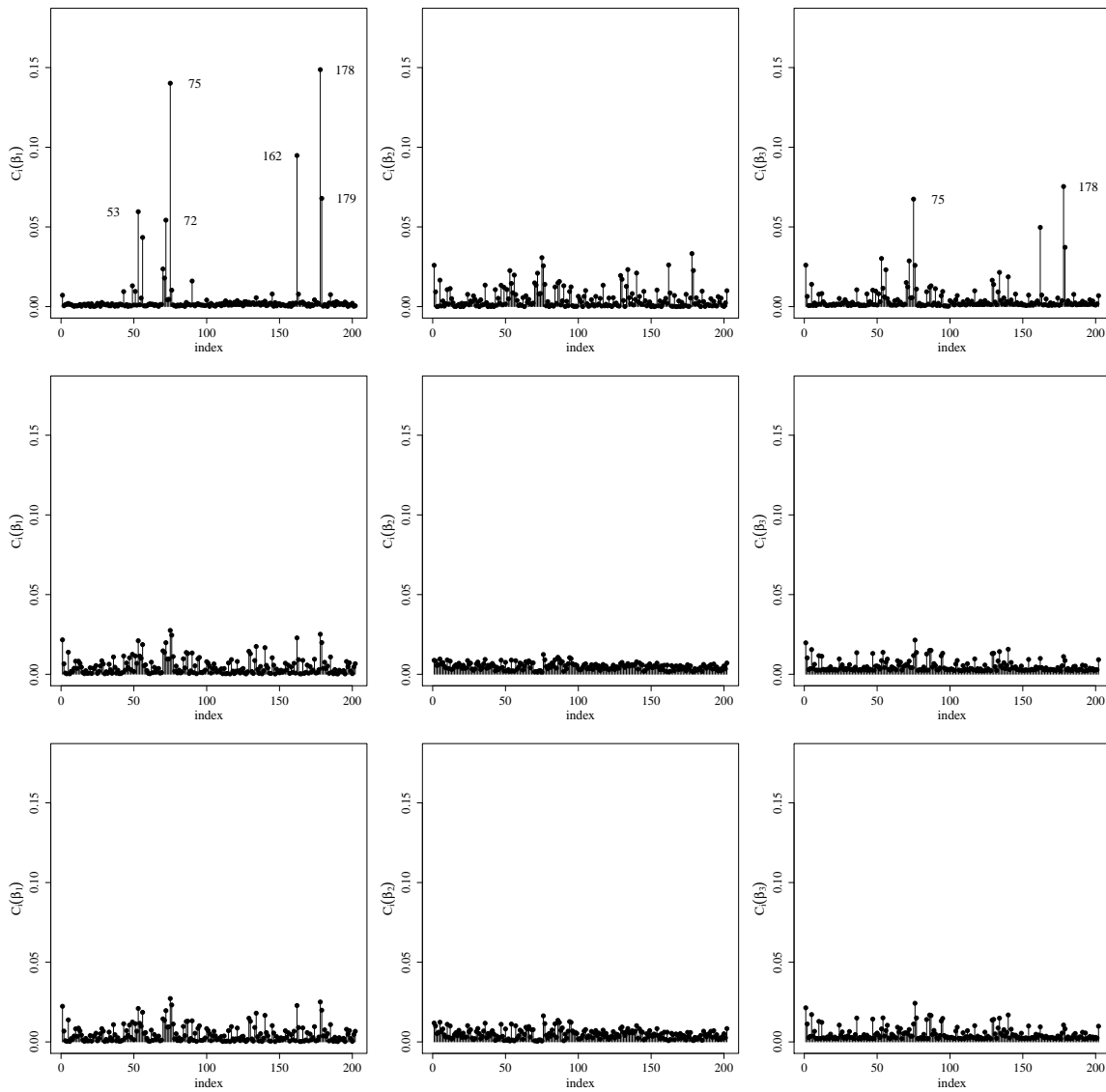
Figure A2. Index plots for $C_i(\widehat{\boldsymbol{\beta}}_1)$ (left), $C_i(\widehat{\boldsymbol{\beta}}_2)$ (center) and $C_i(\widehat{\boldsymbol{\beta}}_3)$ (right) under the weight perturbation (upper), response perturbation (center) and covariate perturbation (lower) schemes for RPSN model for $\tau = 0.25$.
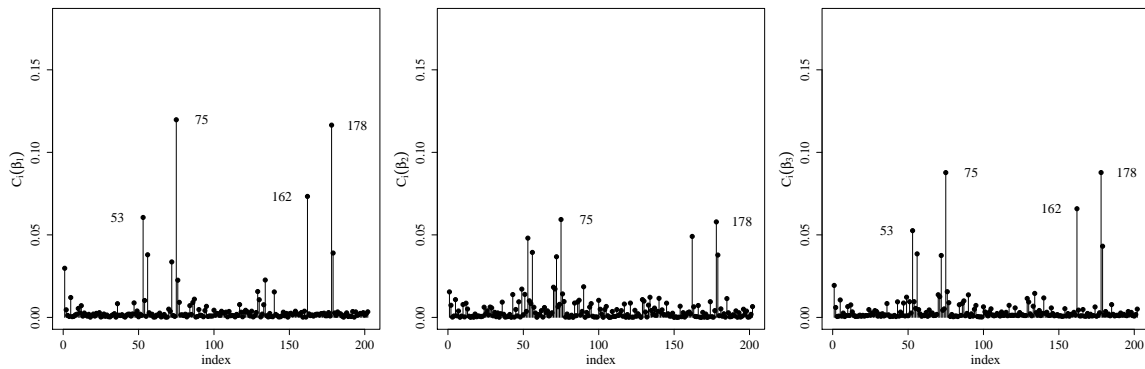

