

Robust Test Statistics Based on Restricted Minimum Rényi's Pseudodistance Estimators

María Jaenada [†], Pedro Miranda [†] and Leandro Pardo ^{*,†}

Department of Statistics and Operation Research, Faculty of Mathematics, Interdisciplinary Mathematical Institute, Complutense University of Madrid, Plaza Ciencias, 3, 28040 Madrid, Spain; mjaenada@ucm.es (M.J.); pmiranda@mat.ucm.es (P.M.)

* Correspondence: lpardo@mat.ucm.es

† These authors contributed equally to this work.

Abstract: The Rao's score, Wald and likelihood ratio tests are the most common procedures for testing hypotheses in parametric models. None of the three test statistics is uniformly superior to the other two in relation with the power function, and moreover, they are first-order equivalent and asymptotically optimal. Conversely, these three classical tests present serious robustness problems, as they are based on the maximum likelihood estimator, which is highly non-robust. To overcome this drawback, some test statistics have been introduced in the literature based on robust estimators, such as robust generalized Wald-type and Rao-type tests based on minimum divergence estimators. In this paper, restricted minimum Rényi's pseudodistance estimators are defined, and their asymptotic distribution and influence function are derived. Further, robust Rao-type and divergence-based tests based on minimum Rényi's pseudodistance and restricted minimum Rényi's pseudodistance estimators are considered, and the asymptotic properties of the new families of tests statistics are obtained. Finally, the robustness of the proposed estimators and test statistics is empirically examined through a simulation study, and illustrative applications in real-life data are analyzed.

Keywords: Rényi's pseudodistance; minimum Rényi's pseudodistance estimators; restricted minimum Rényi's pseudodistance estimators; Rao-type tests; divergence-based tests



Citation: Jaenada, M.; Miranda, P.; Pardo, L. Robust Test Statistics Based on Restricted Minimum Rényi's Pseudodistance Estimators. *Entropy* **2022**, *24*, 616. <https://doi.org/10.3390/e24050616>

Academic Editors: Karagrigoriou Alexandros and Makrides Andreas

Received: 30 March 2022

Accepted: 26 April 2022

Published: 28 April 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Let $(\mathcal{X}, \beta_{\mathcal{X}}, P_{\theta})_{\theta \in \Theta}$ be the statistical space associated with the random variable X , where $\beta_{\mathcal{X}}$ is the σ -field of Borel subsets $A \subset \mathcal{X}$ and $\{P_{\theta}\}_{\theta \in \Theta}$ is a family of probability distributions defined on the measurable space $(\mathcal{X}, \beta_{\mathcal{X}})$, with Θ an open subset of \mathbb{R}^p and $p \geq 1$. We assume that the probability measures P_{θ} are described by densities $f_{\theta}(x) = dP_{\theta}/d\mu(x)$, where μ is a σ -finite measure on $(\mathcal{X}, \beta_{\mathcal{X}})$. Given a random sample X_1, \dots, X_n , of the random variable X with density belonging to the parametric family P_{θ} , the most popular estimator for the model parameter θ is the maximum likelihood estimator (MLE), which maximizes the likelihood function of the assumed model. The MLE has been widely studied in the literature for general statistical models, and it has been shown that, under certain regularity conditions, the sequence of MLEs of θ , $\hat{\theta}_n$, is asymptotically normal and it satisfies some desirable properties, such as consistency and asymptotic efficiency. That is, the MLE is the BAN (best asymptotically normal) estimator. However, in many popular statistical models, the MLE is markedly non-robust against deviations, even very small ones, from the parametric conditions.

To overcome the lack of robustness, minimum distance (or minimum divergence) estimators (MDEs) have been developed. MDEs have received growing attention in statistical inference because of their ability to conciliate efficiency and robustness. In parametric estimation, the role of divergence or distance measures is very intuitive: the estimates of the unknown parameters are obtained by minimizing a suitable divergence measure between the estimated from data and the assumed model distributions. There is a growing

body of literature that recognizes the importance of MDEs in terms of robustness, without a significant loss of efficiency, with respect to the MLE. See, for instance, the works of Beran [1], Tamura and Boes [2], Simpson [3,4], Lindsay [5], Pardo [6], and Basu et al. [7] and the references therein.

