Hacettepe Journal of Mathematics & Statistics

Hacet. J. Math. Stat.
Volume 51 (4) (2022), 1189 – 1210
DOI : 10.15672/hujms.920545

Research Article

Comparison of the alternative parameter estimators of Pearson distributions by robustness criteria

Mustafa Ünlü*[1], Ali Kemal Şehirlioğlu[2]

[1] *Bingöl University, Faculty of Economics and Administrative Sciences, Department of Econometrics, Bingöl, Turkey*

[2] *Dokuz Eylül University,Faculty of Economics and Administrative Sciences, Department of Econometrics, İzmir, Turkey*

Abstract

Pearson's differential equation is used for fitting a distribution to a data set. The differential equation has some alternative moment-based estimators (depending on the transformation to data). The estimator used when no transformation is made on the data set has 4 elements, and the estimators that require any transformation have 3 elements. We describe all elements of the estimators by corresponding vectors. One of the factors affecting the preference of an estimator is robustness. We use covariance matrix, bias, relative efficiency and influence function as our robustness criteria. Our aim is to compare the performance of the estimators of the differential equation for some specific distributions (namely Type I, Type IV, Type VI and Type III). 10,000 samples with specific sizes were selected with replacement. Also, we evaluated the performance of the estimators over real-life data. Considering the results, there is no best estimator in all criteria. Depending on the criterion to be based, the estimator to be preferred varies.

Mathematics Subject Classification (2020). 34A30, 62E10 , 62F35

Keywords. Pearson differential equation, robustness, influence function, variance-covariance matrix

1. Introduction

Data based on observation are generally intended to be expressed by a probability function. One of the methods used for this purpose is the probability density functions (pdf) obtained from the Pearson differential equation whose parameters are estimated by the method of moments. The distributions which are obtained from the parameters of the differential equation are called the Pearson Family of Distributions. Within this family of distributions, there are 13 distributions with various skewness and kurtosis. Type I, Type IV and Type VI distributions are called the main types and there are also transition types (e.g. Type III).

*Corresponding Author.
Email addresses: munlu@bingol.edu.tr (M. Ünlü), kemal.sehirli@deu.edu.tr (A.K. Şehirlioğlu)
Received: 19.04.2021; Accepted: 15.04.2022

The literature about the Pearson distributions can be divided into two groups: while the first group deals with the system itself [4,24,37], the second one reviews some specific distributions in the system [5,7,42].

The original differential equation defined by Pearson has a first-degree polynomial in the numerator and a second-order polynomial in the denominator. However, this structure of the differential equation can be generalized and the system can be modified. Thus, new distributions with different skewness and kurtosis can be derived [10,21,34,38]. There are also alternative methods to fit a distribution in the Pearson system with the original differential equation. Andreev et al. [1] proposed D and λ parameters for distribution selection in the system. Their selection criteria based on the parameters of the original differential equation. Cohen [8] proposed a method for fitting a distribution in the Pearson system for truncated samples. Parrish [30] used loss function approach instead of method of moments to fit a distribution in the Pearson system. Apart from the estimators mentioned above, alternative estimators can be obtained by making some transformations (e.g. location and/or scale) on the data set. We will discuss these alternative estimators throughout this study.

The need for asymmetric and heavy-tailed distributions in statistical modeling increased interest in the Pearson Distribution Family. Type IV distribution in this system is widely used especially for financial data (see [25] and [26]). Besides, Type IV distribution is used in the field of astrophysics (see [43]). Type III distribution mostly was used to model the flood of the rivers (see [3,12,23,36]). Statistical quality control [35], optimization [22], and reliability analysis [28] are some other fields the Pearson system is used.

Type III distribution has a great interest in the literature. Arora and Singh [2] compared direct moments, mixed moments, maximum likelihood method, and entropy method in terms of robustness properties for log-Pearson Type III distribution. They used bias, standard deviation and root mean square error over the large and small sample sizes as comparison criteria. By the results, it was concluded that the method of direct moments and the mixed moments were more robust than the others. Naghavi et al. [27] compared the direct moments, logarithmic moments and mixed moments methods used in parameter estimation of the Type III distribution. They selected root mean square error and mean absolute deviation as the performance criteria and they investigated the most robust estimator among them. Regarding the results, any estimator is not definitely better than the others. Depending on the skewness of the data set and sample size, alternatives are preferred. Koutrouvelis and Canavos [19] compared the method of moments, the simplified conditional method of moments and the mixed method of moments for the Pearson Type III distribution, by Monte Carlo simulation with different skewness values and small and large sample sizes derived from this distribution. The robustness properties of the estimators were compared over the bias and normalized root mean square error.

The robustness properties of the Pearson system itself have not been examined in the literature. However, there are some robust statistics-related studies. Büyükkör and Şehirlioğlu [6] evaluated robustness properties of Type VI distribution via influence function. Sun et al. [39] developed robust clustering approach by using Type VII distributions. Posten [31] examined robustness of the one sample t-test over the Pearson system.

In the previous studies, the characteristics of the estimators of some distributions in the Pearson family of distributions (such as Type III, Type VII) have been discussed. However, there is no study in the literature deals with the estimators of the differential equation as a vector.

The parameter estimates of the Pearson differential equation are based on the method of moments. Four different estimators can be defined by location and/or scale transformation of the data set. These four estimators are: the estimator does not require any transformation on the data set ($\boldsymbol{\theta}_x$), the estimator obtained by setting the mean to zero ($\boldsymbol{\theta}_y$) (location), the estimator obtained by setting the mode to zero ($\boldsymbol{\theta}_z$) (location) and obtained by standardizing (location and scale) the data set ($\boldsymbol{\theta}_t$). The aim of this study is to investigate the robustness measurements of these estimators. In accordance with this purpose, we will obtain influence functions and covariance matrices for the estimators and make performance comparisons among them for different sample sizes over some types of Pearson distributions. Since the parameter estimators of the Pearson differential equation are based on the method of moments, they are very sensitive to outliers. Therefore, it is important to determine which of the alternative estimators are robust for some specific cases. The effects of location/scale transformations on estimators of the differential equations will be evaluated on certain measures. The transformations reduce the number of parameters to be estimated. It is expected that the transformations will yield better results under the specified criteria. We will discuss the validity of the expectations with regard to the simulation study.

The following parts are as follows: Section 2 describes the Pearson Family of distributions, the related differential equation, and some distributions (namely Type I, Type VI, Type IV, and Type III) obtained from this differential equation. Section 3 introduces briefly the robustness criteria by which the estimators examined throughout this study. Section 4, evaluates the performances of the specified distributions under the determined criteria for large and small samples. Section 5 concludes.

2. Pearson family of distributions

To fit a probability distribution to a data set, the Pearson family of distributions might be useful. The distributions of this family are derived from a differential equation.The process starts with the calculation of moments of a data set.Via the moments, parameters of the differential equation can be determined. After that stage, proper distribution can be found in the family. Details of this process are discussed in the following subsections.

2.1. Differential equation

The probability density functions within the distribution system defined by Karl Pearson are derived from the solution of the following differential equation [18]

$$\frac{df}{dx} = \frac{(x - \alpha_x)\, f}{\phi_0 + \phi_1 x + \phi_2 x^2} \tag{2.1}$$

The shape of distribution depends on the parameters (α_x, ϕ_0, ϕ_1 and ϕ_2) of the differential equation. General solution for all distributions in the system can be obtained by the recursive moment equation after some integral operations as follows [18]:

$$n\phi_0 \mu'_{n-1} + \left[(n+1)\,\phi_1 - \alpha_x\right] \mu'_n + \left[(n+2)\,\phi_2 + 1\right] \mu'_{n+1} = 0$$

where $\mu_i^{'}$ is the i-th crude moment. If the equation is solved for $n = 0, 1, 2, 3$ the parameters are obtained in terms of moments as follows:

$$\alpha_x = -\frac{-12{\mu'}_4{\mu'}^3{}_1 + 20{\mu'}_2{\mu'}_3{{\mu'}_1}^2 - 9{\mu'}_1{{\mu'}_2}^3 + 13{\mu'}_1{\mu'}_2{\mu'}_4 - 8{\mu'}_1{{\mu'}_3}^2 - 3{{\mu'}_2}^2{\mu'}_3 - {\mu'}_3{\mu'}_4}{A}$$

$$\phi_0 = -\frac{-3{\mu'}_2{\mu'}^2{}_1{\mu'}_4 + 4{{\mu'}_1}^2{{\mu'}_3}^2 - {{\mu'}_2}^2{\mu'}_3{\mu'}_1 - {\mu'}_1{\mu'}_3{\mu'}_4 + 4{{\mu'}_2}^2{\mu'}_4 - 3{\mu'}_2{{\mu'}_3}^2}{A}$$

$$\phi_1 = -\frac{-6{\mu'}_4{\mu'}^3{}_1 + 8{\mu'}_2{\mu'}_3{{\mu'}_1}^2 - 3{\mu'}_1{{\mu'}_2}^3 + 7{\mu'}_1{\mu'}_2{\mu'}_4 - 2{\mu'}_1{{\mu'}_3}^2 - 3{{\mu'}_2}^2{\mu'}_3 - {\mu'}_3{\mu'}_4}{A}$$

$$\phi_2 = -\frac{-4{\mu'}^3{}_1{\mu'}_3 - 3{{\mu'}_1}^2{{\mu'}_2}^2 + 2{{\mu'}_1}^2{\mu'}_4 - 10{\mu'}_1{\mu'}_2{\mu'}_3 + 6{{\mu'}_2}^3 - 2{\mu'}_2{\mu'}_4 + 3{{\mu'}_3}^2}{A}$$

where

$$A = 8{\mu'_1}^3\mu'_3 - 6{\mu'_1}^2{\mu'_2}^2 + 10{\mu'_1}^2\mu'_4 - 32\mu'_1\mu'_2\mu'_3 + 18{\mu'_2}^3 - 10\mu'_2\mu'_4 + 12{\mu'_3}^2.$$