Figure A3. *Cont.*

Figure A3. Index plots for $C_i(\widehat{\beta}_1)$ (left), $C_i(\widehat{\beta}_2)$ (center) and $C_i(\widehat{\beta}_3)$ (right) under the weight perturbation (upper), response perturbation (center) and covariate perturbation (lower) schemes for RPSN model for $\tau = 0.75$.
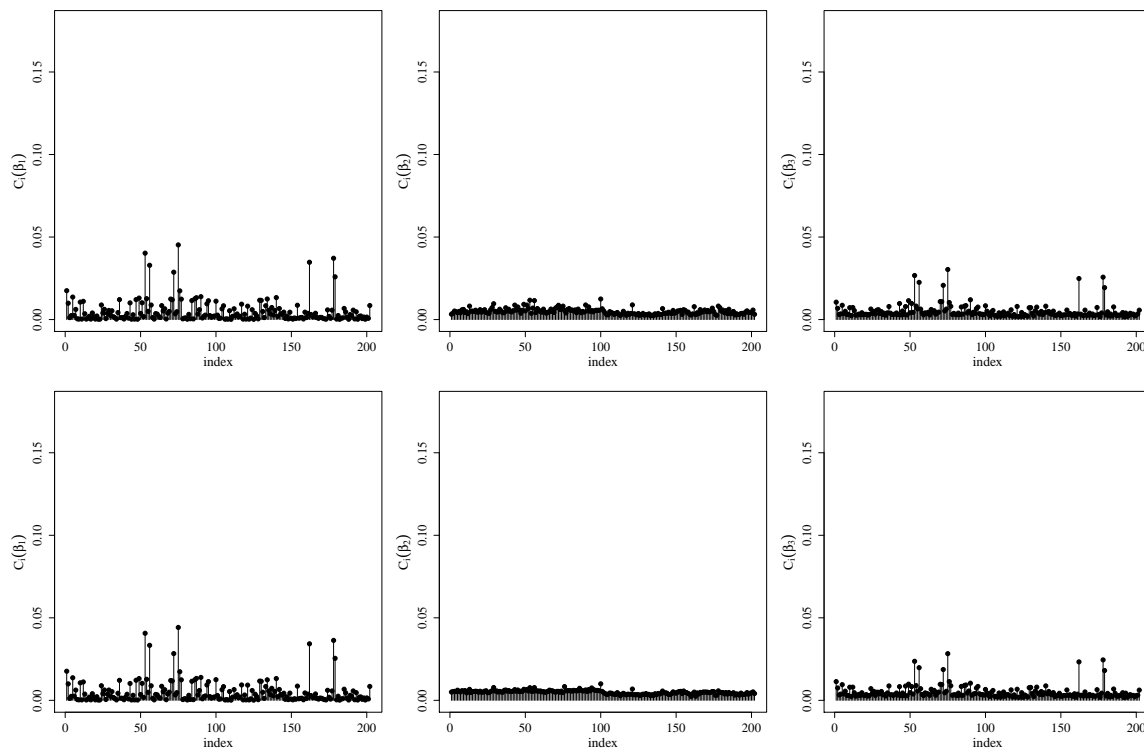