Let G denote the unknown distribution function, with associated density g , underlying the data. The minimum divergence (distance) functional evaluated at G , $T(G)$, is defined as

$$d(g, f_{T(G)}) = \min_{\theta \in \Theta} d(g, f_{\theta}), \quad (1)$$

with $d(g, f_{\theta})$ being a distance or divergence measure between the densities g and f_{θ} . As the true distribution underlying the data is unknown, given a random sample, we could estimate the model parameter θ , substituting in the previous expression the true distribution G by its empirical estimation G_n . Therefore, the MDE of θ is given by

$$\hat{\theta}_n = T(G_n), \quad (2)$$

When dealing with continuous models, it is convenient to consider families of divergence measures for which non-parametric estimators of the unknown density function are not needed. From this perspective, the density power divergence (DPD) family, leading to the minimum density power divergence estimators (MDPDEs) (see Basu et al. [7]), as well as the Rényi's pseudodistance (RP), leading to the minimum Rényi's pseudodistance estimators (MRPE) (see Broniatowski et al. [8]) between others, play an important role. The results presented in Broniatowski et al. [8] in the context of independent and identically distributed random variables were extended for the case of independent but not identically distributed random variables by Castilla et al. [9].

In many situations we have additional knowledge about the true parameter value, as it must satisfy certain constraints. Then, the restricted parameter space has the form

$$\{\theta \in \Theta / g(\theta) = \mathbf{0}_r\}, \quad (3)$$

where $\mathbf{0}_r$ denotes the null vector of dimension r , and $g : \mathbb{R}^p \rightarrow \mathbb{R}^r$ is a vector-valued function such that the $p \times r$ matrix

$$G(\theta) = \frac{\partial g^T(\theta)}{\partial \theta} \quad (4)$$

exists and is continuous in θ , and $\text{rank}(G(\theta)) = r$. Here, superscript T represents the transpose of the matrix. In the following, the restricted parameter space given in (3) is denoted by Θ_0 , as in most situations, it will represent a composite null hypothesis.

The most popular estimator of θ under the non-linear constraint given in (3) is the restricted MLE (RMLE) that maximizes the likelihood function subject to the constraint $g(\theta) = \mathbf{0}_r$ (see Silvey [10]). The RMLE encounters similar robustness problems to the MLE. To overcome such deficiency, the restricted MDPDEs (RMDPDEs) were introduced in Basu et al. [11] and their theoretical robustness properties were later studied in Ghosh [12].

The main purpose in this paper is extending the theory developed for the MRPE to the restricted parameter space setting, yielding to the restricted MRPE (RMPRE), where the parameter space has the form (3). The rest of the paper is as follows: In Section 2, MRPE is introduced. Section 3 presents RMPRE, and its asymptotic distribution as well as its influence function are obtained. In Section 4, two different test statistics for testing composite null hypothesis, based on the RMRPE, are developed, and explicit expressions of the statistics are presented for testing in normal populations. Section 5 presents a simulation study, where the robustness of the proposed estimators and test statistics is empirically shown. Section 6 deals with real-data situations. Finally, some conclusions are presented in Section 7.

2. Minimum Rényi Pseudodistance Estimators

In this section, we introduce the MRPE. We derive the estimating equations of the MRPE and recall its asymptotic distribution.