The distribution of the random variable X can be defined as $X \sim P\left({\mu'}_1, {\mu'}_2, {\mu'}_3, {\mu'}_4\right)$. For a transformation $Y = X - \mu$, where μ is the mean of distribution X, the new random variable can be expressed as $Y \sim P\left(0, \mu_2, \mu_3, \mu_4\right)$ where μ_i is the i-th central moment. The transformation sets the mean of the new variable to the origin. The new estimators of the differential equation can be obtained as setting $\mu'_1 = 0$ in the previous step as follows

$$\alpha_y = \varphi_1 = -\frac{\mu_3\left(3\mu_2^2 + \mu_4\right)}{B}$$

$$\varphi_0 = -\frac{\mu_2\left(4\mu_2\mu_4 - 3\mu_3^2\right)}{B}$$

$$\varphi_2 = -\frac{-6\mu_2^3 + 2\mu_2\mu_4 - 3\mu_3^2}{B}$$

where

$$B = -18\mu_2^3 + 10\mu_2\mu_4 - 12\mu_3^2.$$

Similarly, for the random variable Y, a new transformation can be defined as $Z = Y - \alpha_y$ where α_y is the mode of Y. After the location transformation, the new random variable Z's mode is set to zero [18]. The new differential equation to be obtained by setting the mode to zero is as follows:

$$\frac{df}{dz} = \frac{zf}{\lambda_0 + \lambda_1 z + \lambda_2 z^2} \tag{2.2}$$

This differential equation system has 3 parameters to be estimated. Parameter estimators can be expressed by the following equations:

$$\lambda_0 = -\frac{-4{\mu'}_3{\mu'}_1{}^2 + 3{\mu'}_1{\mu'}_2{}^2 + {\mu'}_3{\mu'}_2}{2C}$$

$$\lambda_1 = -\frac{{\mu'}_1\left(-{\mu'}_2{\mu'}_1 + {\mu'}_3\right)}{C}$$

$$\lambda_2 = -\frac{-4{\mu'}_1{}^3 - 5{\mu'}_1{\mu'}_2 + {\mu'}_3}{2C}$$

where

$$C = 4{\mu'_1}^3 - 6\mu'_2\mu'_1 + 2\mu'_3.$$

Equations of the new parameters in terms of the original differential equation parameters are given below:

$$\lambda_0 = \phi_0 + \alpha_x^2\left(1 + \phi_2\right)$$

$$\lambda_1 = \alpha_x\left(1 + 2\phi_2\right)$$

$$\lambda_2 = \phi_2$$

Another available transformation is to standardizing the random variable. The new random variable can be defined as $T = (X - \mu)/\sigma$ where μ is the mean and σ is the standart deviation of X. The new random variable T can be expressed as $T \sim P\left(0, 1, {\mu'}_3, {\mu'}_4\right)$. The new estimators of the differantial equation now became:

$$\alpha_t = \psi_1 = -\frac{\mu_3\left(3 + \mu_4\right)}{D}$$

$$\psi_0 = -\frac{4\mu_4 - 3\mu_3^2}{D}$$

$$\psi_2 = -\frac{-6 + 2\mu_4 - 3\mu_3^2}{D}$$

where

$$D = -18 + 10\mu_4 - 12\mu_3^2.$$

All of the transformations on the random variable, reduce the number of parameters to be estimated from 4 to 3.

2.2. Main types

The Pearson distribution system is based on the value of the roots

$$r_1 = \frac{-\phi_1 + \sqrt{\phi_1^2 - 4\phi_0\phi_2}}{2\phi_2}$$

and

$$r_2 = \frac{-\phi_1 - \sqrt{\phi_1^2 - 4\phi_0\phi_2}}{2\phi_2}$$

of the equation $\phi_0 + \phi_1 x + \phi_2 x^2 = 0$. If the roots are in $r_2 < 0 < r_1$ form and opposite signed, the distribution of X is Type I, and if the roots are in $r_1 < r_2 < 0$ or $0 < r_1 < r_2$ form with identical sign, the distribution of X is Type VI, and if the roots are complex, the distribution of X is Type IV.

The pdf of Type I distribution, where the random variable is defined on the range $r_2 < x < r_1$ is:

$$f(x) = K(r_1 - x)^{m_1}(x + r_2)^{m_2}$$

where K is the constant of integration and is obtained as follows [11]:

$$K = \frac{1}{Beta\,(m_1+1, m_2+1)\,(r_1 - r_2)^{m_1+m_2+1}}$$

The pdf of Type VI distribution, where the random variable is defined on the range $-r_2 < x < \infty$ is:

$$f(x) = K(r_1 + x)^{m_1}(r_2 + x)^{m_2}$$

The constant of integration K can be calculated as follows [11]:

$$K = \frac{1}{Beta\,(-m_1 - m_2 - 1, m_1 + 1)\,(r_2 - r_1)^{m_1+m_2+1}}$$

The pdf of Type IV distribution, where the random variable is defined on the range $-\infty < x < \infty$ is:

$$f(x) = K\left[(x + r)^2 + s^2\right]^m e^{v.\arctan\theta}$$

The constant of integration K can be obtained by $x + r = s\tan\theta$ transformation [11].

$$K = \frac{s^{-2m-1}}{\exp\left(\frac{v\pi}{2}\right)\int\limits_{-\pi/2}^{\pi/2}(\cos\theta)^{-2m-2}\exp(-v\theta)\,d\theta}$$

Pearson family of distributions consists of three main types. Type III distribution is a transition type of this system and it can be defined by the case that $\phi_2 = 0$ in (2.1) . The pdf of Type III is

$$f(x) = Ke^{(x+r)/a}(x + r)^m,\ -r < x < \infty$$

where r is the root of the denominator in (2.1) (while $\phi_2 = 0$), $m = \frac{r}{\alpha_x} - 1$ and $K = \frac{1}{(-\alpha_x)^{m+1}\Gamma(m+1)}$. Other transition types are out of the scope of this study.

3. Robustness measures

The aim of the robust approach to classical modeling and data analysis is to generate reliable parameter estimates and methods to generate the corresponding hypothesis tests and confidence intervals, not only when the data fits exactly to any given distribution, but also when they fit approximately a distribution [14, 16] . Robust methods also aim to provide the best fit for the majority of data. If there are no outliers in the data set, robust methods give almost the same results as classical methods. As a result of the goodness of fit with the majority of the data set, robust methods are very reliable in determining outliers [14]. For these reasons, alternative estimators of the differential equation will be evaluated on some robustness measures.

Some of the tools used to determine the robustness characteristics of the estimators are introduced below.

Definition 3.1. Bias

Bias is defined as the expected value of the difference between an estimator and a parameter [32].

$$\text{Bias} = E\left(\hat{\theta}_n - \theta\right)$$

For a vector of estimators, bias can be calculated by Euclidean distance.

Definition 3.2. Asymptotic Relative Efficiency

For estimator vectors $\hat{\theta}_1$ and $\hat{\theta}_2$ on $\mathbb{R}^k$, asymptotic relative efficiency (ARE) can be defined as follows [33]

$$ARE\left(\hat{\boldsymbol{\theta}}_2, \hat{\boldsymbol{\theta}}_1\right) = \left(\frac{|\sum_1|}{|\sum_2|}\right)^{1/k}$$

where $\sum_i$ is the determinant of the corresponding covariance matrix. For an estimator on $\mathbb{R}$, it can be calculated as the ratio between variances.

Definition 3.3. Influence Function

Let T be a real valued functional on a subset of all probability measures of the real number space $\mathbb{R}$ and let F be a probability measure on $\mathbb{R}$. The probability measure Δ_x for any point $x \in \mathbb{R}$ has point mass 1. The joint distribution of F and Δ_x for $0 < \varepsilon < 1$ can be expressed as $(1-\varepsilon)\, F + \varepsilon\Delta_x$. Influence function for the estimator T under F is [15]

$$IF\,(T, F; x) = \lim_{\varepsilon\downarrow 0} \left\{T\left[(1-\varepsilon)\, F + \varepsilon\Delta_x\right] - T\,(F)\right\} \Big/ \varepsilon \tag{3.1}$$

For any $W = (1-\varepsilon)\, F + \varepsilon\Delta_x$ distribution, where $0 \le \varepsilon \le 1$ let $\mu'\,(W)$ be the mean of the contaminated distribution. Then the following equation holds for crude moments:

$$\mu'_r\,(W) = (1-\varepsilon)\, \mu'_r\,(F) + \varepsilon x^r \tag{3.2}$$

And for central moments,

$$\mu_s\,(W) = (1-\varepsilon) \int_{-\infty}^{\infty} \left[t - \mu'\,(W)\right]^s dF\,(t) + \varepsilon\left[x - \mu'\,(W)\right]^s \tag{3.3}$$

To substitute (3.2) into (3.3),the expression becomes [13]

$$\begin{aligned}\mu_s\,(W) = (1-\varepsilon) \int_{-\infty}^{\infty} \left[t - \mu'\,(F)\right]^s dF\,(t) - \varepsilon\left[x - \mu'\,(F)\right]^s dF\,(t)\\ + \varepsilon\{(1-\varepsilon)\left[x - \mu'\,(F)\right]\}^s\end{aligned} \tag{3.4}$$

By the general structure of the central moment of the contaminated distribution given by (3.4), the second central moment of the contaminated distribution can be found as follows:

$$\mu_2(W) = (1-\varepsilon)\left\{\mu_2(F) + \varepsilon^2[x - \mu'(F)]^2\right\} + \varepsilon(1-\varepsilon)^2[x - \mu'(F)]^2$$

By the general structure of the crude moment of the contaminated distribution given by (3.2), multiplicative moment structures such as $\mu'_s\mu'_r$ and $\mu'_s\mu'_r\mu'_u$ can be found respectively as follows [41]:

$$\begin{aligned}\mu'_s(W)\,\mu'_r(W) = (1-\varepsilon)^2\mu'_s(F)\,\mu'_r(F) + \varepsilon(1-\varepsilon)\,\mu'_s(F)\,x^r \\ + \varepsilon(1-\varepsilon)\,\mu'_r(F)\,x^s + \varepsilon^2 x^s x^r\end{aligned} \tag{3.5}$$

$$\begin{aligned}\mu'_s(W)\,\mu'_r(W)\,\mu'_u(W) &= (1-\varepsilon)^3\mu'_s(F)\,\mu'_r(F)\,\mu'_u(F) \\ &+ \varepsilon(1-\varepsilon)^2\left[\mu'_s(F)\,\mu'_u(F)\,x^r + \mu'_r(F)\,\mu'_u(F)\,x^s + \mu'_s(F)\,\mu'_r(F)\,x^u\right] \\ &+ \varepsilon^2(1-\varepsilon)\left[\mu'_u(F)\,x^{r+s} + \mu'_s(F)\,x^{r+u} + \mu'_r(F)\,x^{s+u}\right] + \varepsilon^3 x^{s+r+u}\end{aligned} \tag{3.6}$$

A similar approach can be used for central moments.

Definition 3.4. Covariance Matrices and Influence Functions of the Estimators

When investigating the variance-covariance of linear functions of moment estimators, the approaches defined by [9] and [18] can be used. However, if the function of interest is a nonlinear function of moment estimators (e.g. correlation coefficients, skewness and kurtosis measures), the specified approaches cannot be used. The approach proposed by [9] for the variance-covariance structures of functions which contains ratios of certain moments is explained by the following formulas. Regarding Cramer's theorem, a function of two central moment estimators m_v and m_ρ is expressed by $H(m_v, m_\rho)$ [9]. Expected value and variance of the random variable $H(m_v, m_\rho)$ is as follows, repectively:

$$E(H) = H_0 + O\left(\frac{1}{n}\right)$$

$$\operatorname{var}(H) = \sum_{i=1}^{n}\operatorname{var}(m_i)\left(\frac{\partial H}{\partial m_i}\right)^2 + 2\sum\sum_{i<j}\operatorname{cov}(m_i, m_j)\left(\frac{\partial H}{\partial m_i}\right)\left(\frac{\partial H}{\partial m_j}\right) + O\left(\frac{1}{n^{3/2}}\right)$$

where H_0 is the H functions value at points $m_v = \mu_v$ and $m_\rho = \mu_\rho$ [9] .