Figure A4. *Cont.*

Figure A4. Index plots for $C_i(\widehat{\boldsymbol{\beta}}_1)$ (left), $C_i(\widehat{\boldsymbol{\beta}}_2)$ (center) and $C_i(\widehat{\boldsymbol{\beta}}_3)$ (right) under the weight perturbation (upper), response perturbation (center) and covariate perturbation (lower) schemes for RPSN model for $\tau = 0.9$.
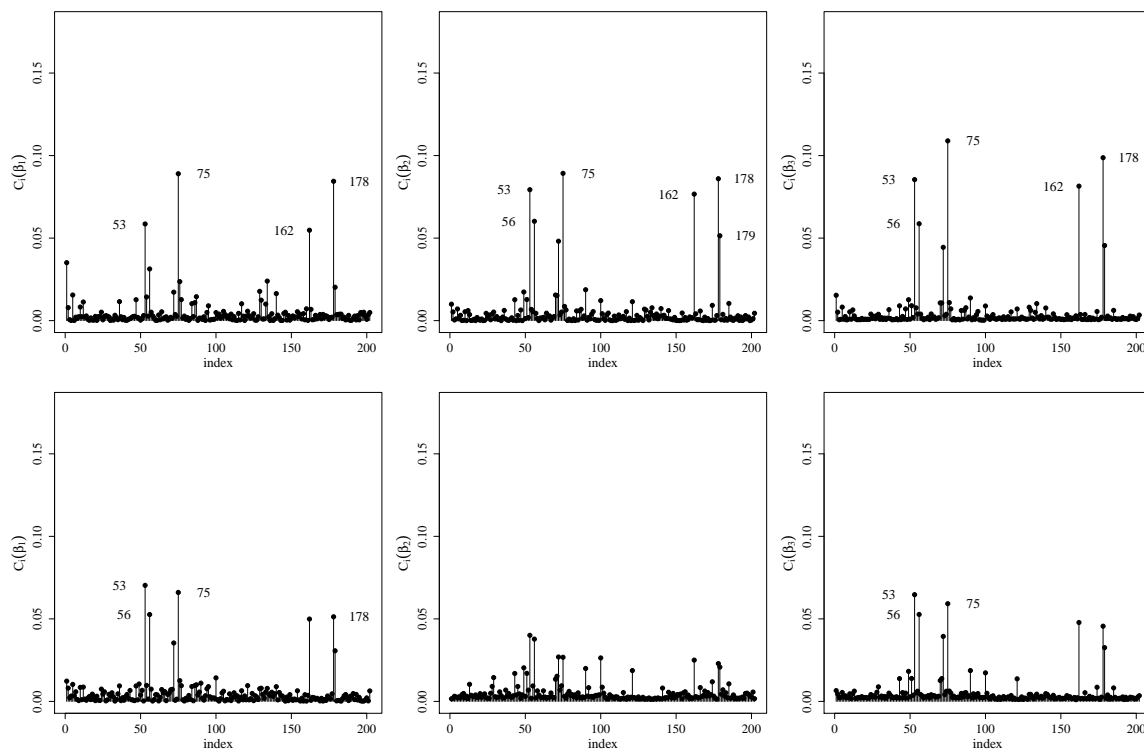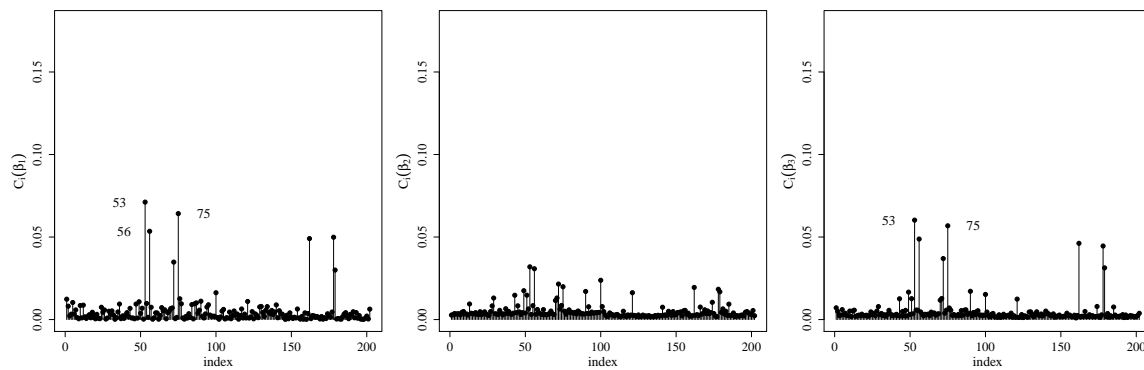
Table A1. RCs (in %) in ML estimates and their corresponding SEs for the indicated parameter and respective *p*-values for the athletes dataset when observation 75 and 178 are dropped separately.