For a measure taking into account the variance-covariance of the estimators such as ARE, the covariance between two different H functions must be obtained. For this purpose, the following corollary was obtained by using Cramer's H theorem.

Corollary 3.5. *Covariance between two H functions*

Let H_1 and H_2 be functions of m_v and m_ρ. The covariance between these two functions is [41]

$$Cov(H_1, H_2) = \sum_{i=1}^{n}\operatorname{var}(m_i)\left(\frac{\partial H_1}{\partial m_i}\right)\left(\frac{\partial H_2}{\partial m_i}\right) + 2\sum\sum_{i<j}\operatorname{cov}(m_i, m_j)\left(\frac{\partial H_1}{\partial m_i}\right)\left(\frac{\partial H_2}{\partial m_j}\right)$$

This corollary can be extended to the situations for H function has 3 or more argumenets. Parameter estimators of the Pearson differential equation will change depending on the transformation of the random variable that was mentioned in Section 2.1. Alternative estimators can be expressed as:

$$\boldsymbol{\theta}_x = \begin{bmatrix}\alpha_x \\ \varphi_0 \\ \varphi_1 \\ \varphi_2\end{bmatrix}, \quad \boldsymbol{\theta}_y = \begin{bmatrix}\alpha_y \\ \phi_0 \\ \phi_2\end{bmatrix}, \quad \boldsymbol{\theta}_z = \begin{bmatrix}\lambda_0 \\ \lambda_1 \\ \lambda_2\end{bmatrix}, \quad \text{and} \quad \boldsymbol{\theta}_t = \begin{bmatrix}\alpha_t \\ \psi_0 \\ \psi_2\end{bmatrix}$$

Parameter estimators of the differential equation are nonlinear functions of sample moments. For instance, $\boldsymbol{\theta}_x$ consist of $\alpha_x = f(\mu'_1, \mu'_2, \mu'_3, \mu'_4)$, $\phi_i = f(\mu'_1, \mu'_2, \mu'_3, \mu'_4)$ for $i = 0, 1, 2$. Since these functions depend on moments, they can also be expressed as a function of the random variable. The influence function on (3.1) can be extended to estimator vectors. From this perspective, the influence functions of the estimator vectors are in the form of [17]

$$IF\left(\boldsymbol{\theta}, F; x\right) = \lim_{\varepsilon \downarrow 0} \frac{\boldsymbol{\theta}\left(W\right) - \boldsymbol{\theta}\left(F\right)}{\varepsilon}$$

The influence function of $\boldsymbol{\theta}_x$ is:

$$\begin{aligned} IF\left(\boldsymbol{\theta}_x, F; x\right) &= \tfrac{d}{d\varepsilon}\boldsymbol{\theta}_x\left(W\right)|_{\varepsilon=0} \\ &= \begin{bmatrix} IF\left(\alpha_x, F; x\right) \\ IF\left(\phi_0, F; x\right) \\ IF\left(\phi_1, F; x\right) \\ IF\left(\phi_2, F; x\right) \end{bmatrix} \end{aligned}$$

First influence function of this vector can be obtained by using (3.5) and (3.6) as follows:

$$\begin{aligned} IF\left(\alpha_x, F; x\right) = -2\alpha_x &+ \frac{\xi_1(\mu) + x\xi_2(\mu) + x^2\xi_3(\mu) + x^3\xi_4(\mu) + x^4\xi_5(\mu)}{A} \\ &- \frac{\alpha_x\left[\xi_6(\mu) + x\xi_7(\mu) + x^2\xi_8(\mu) + x^3\xi_9(\mu)\right]}{A} \end{aligned} \tag{3.7}$$

where

$$\begin{aligned}
\xi_1(\mu) &= 13\mu_1'\mu_2'\mu_4' - 24\mu_1'^3\mu_4' - 3\mu_2'^2\mu_3' - 8\mu_1'\mu_3'^2 - 18\mu_1'\mu_2'^3 + 40\mu_1'^2\mu_2'\mu_3' \\
\xi_2(\mu) &= 8\mu_3'^3 + 9\mu_2'^3 - 13\mu_2'\mu_4' + 36\mu_1'^2\mu_4' - 40\mu_1'\mu_2'\mu_3' \\
\xi_3(\mu) &= 6\mu_2'\mu_3' - 13\mu_1'\mu_4' + 27\mu_1'\mu_2'^2 - 20\mu_1^2\mu_3' \\
\xi_4(\mu) &= 3\mu_2'^2 + \mu_4' + 16\mu_1'\mu_3' - 20\mu_1'^2\mu_2' \\
\xi_5(\mu) &= 12\mu_1'^3 + \mu_3' - 13\mu_1'\mu_2' \\
\xi_6(\mu) &= 20\mu_2'\mu_4' - 54\mu_2' - 24\mu_3'^2 - 32\mu_1^3\mu_3' - 10\mu_1'^2\mu_4' + 96\mu_1'\mu_2'\mu_3' + 24\mu_1'^2\mu_2' + 24\mu_1'^2\mu_3' \\
\xi_7(\mu) &= 20\mu_1'\mu_4' - 32\mu_2'\mu_3' - 12\mu_1' \\
\xi_8(\mu) &= 54\mu_2'^2 - 10\mu_4' - 32\mu_1'\mu_3' - 12\mu_1'^2\mu_2' \\
\xi_9(\mu) &= 8\mu_1'^3 + 24\mu_3' - 32\mu_1'\mu_2'
\end{aligned}$$

and A is defined in Section 2.1. Similar approaches can be used for other components and other estimators (see AppendixC).

The covariance matrix of the estimator $\boldsymbol{\theta}_x$ is:

$$\sum_x = \begin{bmatrix} Var\left(\alpha_x\right) & Cov\left(\alpha_x, \phi_0\right) & Cov\left(\alpha_x, \phi_1\right) & Cov\left(\alpha_x, \phi_2\right) \\ Cov\left(\alpha_x, \phi_0\right) & Var\left(\phi_0\right) & Cov\left(\phi_0, \phi_1\right) & Cov\left(\phi_0, \phi_2\right) \\ Cov\left(\alpha_x, \phi_1\right) & Cov\left(\phi_0, \phi_1\right) & Var\left(\phi_1\right) & Cov\left(\phi_1, \phi_2\right) \\ Cov\left(\alpha_x, \phi_2\right) & Cov\left(\phi_0, \phi_2\right) & Cov\left(\phi_1, \phi_2\right) & Var\left(\phi_2\right) \end{bmatrix}$$

Diagonal elements of this matrix is obtained by Cramers H theorem and off-diagonals by recalling (see Appendix B). With a similar approach, covariance matrices of other estimators can be obtained.

4. Applications

We will evaluate the performances of the estimators by both a simulation study and a real-life data example in this section.

4.1. Simulation study

We use pearsrnd code in Matlab to create 4 different distributions (Type I, Type IV Type VI, and Type III) with 1,000 observations in order to compare the performances of the alternative estimators on the corresponding distributions. All this distributions can be Bell-shaped. The probability density functions of these distributions and the range of random variables are as follows

Type I:

$$f(x) = 5.2193 \times 10^{-27}(22.9504 - x)^{17.6493}(x - 5.8799)^{1.4153}$$

$$-5.8799 < x < 22.9504$$

Type IV:

$$f(x) = 473.2603\left[(x + 4.2122)^2 + 2.1815^2\right]^{-5.8076} \exp\left[7.3356. \arctan\left(\frac{x + 4.2122}{2.1815}\right)\right]$$

$$-\infty < x < \infty$$

Type VI:

$$f(x) = 8.5257 \times 10^{67}(x - 22.3818)^{-53.9867}(x - 4.7463)^{5.0569}$$

$$-4.7463 < x < \infty$$

Type III:

$$f(x) = 0.394e^{(x+4.6713)/0.7932}(x + 4.7463)^{4.8892}$$

$$-4.6713 < x < \infty$$

To compare the big-small sample performances of these distributions, 10,000 samples were selected from $N = 1,000$ sized populations with replacement for $n = 50$, $n = 100$, $n = 200$, $n = 400$ and $n = 800$ sample sizes. These sample sizes can be seen in the literature [23]. Thus, we have data matrices with dimensions of $50 \times 10,000$, $100 \times 10,000$, $200 \times 10,000$, $400 \times 10,000$ and $800 \times 10,000$. For a population of size 1000, the required sample size is 278 to represent the corresponding population [20].

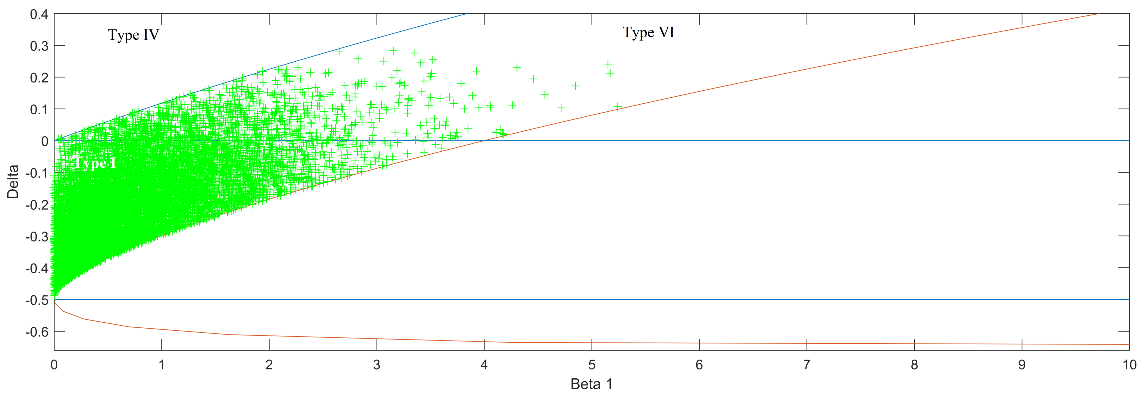


Figure 1. Locations of sample distributions on Craig chart for Type I distribution ($n = 50$).

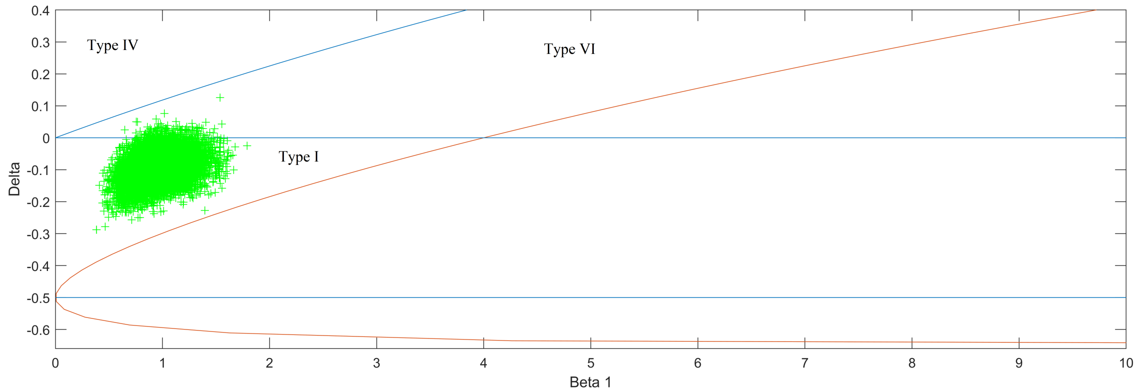


Figure 2. Locations of sample distributions on Craig chart for Type I distribution ($n = 800$).

Another way to choose a distribution suitable for the data set from the Pearson system is the Craig chart. The axes consist of δ and β_1 where $\beta_1 = \frac{\mu_3^2}{\mu_2^3}$, $\beta_2 = \frac{\mu_4}{\mu_2^2}$,and $\delta = \frac{2\beta_2 - 3\beta_1 - 6}{\beta_2 + 3}$ [40]. For the estimator $\boldsymbol{\theta}_y$ samples derived from Type I can be seen on Craigs chart in Figure 1. Some of the sample distributions are located in the Type VI area for $n = 50$ while for $n = 800$ most of them are in the Type I area. Similar results can be seen for other type of distributions.

The influence functions of all four estimators are unlimited (see Figure 3). The functional structures of the influence functions of all estimators except ϕ_0 and λ_0 are not affected by the distribution. As the type of distribution changes, the effect of an observation on the estimator changes. For all estimators, the effect of an observation is the lowest in the Type I distribution, while is the highest in the Type IV distribution.

The variance of all components of all estimators are the smallest in the Type IV distribution. The lowest variability among the components is seen in the coefficient of the quadratic variable ($\hat{\phi}_2$,$\hat{\varphi}_2$,$\hat{\lambda}_2$ and $\hat{\psi}_2$) of all the estimators for all distributions and all sample sizes (see Table 3-6). There is an almost reciprocal linear relationship between the variance and the number of observations. The relationship can be expressed as

$$Var_{k\times n} \cong \frac{Var_n}{k}$$

for $k = 1, 2, 4, 8, 16$ and $n = 50$ (see Appendix D). We use the indice k to adress our specific sample sizes. Correlations between the components of the parameter estimation vectors $\boldsymbol{\theta}_x$,$\boldsymbol{\theta}_y$,$\boldsymbol{\theta}_z$ and $\boldsymbol{\theta}_t$ are obtained by using the corollary in Section 2. The highest degree of linear relationship between the components can be seen between $\hat{\alpha}_x$ for $\boldsymbol{\theta}_x$,$\hat{\alpha}_y$ and $\hat{\varphi}_0$ for $\boldsymbol{\theta}_y$, $\hat{\alpha}_t$ and $\hat{\psi}_2$ for $\boldsymbol{\theta}_t$ (see Table 7-10). The sample size does not affect the correlation coefficient between these components. Another way to examine the linear relationship between the components of the estimators is to express them into a correlation matrix. The determinant of the correlation matrix reveals the dependency between the components. As the correlations are independent of the sample sizes, determinants of the correlation matrices are also independent of the sample sizes. Determinants of the correlation matrices for $n = 800$ is given in Table 1.

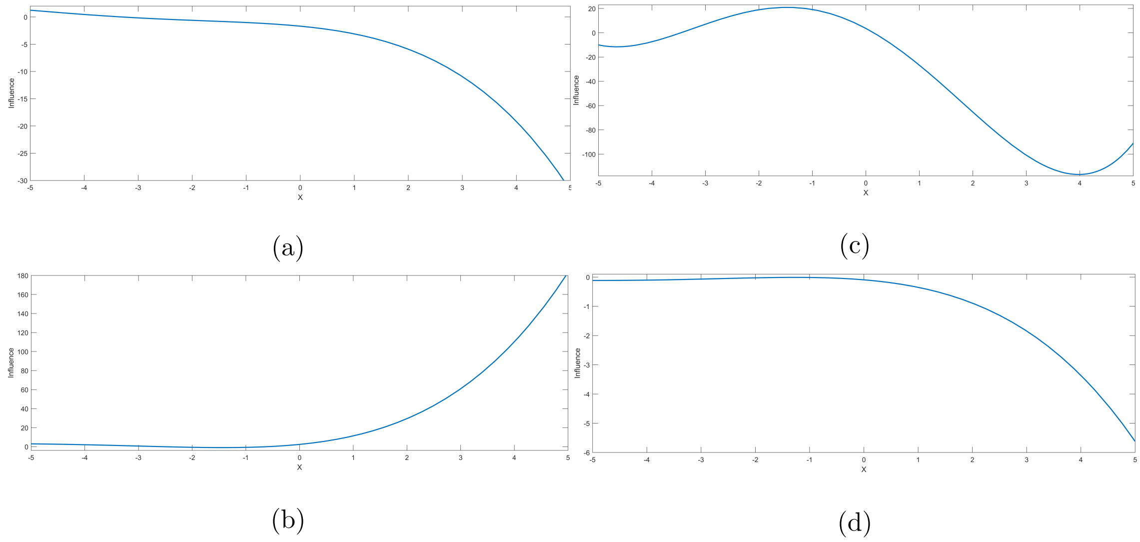


Figure 3. Influence Function of α_x(a), ϕ_0(b), ϕ_1(c) and ϕ_2(c) for Type I distribution.

Table 1. Determinants of the Correlation Matrices of the Estimators

	Type I	**Type VI**	**Type IV**	**Type III**
$\det \left\| P_{\hat{\theta}_x} \right\|$	0.0045	0.0007	0.0026	0.0473
$\det \left\| P_{\hat{\theta}_y} \right\|$	0.0726	0.0856	0.2250	0.2963
$\det \left\| P_{\hat{\theta}_z} \right\|$	0.0722	0.0663	0.1094	0.2933
$\det \left\| P_{\hat{\theta}_t} \right\|$	0.0720	0.0853	0.2279	0.2959

The mean vectors obtained from 10,000 samples from the given sample sizes of four estimators, and norms of the difference vectors defining the difference of these estimates from the population parameters are given in Table 11, Table 12, Table 13 and Table 14 (Appendix D), respectively. The norm of the bias vector is calculated by the Euclidean distance. With reference to the results obtained from the Type I and Type III distributions, the mean vectors of $\boldsymbol{\theta}_t$ in all sample sizes, is closer to the parameter vector. For Type VI and Type IV distributions, $\boldsymbol{\theta}_z$ is closer to population parameters. Generally, a negative linear relationship is valid between the biases and the sample sizes. In Table 15 (Appendix D), the asymptotic relative efficiencies of alternative estimators were compared. Some obvious results are as follows: The dispersion of $\boldsymbol{\theta}_y$ is less in all distributions and in all sample sizes compared to $\boldsymbol{\theta}_x$. The dispersion of $\boldsymbol{\theta}_z$ is smaller than $\boldsymbol{\theta}_x$ in all distributions and in all sample sizes.

4.2. Real-life data

Our sample is the daily brent oil returns between 31 December 2009 and 25 February 2015 [29]. The data consist of 2215 observations.

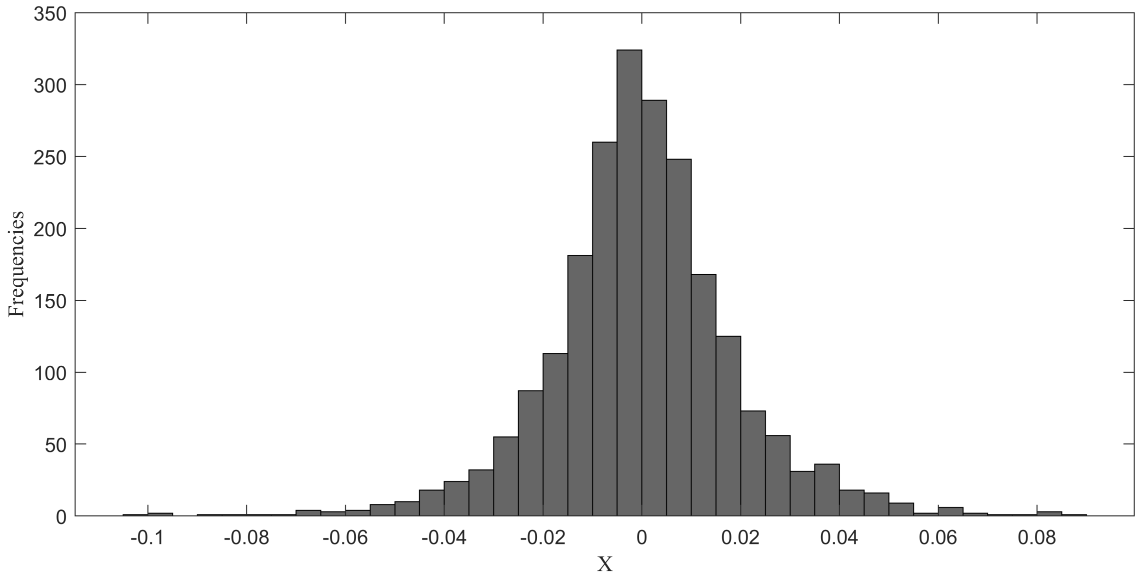


Figure 4. Histogram of the sample data.

Histogram of our sample data is given by Figure 4. Since the distribution of sample data seems to be unbounded on both ends, it might be fitted by Type IV distribution. Defining a suitable probability density function has great importance for any kind of posterior analysis. Our analysis will reveal the best available option among the alternative estimators for this particular instance.

Crude moments of the data are follows:

$$\mu_1' = 8.8650 \times 10^{-5}, \mu_2' = 3.7055 \times 10^{-4}, \mu_3' = -2.1047 \times 10^{-7}, \mu_4' = 7.9072 \times 10^{-7}.$$

Differential equation parameters are follows:

$$\alpha_x = 2.7324 \times 10^{-4}, \phi_0 = -2.1568 \times 10^{-4}, \phi_1 = 2.0930 \times 10^{-4}, \phi_2 = -0.1393$$

Since the roots of the denominator of the (2.1) are complex, the data fits the Type IV distribution. Corresponding pdf can be expressed as:

$$f(x) = 1.9863 \times 10^{-9} \left[(x - 7.5110 \times 10^{-4})^2 + 0.0393^2\right]^{-3.5887} \times \exp\left[-0.0871 \times \arctan\left(\frac{x - 7.5110 \times 10^{-4}}{0.0393}\right)\right]$$

Determinants of the estimators' correlation matrices are given in Table 2. Similar to the simulation results, determinant of the correlation matrix of $\boldsymbol{\theta}_y$ and $\boldsymbol{\theta}_t$ is greater than $\boldsymbol{\theta}_x$ and $\boldsymbol{\theta}_z$. However, determinant $\boldsymbol{\theta}_z$ is the minimum among others for the Type IV distribution.