Dropped			τ				
Cases		Parameter	0.10	0.25	0.50	0.75	0.90
75	RC		5.31	7.22	10.82	16.2	22.57
	RCSE	$\beta_{11}(\tau)$	0.23	0.20	0.17	0.11	0.04
	p-value		<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
	RC		1.82	5.03	10.02	16.09	22.08
	RCSE	$\beta_{12}(\tau)$	0.15	0.05	0.08	0.07	0.17
	p-value		<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
	RC		6.77	9.84	14.27	19.20	23.84
	RCSE	$\beta_{21}(\tau)$	0.65	0.93	1.05	0.71	0.33
	p-value		0.0118	0.0105	0.0095	0.0086	0.0078
178	RC		0.72	2.62	6.30	11.88	18.50
	RC_{SE}	$\beta_{11}(\tau)$	0.17	0.15	0.12	0.07	0.00
	p-value		<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
	RC		0.12	3.36	8.60	14.88	21.06
	RC_{SE}	$\beta_{12}(\tau)$	0.13	0.06	0.09	0.07	0.18
	p-value		<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
	RC		22.91	25.43	29.09	33.17	37.01
	RC_{SE}	$\beta_{21}(\tau)$	0.75	0.47	0.31	0.61	1.61
	p-value		0.0449	0.0418	0.0393	0.0371	0.0352

Table A2. RCs (in %) in ML estimates and their corresponding SEs for the indicated parameter and respective *p*-values for the athletes dataset when observations {75, 178} and {75, 162, 178} are dropped separately.

Dropped Cases		Parameter	τ				
			0.10	0.25	0.50	0.75	0.90
75 and 178	RC		6.30	8.16	11.69	17.01	23.29
	RCSE	$\beta_{11}(\tau)$	0.41	0.39	0.34	0.28	0.19
	p-value		<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
	RC		1.75	5.32	10.58	16.84	22.97
	RCSE	$\beta_{12}(\tau)$	0.29	0.08	0.03	0.18	0.27
	p-value		<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
	RC		31	33.34	36.67	40.38	43.87
	RCSE	$\beta_{21}(\tau)$	0.04	0.27	0.42	0.11	0.91
	p-value		0.0674	0.0633	0.0600	0.0572	0.0546
75, 162 and 178	RC		5.43	7.27	10.80	16.13	22.45
	RCSE	$\beta_{11}(\tau)$	0.57	0.54	0.50	0.43	0.34
	p-value		<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
	RC		1.36	5.12	10.53	16.91	23.14
	RCSE	$\beta_{12}(\tau)$	0.39	0.18	0.12	0.26	0.35
	p-value		<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
	RC		43.37	45.46	48.35	51.53	54.51
	RCSE	$\beta_{21}(\tau)$	0.29	0.61	0.77	0.46	0.56
	p-value		0.1300	0.1251	0.1212	0.1178	0.1149