Table 2. Determinants of the Correlation Matrices of the Estimators for Type IV Distribution

$\det\left\|P_{\hat{\theta}_x}\right\|$	$\det\left\|P_{\hat{\theta}_y}\right\|$	$\det\left\|P_{\hat{\theta}_z}\right\|$	$\det\left\|P_{\hat{\theta}_t}\right\|$
0.1684	0.7203	0.00002	0.7214

Influence funcitons of the estimators are given in Appendix E. Coherent to the simulation results, all of the influence functions are unbounded. Therefore, there is no difference among the influence functions. Since det $\left|P_{\hat{\theta}_t}\right|$ is the maximum, the best alternative is $\boldsymbol{\theta}_t$ for the Type IV distribution.

5. Discussion and conclusion

While the estimators obtained for any Pearson distribution parameters were handled individually in previous studies, the differential equation parameter estimators were handled both individually and as a vector, and evaluations were performed on these estimator vectors in this study.

A robust estimator is expected to have a bounded influence function. All the moment estimators are known to have unbounded influence functions. We have proved that our moment estimators also have unbounded influence functions as expected. There is no difference among these estimators in terms of influence functions.

The H-theorem, and the corollary obtained from this theorem were used to obtain the variances and covariances of the estimators. When the variances obtained from the H-theorem are examined, it is seen that they are negatively linearly correlated with the sample size. Components of $\boldsymbol{\theta}_t$ has the lowest variance among others.

Similar analysis were made for correlations between components. It is seen that for all estimators, the correlations between these components are independent of the sample sizes. The degree of linear relationship between components can vary depending on the distribution. The most favorable situation for the correlation between parameter that defines mode (α_x,α_y,α_t) and other location parameters r_1 and r_2 which are functions of other parameters (i.e $\phi_i, \varphi_i, \psi_i$ for $i = 0, 1, 2$) is when it is lowest. These two location measurements are desired to be as independent of each other as possible. This can be seen on $\boldsymbol{\theta}_y$ and $\boldsymbol{\theta}_t$ among all the estimators. Another approach is to considering the estimators as a vector. In this approach, linear relationship can be expressed by correlation matrix. The greater the determinant of the correlation matrix obtained by the corollary, the weaker the linear relationship between the variables. Our real-life data example revealed that for Type IV distribution, $\boldsymbol{\theta}_t$ is the best available estimator This result is consistent with the results obtained for the individual components. It should be noted that when choosing among the estimators, it is desired that the one with a higher determinant also has low variance value at the same time.

Biases of the estimators were examined by the norm of the difference vector. It is seen that location and scale transformation have a positive effect on both biases and ARE. For all estimators, as the sample size increases in all distributions, the norm of the deviation vector decreases as expected.

As a result of our analyzes, the following statements for the specific distributions are true: For Type I distribution, considering the determinants of the correlation matrices, all estimators except $\boldsymbol{\theta}_x$ recommended. Regarding ARE between estimators, however, $\boldsymbol{\theta}_t$ is the best alternative for Type I. $\boldsymbol{\theta}_y$ and $\boldsymbol{\theta}_t$ are the best alternatives with respect to determinants of the correlation matrices for Type VI. Considering ARE between these two, $\boldsymbol{\theta}_y$ preferred.Similar evaluations can be made for the other two distributions. In conclusion, $\boldsymbol{\theta}_y$ and $\boldsymbol{\theta}_z$ can be recommended for Type IV and Type III, respectively.

The number of inflection points is of great importance for Pearson system. If the inflection points are more than two, the distribution cannot be fitted by Pearson distribution family. This problem usually occurs for small sample sizes (especially for $n \leq 50$). This problem can be overcome by increasing the degree of the polynomial in the denominator of the differential equation and obtaining new distributions. The disadvantage of this approach is that the $5th$ and $6th$ order moments must be used. It should be noted that the standard errors of higher-order moments are bigger.

The sensitivity analysis is one of the methods that can be used for a robustness measure. In future studies, how sensitive the estimators are to an additional observation can be examined. Another analysis that can be made is to measure the distance between the population pdf and the sample pdfs. There are some available metrics such as Bhattacharya distance to calculate the distance between two pdfs. But the main limitation of this metric is that the two pdf should be in the same domain. For small sample sizes, the distribution of the data set tends to change its range. To overcome this problem, based on the marginal distributions of the parameters, bivariate distributions of these parameters can be obtained. Thus, a joint confidence region can be obtained with an ellipse with axes β_1 and δ, or β_1 and β_2.

References

[1] A. Andreev , A. Kanto and P. Malo, *Computational examples of a new method for distribution selection in the Pearson system*, J. Appl. Stat. **34** (4), 487-506, 2007.

[2] K. Arora and V.P. Singh, *A comparative evaluation of the estimators of the log Pearson type (LP) 3 distribution*, J. Hydrol. **105** (1-2), 19-37, 1989.

[3] F. Ashkar and B. Bobée, *The generalized method of moments as applied to problems of flood frequency analysis: some practical results for the log-Pearson type 3 distribution*, J. Hydrol. **90** (3-4), 199-217, 1987.

[4] K.O. Bowman and L.R. Shenton, *Approximate percentage points for Pearson distributions*, Biometrika **66** (1), 147-151, 1979.

[5] K.O. Bowman and L.R. Shenton, *Notes on the distribution of $\sqrt{b_1}$ in sampling from pearson distributions*, Biometrika **60** (1), 155-167, 1973.

[6] Y. Büyükkör and A.K. Şehirlioğlu, *A Robust Regression Method Based on Pearson Type VI Distribution*, in: Advances in Econometrics, Operational Research, Data Science and Actuarial Studies, 117-142, Switzerland:Springer, 2022.

[7] S. Chen and H. Nie, *Lognormal sum approximation with a variant of type IV Pearson distribution*, IEEE Commun. Lett. **12** (9), 630-632, 2008.

[8] A.C. Cohen, *Estimation of parameters in truncated Pearson frequency distributions*, Ann. Math. Stat. **22** (2), 256-265, 1951.

[9] H. Cramér, *Mathematical Methods of Statistics*, NJ: Princeton University Press, 1946.

[10] K.A. Dunning and J.N. Hanson, *Generalized Pearson distributions and nonlinear programing*, J. Stat. Comput. Simul. **6** (2), 115-128, 1977.

[11] W.P. Elderton and N.L. Johnson, *System of Frequency Curves*, USA: Cambridge University Press, 1969.

[12] J.F. England, J.D. Salas and R.D. Jarrett, *Comparisons of two moments-based estimators that utilize historical and paleoflood data for the log Pearson type III distribution*, Water Resour. Res. **39** (9), 1-16, 2003.

[13] A.M. Fiori and M. Zenga, *The meaning of kurtosis, the influence function and an early intuition by L. Faleschini*, Statistica **65** (2), 135-144, 2005.

[14] F.R. Hampel, E.M. Ronchetti, P.J. Rousseeuw and W.A. Stahel, *Robust Statistics, The Approach Based on Influence Functions*, New York: Wiley, 1986.

[15] F.R. Hampel, *The influence curve and its role in robust estimation*, J. Amer. Statist. Assoc. **69** (346), 383-393, 1974.

[16] P.J. Huber and E.M. Ronchetti, *Robust Statistics*, New York: John Wiley and Sons, 2009.

[17] T. Isogai, *On using influence functions for testing multivariate normality*, J. Amer. Statist. Assoc. **41** (1), 169-186, 1989.

[18] M.G. Kendall, A. Stuart, J.K. Ord, S.F. Arnold, A. O'Hagan and J. Forster, *Kendall's Advanced Theory of Statistics (Vol.1)*, London: Griffin, 1987.

[19] I.A. Koutrouvelis and G.C. Canavos, *Estimation in the Pearson type 3 distribution*, Water Resour. Res. **35** (9), 2693-2704, 1999.

[20] R.V. Krejcie and D.W. Morgan, *Determining sample size for research activities*, Educ. Psychol. Meas. **30** (3), 607-610, 1970.

[21] B. Lahcene, *On Pearson families of distributions and its applications*, Afr. J. Math. Comput. Sci. Res. **6** (5), 108-117, 2013.

[22] K.W. Liao and N.I.D.R. Biton, *A heuristic optimization considering probabilistic constraints via an equivalent single variable Pearson distribution system*, Appl. Soft Comput. **2019** (78), 670-684, 2019.

[23] N.C. Matalas and J.R. Wallis, *Eureka! It fits a Pearson type: 3 distribution*, Water Resour. Res. **9** (2), 281-298, 1973.

[24] Y. Nagahara, *A method of simulating multivariate nonnormal distributions by the Pearson distribution system and estimation*, Comput Stat Data Anal **40** (2004), 1-29, 2004.

[25] Y. Nagahara, *Non-Gaussian filter and smoother based on the Pearson distribution system*, J. Time Ser. Anal. **24** (6), 721-738, 2003.

[26] Y. Nagahara, *The PDF and CF of Pearson type IV distributions and the ML estimation of the parameters*, Stat Probab Lett **43** (3), 251-264, 1999.

[27] B. Naghavi, J. Cruise and K. Arora *A comparative evaluation of three estimators of log Pearson type 3 distribution*, Transp. Res. Rec. **1279**, 103-112, 1990.

[28] N.U. Nair and P.G. Sankaran, *Characterization of the Pearson family of distributions*, IEEE Trans Reliab **40** (1), 75-77, 1991.

[29] O. Özdemir and H. Emeç, *CDS Primleri, Hisse Senedi Piyasası ve Petrol Piyasası Arasındaki Oynaklık Yayılımı* in Ekonometride Ampirik Çalışmalar, 137-160, Ankara: Nobel, 2020.

[30] R.S. Parrish, *On an integrated approach to member selection and parameter estimation for Pearson distributions*, Comput Stat Data Anal **1**, 239-255, 1983.

[31] H.O. Posten, *The robustness of the one-sample t-test over the Pearson system*, J. Stat. Comput. Simul. **9** (2), 133-149, 1979.

[32] M.H. Quenouille, *Notes on bias in estimation*, Biometrika **43** (3/4), 353-360, 1956.

[33] R. Serfling,*Asymptotic Relative Efficiency in Estimation*, International Encyclopedia Of Statistical Science, 68-72, 2011.

[34] M. Shakil, B.M.G. Kibria and J.N. Singh, *A new family of distributions based on the generalized Pearson differential equation with some applications*, Austrian J. Stat. **39** (3), 259278, 2010.

[35] M. Shauly and Y. Parmet, *Comparison of Pearson distribution system and response modeling methodology (RMM) as models for process capability analysis of skewed data*, Qual. Reliab. Eng. Int. **2011** (27), 681-687, 2011.

[36] V.P. Singh and K. Singh, *Parameter estimation for log-Pearson type III distribution by POME*, J Hydraul Eng **114** (1), 112-122, 1988.

[37] H. Solomon and M.A. Stephens, *Approximations to density functions using Pearson curves*, J. Amer. Statist. Assoc. **73** (361), 153-160, 1978.

[38] S. Stavroyiannis, *On the generalised Pearson distribution for application in financial time series modelling*, Glob. Bus. Econ. Rev **16** (1), 1-14, 2014.

[39] J. Sun, A. Kabán and J.M. Garibaldi, *Robust mixture clustering using Pearson type VII distribution*, Pattern Recognit. Lett. **31** (2010), 2447-2454, 2010.

[40] A.K. Şehirlioğlu and S. Dündar, *Pearson Dağılış Ailesi*,İzmir: Ege Üniversitesi Basımevi, 2014.

[41] M. Ünlü, Comparison of The Parameter Estimators of The Main Types of Pearson Distributions by Robustness Criteria, PhD Thesis, Dokuz Eylül University, 2019.

[42] R. Willink, *A closed-form expression for the Pearson type IV distribution function*, Aust N Z J Stat **50** (2), 199-205, 2008.

[43] Z. Winiewski, *M-estimation with probabilistic models of geodetic observations*, J. Geod. **88** (10), 941-957, 2014.

Appendix A. Proof of the corollary

The H function defined by Cramer can be expressed by bivariate Taylor expansion. By this approach the H function can be expressed as:

$$H\left(m_v, m_\rho\right) = H\left(\mu_v, \mu_\rho\right) + \tfrac{\partial H(m_v, m_\rho)}{\partial \mu_v}\left(m_v - \mu_v\right) + \tfrac{\partial H(m_v, m_\rho)}{\partial \mu_\rho}\left(m_\rho - \mu_\rho\right)$$
$$+\tfrac{1}{2!}\begin{bmatrix} \tfrac{\partial^2 H(m_v, m_\rho)}{\partial {\mu_v}^2}\left(m_v - \mu_v\right)^2 \\ +2\tfrac{\partial H(m_v, m_\rho)}{\partial \mu_v}\tfrac{\partial H(m_v, m_\rho)}{\partial \mu_\rho}\left(m_v - \mu_v\right)\left(m_\rho - \mu_\rho\right) \\ +\tfrac{\partial^2 H(m_v, m_\rho)}{\partial {\mu_\rho}^2}\left(m_\rho - \mu_\rho\right)^2 \end{bmatrix} + \cdots$$

at points $m_v = \mu_v$ and $m_\rho = \mu_\rho$. After taking expected value and variance of both sides, expected value and variance of H function can be obtained. By this information, covariance between two H functions can be obtained.

Appendix B. Variances and covariances

The variances and the covariances between the components of $\boldsymbol{\theta}_x$ can be obtanied as follows for $k, l = 0, 1, 2$:

$$\operatorname{var}\left(\alpha_x\right) = \sum_{i=1}^{4} \operatorname{var}\left({m'}_i\right)\left(\frac{\partial \alpha_x}{\partial {m'}_i}\right)^2 + 2\sum\sum_{i<j} \operatorname{cov}\left({m'}_i, {m'}_j\right)\left(\frac{\partial \alpha_x}{\partial {m'}_i}\right)\left(\frac{\partial \alpha_x}{\partial {m'}_j}\right)$$

$$\operatorname{var}\left(\varphi_k\right) = \sum_{i=1}^{4} \operatorname{var}\left({m'}_i\right)\left(\frac{\partial \varphi_k}{\partial {m'}_i}\right)^2 + 2\sum\sum_{i<j} \operatorname{cov}\left({m'}_i, {m'}_j\right)\left(\frac{\partial \varphi_k}{\partial {m'}_i}\right)\left(\frac{\partial \varphi_k}{\partial {m'}_j}\right)$$

$$\operatorname{cov}\left(\alpha_x, \varphi_k\right) = \sum_{i=1}^{4} \operatorname{var}\left({m'}_i\right)\left(\frac{\partial \alpha_x}{\partial {m'}_i}\right)\left(\frac{\partial \varphi_k}{\partial {m'}_i}\right) + 2\sum\sum_{i<j} \operatorname{cov}\left({m'}_i, {m'}_j\right)\left(\frac{\partial \alpha_x}{\partial {m'}_i}\right)\left(\frac{\partial \varphi_k}{\partial {m'}_j}\right)$$

$$\operatorname{cov}\left(\varphi_k, \varphi_l\right) = \sum_{i=1}^{4} \operatorname{var}\left({m'}_i\right)\left(\frac{\partial \varphi_k}{\partial {m'}_i}\right)\left(\frac{\partial \varphi_l}{\partial {m'}_i}\right) + 2\sum\sum_{i<j} \operatorname{cov}\left({m'}_i, {m'}_j\right)\left(\frac{\partial \varphi_k}{\partial {m'}_i}\right)\left(\frac{\partial \varphi_l}{\partial {m'}_j}\right)$$

Appendix C. Influence function

$$IF\left(\alpha_x, F; x\right) = \lim_{z \downarrow 0} \frac{1}{\varepsilon}\left[-\frac{\varsigma_1 + \varsigma_2 + \varsigma_3}{\omega_1 + \omega_2 + \omega_3} - \frac{\upsilon_1 + \upsilon_2 + \upsilon_3}{\tau_1 + \tau_2 + \tau_3}\right]$$

where

$$\begin{aligned}
\varsigma_1 &= -12\mu_4'(W)\mu_1'^3(W) + 20\mu_2'(W)\mu_3'(W)\mu_1'^2(W) \\
\varsigma_2 &= -9\mu_1'(W)\mu_2'^3(W) + 13\mu_1'(W)\mu_2'(W)\mu_4'(W) \\
\varsigma_3 &= -8\mu_1'(W)\mu_3'^2(W) - 3\mu_2'^2(W)\mu_3'(W) - \mu_3'(W)\mu_4'(W) \\
\omega_1 &= 8\mu_1'^3(W)\mu_3'(W) - 6\mu_1'^2(W)\mu_2'^2(W) \\
\omega_2 &= 10\mu_1'^2(W)\mu_4'(W) - 32\mu_1'(W)\mu_2'(W)\mu_3'(W) \\
\omega_3 &= 18\mu_2'^3(W) - 10\mu_2'(W)\mu_4'(W) + 12\mu_3'^2(W) \\
\upsilon_1 &= -12\mu_4'(F)\mu_1'^3(F) + 20\mu_2'(F)\mu_3'(F)\mu_1'^2(F) \\
\upsilon_2 &= -9\mu_1'(F)\mu_2'^3(F) + 13\mu_1'(F)\mu_2'(F)\mu_4'(F) \\
\upsilon_3 &= -8\mu_1'(F)\mu_3'^2(F) - 3\mu_2'^2(F)\mu_3'(F) - \mu_3'(F)\mu_4'(F) \\
\tau_1 &= 8\mu_1'(F)\mu_3'(F) - 6\mu_1'^2(F)\mu_2'^2(F) + 10\mu_1'(F)\mu_4'(F) \\
\tau_2 &= -32\mu_1'(F)\mu_2'(F)\mu_3'(F) + 18\mu_2'^3(F) \\
\tau_3 &= -10\mu_2'(F)\mu_4'(F) + 12\mu_3'^2(F)
\end{aligned}$$

By (3.4), we can write:

$$IF\left(\alpha_x, F; x\right) = \lim_{z\downarrow 0} \frac{1}{\varepsilon}\left[-\frac{\gamma_1+\gamma_2+\gamma_3+\gamma_4+\gamma_5+\gamma_6+\gamma_7}{\rho_1+\rho_2+\rho_3+\rho_4+\rho_5+\rho_6} - \frac{\kappa_1+\kappa_2+\kappa_3}{\pi_1+\pi_2+\pi_3}\right]$$

where

$$\begin{aligned}
\gamma_1 &= \left[9\mu_1'(F)(1-\varepsilon) - 9\varepsilon x\right]\left[-\varepsilon x^2 - \mu_2'(F)(1-\varepsilon)\right]^3 \\
\gamma_2 &= -\left[8\mu_1'(F)(1-\varepsilon) - 8\varepsilon x\right]\left[\varepsilon x^3 - \mu_3'(F)(1-\varepsilon)\right]^3 \\
\gamma_3 &= -\left[\varepsilon x^3 - \mu_3'(F)(1-\varepsilon)\right]\left[\varepsilon x^4 - \mu_4'(F)(1-\varepsilon)\right] \\
\gamma_4 &= \left[\mu_1'(F)(1-\varepsilon) - \varepsilon x\right]^4\left[12\varepsilon x^4 - 12\mu_4'(F)(1-\varepsilon)\right] \\
\gamma_5 &= 3\left[\varepsilon x^2 - \mu_2'(F)(1-\varepsilon)\right]^2\left[\varepsilon x^3 - \mu_3'(F)(1-\varepsilon)\right] \\
\gamma_6 &= \left[\mu_1'(F)(1-\varepsilon) - \varepsilon x\right]^2\left[\varepsilon x^3 - \mu_3'(F)(1-\varepsilon)\right]\left[20\varepsilon x^2 - 20\mu_2'(F)(1-\varepsilon)\right] \\
\gamma_7 &= \left[12\mu_1'(F)(1-\varepsilon) - 12\varepsilon x\right]\left[\varepsilon x^2 - \mu_2'(F)(1-\varepsilon)\right]\left[\varepsilon x^3 - \mu_3'(F)(1-\varepsilon)\right] \\
\rho_1 &= 8\left[\varepsilon x - \mu_1'(F)(1-\varepsilon)\right]^3\left[-\varepsilon x^3 - \mu_3'(F)(1-\varepsilon)\right] \\
\rho_2 &= 10\left[\mu_1'(F)(1-\varepsilon) - \varepsilon x\right]^2\left[\varepsilon x^4 - \mu_4'(F)(1-\varepsilon)\right] \\
\rho_3 &= \left[\varepsilon x^4 - \mu_4'(F)(1-\varepsilon)\right]\left[10\varepsilon x^2 - 10\mu_2'(F)(1-\varepsilon)\right] \\
\rho_4 &= 6\left[\varepsilon x - \mu_1'(F)(1-\varepsilon)\right]\left[\varepsilon x^2 - \mu_2'(F)(1-\varepsilon)\right]^2 \\
\rho_5 &= 18\left[\varepsilon x^2 - \mu_2'(F)(1-\varepsilon)\right]^3 - 12\left[\varepsilon x^3 - \mu_3'(F)(1-\varepsilon)\right]^2 \\
\rho_6 &= \left[32\mu_1'(F)(1-\varepsilon) - 32\varepsilon x\right]\left[\varepsilon x^2 - \mu_2'(F)(1-\varepsilon)\right]\left[\varepsilon x^3 - \mu_3'(F)(1-\varepsilon)\right]
\end{aligned}$$

and

$$\kappa_1 = -12\mu_4'(F)\mu_1'^3(F) + 20\mu_2'(F)\mu_3'(F)\mu_1'^2(F)$$
$$\kappa_2 = -9\mu_1'(F)\mu_2'^3(F) + 13\mu_1'(F)\mu_2'(F)\mu_4'(F)$$
$$\kappa_3 = -8\mu_1'(F)\mu_3'^2(F) - 3\mu_2'^2(F)\mu_3'(F) - \mu_3'(F)\mu_4'(F)$$
$$\pi_1 = 8\mu_1'\mu_3' - 6\mu_1'^2\mu_2'^2 + 10\mu_1'^2\mu_4'$$
$$\pi_2 = -32\mu_1'(F)\mu_2'(F)\mu_3'(F) + 18\mu_2'^3(F)$$
$$\pi_3 = -10\mu_2'(F)\mu_4'(F) + 12\mu_3'^2(F)$$

After differentiating with respect to ϵ, (3.14) can be obtained.