References

1. Azzalini, A. A class of distributions which includes the normal ones. *Scand. J. Stat.* **1985**, *12*, 171–178.
2. Azzalini, A. Further results on a class of distributions which includes the normal ones. *Statistica* **1986**, *46*, 199–208.
3. Arellano-Valle, R.B.; Gómez, H.W.; Quintana, F.A. A New Class of Skew-Normal Distributions. *Commun. Stat. Theory Methods* **2004**, *33*, 1465–1480. [CrossRef]
4. Pewsey, A.; Gómez, H.W.; Bolfarine, H. Likelihood-based inference for power distributions. *Test* **2012**, *21*, 775–789. [CrossRef]
5. Lehmann, E.L. The power of rank tests. *Ann. Math. Statist.* **1953**, *24*, 23–43. [CrossRef]
6. Durrans, S.R. Distributions of fractional order statistics in hydrology. *Water Resour. Res.* **1992**, *28*, 1649–1655. [CrossRef]
7. Gupta, D.; Gupta, R.C. Analyzing skewed data by power normal model. *Test* **2008**, *17*, 197–210. [CrossRef]
8. Castillo, N.O.; Gallardo, D.I.; Bolfarine, H.; Gómez, H.W. Truncated power-normal distribution with application to non-negative measurements. *Entropy* **2018**, *20*, 433. [CrossRef]
9. Martínez-Flórez, G.; Arnold, B.C.; Bolfarine, H.; Gómez, H.W. The alpha-power tobit model. *Commun. Stat. Theory Methods* **2013**, *42*, 633–643. [CrossRef]
10. Martínez-Flórez, G.; Bolfarine, H.; Gómez, H.W. Doubly censored power-normal regression models with inflation. *Test* **2015**, *24*, 265–286. [CrossRef]

11. Martínez-Flórez, G.; Bolfarine, H.; Gómez, H.W. Skew-normal alpha-power model. *Statistics* **2014**, *48*, 1414–1428. [CrossRef]

12. Martínez-Flórez, G.; Bolfarine, H.; Gómez, H.W. The log alpha-power asymmetric distribution with application to air pollution. *Environmetrics* **2014**, *25*, 44–56. [CrossRef]

13. Martínez-Flórez, G.; Bolfarine, H.; Gómez, H.W. Asymmetric regression models with limited responses with an application to antibody response to vaccine. *Biom. J.* **2013**, *55*, 156–172. [CrossRef] [PubMed]

14. Cook, R.D. Detection of influential observation in linear regression. *Technometrics* **1977**, *19*, 15–18.

15. Cook, R.D.; Weisberg, S. *An Introduction to Regression Graphics;* Wiley: New York, NY, USA, 1994.

16. Galarza, C.E.; Lachos, V.H.; Barbosa, C.; Castro, L.M. Robust quantile regression using a generalized class of skewed distributions. *Stat* **2017**, *6*, 113–130.

17. Akaike, H. A new look at the statistical model identification. *IEEE Trans. Autom. Control* **1974**, *19*, 716–723. [CrossRef]

18. Kolmogorov, A.N. Sulla determinazione empirica di una legge di distribuzionc, *Giorn. Ist. Ital. Attuar.* **1933**, *4*, 83–91.

19. Dunn, P.; Smyth G. Randomized quantile residuals. *J. Comput. Graph. Stat.* **1996**, *5*, 236–244.