Appendix D. Simulation results

Table 3. The Variances of the Components of $\boldsymbol{\theta}_x$

$n = 50$	**Type I**	**Type VI**	**Type IV**	**Type III**
$\mathrm{Var}(\hat{\alpha}_x)$	0.3582	0.0828	0.0507	0.5584
$\mathrm{Var}\left(\hat{\phi}_0\right)$	6.3323	0.6781	0.5182	8.9231
$\mathrm{Var}\left(\hat{\phi}_1\right)$	0.3518	0.4445	0.1142	0.4950
$\mathrm{Var}\left(\hat{\phi}_2\right)$	0.0141	0.0225	0.0034	-

Table 4. The Variances of the Components of $\boldsymbol{\theta}_y$

$n = 50$	**Type I**	**Type VI**	**Type IV**	**Type III**
$\mathrm{Var}(\hat{\alpha}_y)$	0.2929	0.0679	0.0516	0.3959
$\mathrm{Var}(\hat{\varphi}_0)$	3.0491	0.2202	0.0649	7.7625
$\mathrm{Var}(\hat{\varphi}_2)$	0.0137	0.0214	0.0036	-

Table 5. The Variances of the Components of $\boldsymbol{\theta}_z$

$n = 50$	Type I	Type VI	Type IV	Type III
$\mathrm{Var}\left(\hat{\lambda}_0\right)$	2.3626	20.4056	29.7875	6.19×10^8
$\mathrm{Var}\left(\hat{\lambda}_1\right)$	0.5206	148.3669	0.0136	40.4044
$\mathrm{Var}\left(\hat{\lambda}_2\right)$	0.0415	0.0700	0.0400	–

Table 6. The Variances of the Components of $\boldsymbol{\theta}_t$

$n = 50$	**Type I**	**Type VI**	**Type IV**	**Type III**
$\mathrm{Var}(\hat{\alpha}_t)$	0.0902	0.0683	0.0376	0.1240
$\mathrm{Var}\left(\hat{\psi}_0\right)$	0.3182	0.2654	0.0681	0.6483
$\mathrm{Var}\left(\hat{\psi}_2\right)$	0.0216	0.0209	0.0020	-

Table 7. The Correlations Between of the Components of $\boldsymbol{\theta}_x$

$n = 800$	**Type I**	**Type VI**	**Type IV**	**Type III**
Corr $\left(\hat{\alpha}_x, \hat{\phi}_0\right)$	0.8557	0.6458	0.8316	0.7549
Corr $\left(\hat{\alpha}_x, \hat{\phi}_1\right)$	0.4135	-0.0961	0.4963	0.9036
Corr $\left(\hat{\alpha}_x, \hat{\phi}_2\right)$	-0.5324	-0.4259	-0.1864	-
Corr $\left(\hat{\phi}_0, \hat{\phi}_1\right)$	0.3483	0.5512	0.7414	0.8601
Corr $\left(\hat{\phi}_0, \hat{\phi}_2\right)$	-0.6710	0.1955	-0.0922	-
Corr $\left(\hat{\phi}_1, \hat{\phi}_2\right)$	0.4427	0.9237	0.7640	-

Table 8. The Correlations Between of the Components of $\boldsymbol{\theta}_y$

$n = 800$	**Type I**	**Type VI**	**Type IV**	**Type III**
Corr $(\hat{\alpha}_y, \hat{\varphi}_0)$	0.8127	0.7431	0.6640	0.8389
Corr $(\hat{\alpha}_y, \hat{\varphi}_2)$	-0.6037	-0.4494	-0.1864	-
Corr $(\hat{\varphi}_0, \hat{\varphi}_2)$	-0.8690	-0.8553	-0.6847	-

Table 9. The Correlations Between of the Components of $\boldsymbol{\theta}_z$

$n = 800$	**Type I**	**Type VI**	**Type IV**	**Type III**
Corr $\left(\hat{\lambda}_0, \hat{\lambda}_1\right)$	0.8033	0.8101	0.7337	0.8406
Corr $\left(\hat{\lambda}_0, \hat{\lambda}_2\right)$	-0.8199	-0.8733	-0.8655	-
Corr $\left(\hat{\lambda}_1, \hat{\lambda}_2\right)$	-0.8686	-0.8312	-0.7156	-

Table 10. The Correlations Between of the Components of $\boldsymbol{\theta}_t$

$n = 800$	**Type I**	**Type VI**	**Type IV**	**Type III**
Corr $\left(\hat{\alpha}_t, \hat{\psi}_0\right)$	0.8120	0.7479	0.6554	0.8391
Corr $\left(\hat{\alpha}_t, \hat{\psi}_2\right)$	-0.6052	-0.4628	-0.1827	-
Corr $\left(\hat{\psi}_0, \hat{\psi}_2\right)$	-0.8705	-0.8569	-0.6883	-

Table 11. Norms of the Difference Vectors for Type I Distribution

n	norm $\left(\boldsymbol{\theta}_x - \hat{\boldsymbol{\theta}}_x\right)$	norm $\left(\boldsymbol{\theta}_y - \hat{\boldsymbol{\theta}}_y\right)$	norm $\left(\boldsymbol{\theta}_z - \hat{\boldsymbol{\theta}}_z\right)$	norm $\left(\boldsymbol{\theta}_t - \hat{\boldsymbol{\theta}}_t\right)$
50	1.1680	1.6999	0.9120	0.5911
100	0.9018	1.0298	0.5521	0.3260
200	0.5376	0.5373	0.2872	0.1690
400	0.2577	0.2668	0.1388	0.0745
800	0.1297	0.1183	0.0634	0.0336

Table 12. Norms of the Difference Vectors for Type VI Distribution

n	norm $\left(\boldsymbol{\theta}_x - \hat{\boldsymbol{\theta}}_x\right)$	norm $\left(\boldsymbol{\theta}_y - \hat{\boldsymbol{\theta}}_y\right)$	norm $\left(\boldsymbol{\theta}_z - \hat{\boldsymbol{\theta}}_z\right)$	norm $\left(\boldsymbol{\theta}_t - \hat{\boldsymbol{\theta}}_t\right)$
50	33.3814	1.0459	978.09	0.3402
100	0.5835	0.4252	0.3479	0.4463
200	0.2588	0.2049	0.1622	0.2061
400	0.1256	0.1010	0.0828	0.1042
800	0.0619	0.0543	0.0423	0.0519

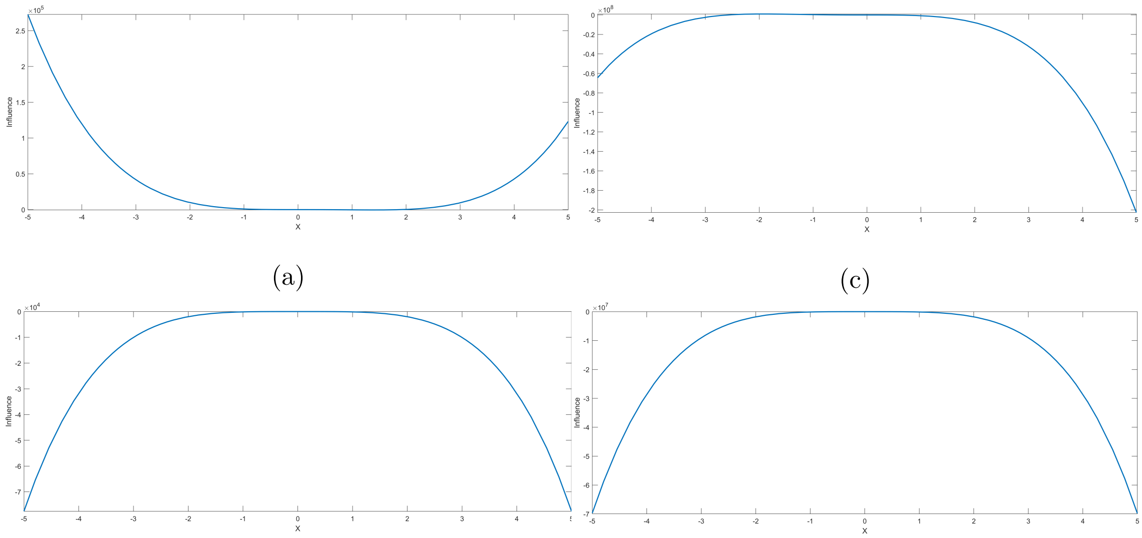
Table 13. Norms of the Difference Vectors for Type IV Distribution

n	norm $\left(\boldsymbol{\theta}_x - \hat{\boldsymbol{\theta}}_x\right)$	norm $\left(\boldsymbol{\theta}_y - \hat{\boldsymbol{\theta}}_y\right)$	norm $\left(\boldsymbol{\theta}_z - \hat{\boldsymbol{\theta}}_z\right)$	norm $\left(\boldsymbol{\theta}_t - \hat{\boldsymbol{\theta}}_t\right)$
50	0.4385	0.1339	0.0759	0.1478
100	0.2304	0.0613	0.0360	0.0718
200	0.1221	0.0271	0.0163	0.0396
400	0.0592	0.0184	0.0124	0.0258
800	0.0315	0.0146	0.0100	0.0165

Table 14. Norms of the Difference Vectors for Type III Distribution

n	norm $\left(\boldsymbol{\theta}_x - \hat{\boldsymbol{\theta}}_x\right)$	norm $\left(\boldsymbol{\theta}_y - \hat{\boldsymbol{\theta}}_y\right)$	norm $\left(\boldsymbol{\theta}_z - \hat{\boldsymbol{\theta}}_z\right)$	norm $\left(\boldsymbol{\theta}_t - \hat{\boldsymbol{\theta}}_t\right)$
50	3.7951	6.2196	6.6e+07	2.8056
100	3.5491	3.6116	1.7e+03	0.8956
200	1.9712	2.0185	11.6367	0.6075
400	1.5363	1.5400	0.8167	0.4770
800	1.2904	1.3025	0.7063	0.4058

Appendix E. Real-life data results



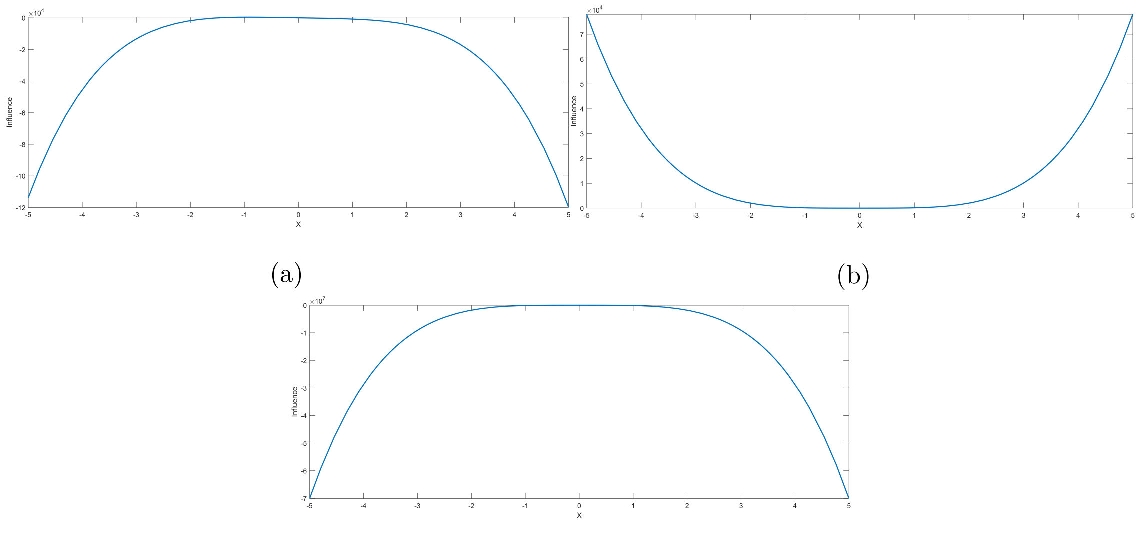
(a) (c)

(b) (d)

Figure 5. Influence Function of α_x(a), ϕ_0(b), ϕ_1(c) and ϕ_2(c) for Type IV distribution.

Table 15. Relative Efficiency Comparison of the Estimators

$n = 50$	**Type I**	**Type VI**	**Type IV**	**Type III**
$\boldsymbol{\theta}_x/\boldsymbol{\theta}_y$	0.4746	0.3361	0.3143	0.4809
$\boldsymbol{\theta}_x/\boldsymbol{\theta}_z$	0.3605	0.0493	0.6094	3.9×10^{-4}
$\boldsymbol{\theta}_y/\boldsymbol{\theta}_z$	0.6932	0.0773	2.4180	2.3×10^{-5}
$\boldsymbol{\theta}_x/\boldsymbol{\theta}_t$	1.0652	0.3086	0.2869	1.5993
$\boldsymbol{\theta}_y/\boldsymbol{\theta}_t$	2.9388	0.8788	0.8865	6.0651
$\boldsymbol{\theta}_t/\boldsymbol{\theta}_z$	0.2359	0.8340	0.2674	3.9×10^{-6}
$n = 100$	**Type I**	**Type VI**	**Type IV**	**Type III**
$\boldsymbol{\theta}_x/\boldsymbol{\theta}_y$	0.3920	0.2783	0.2647	0.3698
$\boldsymbol{\theta}_x/\boldsymbol{\theta}_z$	0.3732	0.0832	0.2465	0.0414
$\boldsymbol{\theta}_y/\boldsymbol{\theta}_z$	0.9364	0.1998	0.9089	0.0374
$\boldsymbol{\theta}_x/\boldsymbol{\theta}_t$	0.9731	0.2616	0.2245	1.3017
$\boldsymbol{\theta}_y/\boldsymbol{\theta}_t$	3.2499	0.9190	0.8163	6.6039
$\boldsymbol{\theta}_t/\boldsymbol{\theta}_z$	0.2877	1.0779	0.9073	0.0057
$n = 200$	**Type I**	**Type VI**	**Type IV**	**Type III**
$\boldsymbol{\theta}_x/\boldsymbol{\theta}_y$	0.3340	0.2310	0.2202	0.2817
$\boldsymbol{\theta}_x/\boldsymbol{\theta}_z$	0.4115	0.2685	0.1949	0.1080
$\boldsymbol{\theta}_y/\boldsymbol{\theta}_z$	1.3211	1.2224	0.8495	0.2372
$\boldsymbol{\theta}_x/\boldsymbol{\theta}_t$	0.8441	0.2304	0.1904	1.0139
$\boldsymbol{\theta}_y/\boldsymbol{\theta}_t$	3.4740	0.9796	0.8275	6.8265
$\boldsymbol{\theta}_t/\boldsymbol{\theta}_z$	0.3793	1.2424	0.8585	0.0347
$n = 400$	**Type I**	**Type VI**	**Type IV**	**Type III**
$\boldsymbol{\theta}_x/\boldsymbol{\theta}_y$	0.2746	0.1928	0.1842	0.2238
$\boldsymbol{\theta}_x/\boldsymbol{\theta}_z$	0.3387	0.2367	0.1684	0.3208
$\boldsymbol{\theta}_y/\boldsymbol{\theta}_z$	1.3231	1.3143	0.8871	1.7163
$\boldsymbol{\theta}_x/\boldsymbol{\theta}_t$	0.7254	0.1872	0.1620	0.8173
$\boldsymbol{\theta}_y/\boldsymbol{\theta}_t$	3.6523	0.9616	0.8427	6.9788
$\boldsymbol{\theta}_t/\boldsymbol{\theta}_z$	0.3623	1.3667	1.0527	0.2459
$n = 800$	**Type I**	**Type VI**	**Type IV**	**Type III**
$\boldsymbol{\theta}_x/\boldsymbol{\theta}_y$	0.2307	0.1630	0.1566	0.1763
$\boldsymbol{\theta}_x/\boldsymbol{\theta}_z$	0.3045	0.2140	0.1471	0.2719
$\boldsymbol{\theta}_y/\boldsymbol{\theta}_z$	1.4480	1.4378	0.9201	1.9150
$\boldsymbol{\theta}_x/\boldsymbol{\theta}_t$	0.6201	0.1604	0.1390	0.6511
$\boldsymbol{\theta}_y/\boldsymbol{\theta}_t$	3.7197	0.9837	0.8588	7.0976
$\boldsymbol{\theta}_t/\boldsymbol{\theta}_z$	0.3877	1.4654	1.0994	0.2698



(a) (b)

(c)

Figure 6. Influence Function of α_y(a), φ_0(b), and φ_2(c) for Type IV distribution.

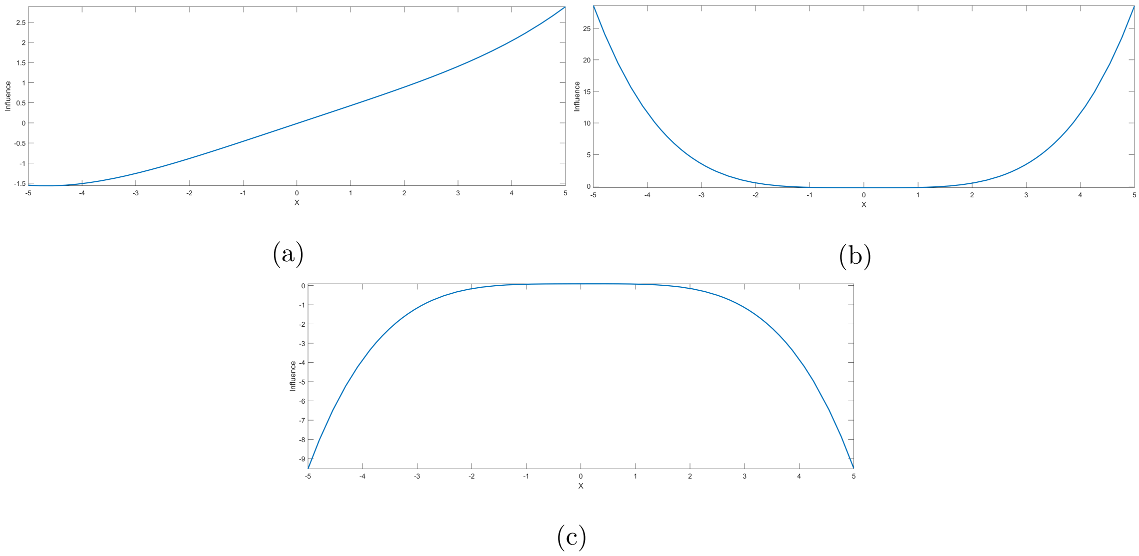


Figure 7. Influence Function of ψ_0(a), ψ_1(b), and ψ_2(c) for Type IV distribution.

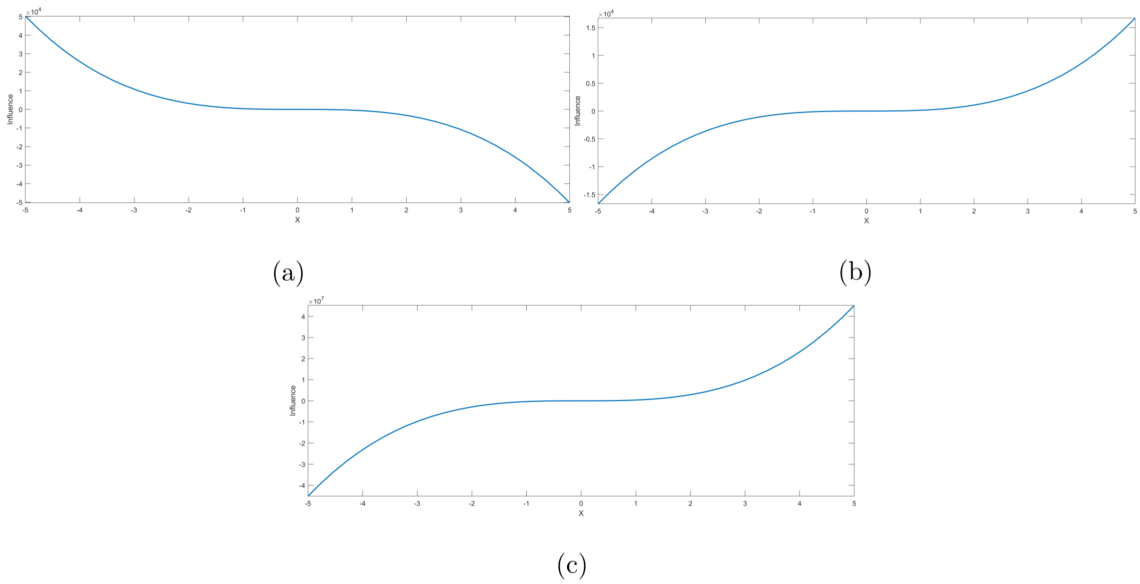


Figure 8. Influence Function of λ_0(a), λ_1(b), and λ_2(c) for Type IV distribution.