Let X_1, \dots, X_n be a random sample of size n from a population having true and unknown density function g , modeled by a parametric family of densities f_θ with $\theta \in \Theta \subset \mathbb{R}^p$. The RP between the densities f_θ and g is given, for $\tau > 0$, by

$$R_\tau(f_\theta, g) = \frac{1}{\tau + 1} \log \left(\int f_\theta(x)^{\tau+1} dx \right) + \frac{1}{\tau(\tau + 1)} \log \left(\int g(x)^{\tau+1} dx \right) - \frac{1}{\tau} \log \left(\int f_\theta(x)^\tau g(x) dx \right).$$

The RP can be defined for $\tau = 0$ taking continuous limits, yielding the expression

$$R_0(f_\theta, g) = \lim_{\tau \downarrow 0} R_\tau(f_\theta, g) = \int g(x) \log \frac{g(x)}{f_\theta(x)} dx.$$

Then, the RP coincides with the Kullback–Leibler divergence (KL) between g and f_θ , at $\tau = 0$ (see Pardo, 2006).

The RP was considered for the first time by Jones et al. [13]. Later Broniatowski et al. [8] established some useful properties of the divergence, such as the positivity of the RP for any two densities and for all values of the parameter τ , $R_\tau(f_\theta, g) \geq 0$ and uniqueness of the minimum RP within a parametric family, that is, $R_\tau(f_\theta, g) = 0$ if and only if $f_\theta = g$. The last property justifies the definition of the MRPEs as the minimizer of the RP between the assumed distribution and the empirical distribution of the data. It is interesting to note that the so-called RP by Broniatowski et al. [8] had been previously considered by Fujisawa and Eguchi [14] under the name of γ -cross entropy. In that paper, some appealing robustness properties of the estimators based on such entropy are shown.

Given a sample X_1, \dots, X_n , from Broniatowski et al. [8] it can be seen that minimizing $R_\tau(f_\theta, g)$ leads to the following definition.

Definition 1. Let $(\mathcal{X}, \beta_{\mathcal{X}}, f_\theta)_{\theta \in \Theta \subset \mathbb{R}^p}$ be a statistical space. The MRPE based on the random sample X_1, \dots, X_n for the unknown parameter θ is given, for $\tau > 0$, by

$$\hat{\theta}_\tau(X_1, \dots, X_n) = \arg \sup_{\theta \in \Theta} \frac{\sum_{i=1}^n f_\theta(X_i)^\tau}{C_\tau(\theta)}, \quad (5)$$

where

$$C_\tau(\theta) = \left(\int f_\theta(x)^{\tau+1} dx \right)^{\frac{\tau}{\tau+1}}.$$

Further, at $\tau = 0$, $\hat{\theta}_0(X_1, \dots, X_n)$ minimizes the KL divergence, and thus the MRPE coincides with the MLE for $\tau = 0$. Based on the previous definition (5), differentiating, we obtain that the estimating equations of the MRPE are given by

$$\sum_{i=1}^n \Psi_\tau(x_i; \theta) = \mathbf{0}_p, \quad (6)$$

with

$$\begin{aligned} \Psi_\tau(x; \theta) &= f_\theta(x)^\tau (\mathbf{u}_\theta(x) - \mathbf{c}_\tau(\theta)), \\ \mathbf{u}_\theta(x) &= (u_{\theta_1}(x), \dots, u_{\theta_p}(x))^T, \quad u_{\theta_i}(x) = \frac{\partial}{\partial \theta_i} \log f_\theta(x), \\ \frac{\partial C_\tau(\theta)}{\partial \theta} &= C_\tau(\theta) \mathbf{c}_\tau(\theta) \tau, \end{aligned} \quad (7)$$

being

$$\mathbf{c}_\tau(\boldsymbol{\theta}) = \frac{1}{\kappa_\tau(\boldsymbol{\theta})} \boldsymbol{\xi}_\tau(\boldsymbol{\theta}) = (c_{\tau,1}(\boldsymbol{\theta}), \dots, c_{\tau,p}(\boldsymbol{\theta}))^T, \quad (8)$$

$$\boldsymbol{\xi}_\tau(\boldsymbol{\theta}) = \int f_\theta(x)^{\tau+1} \mathbf{u}_\theta(x) dx, \quad (9)$$

$$\kappa_\tau(\boldsymbol{\theta}) = \int f_\theta(x)^{\tau+1} dx. \quad (10)$$

The MRPE is an M-estimator and thus its asymptotic distribution and influence function (IF) can be obtained based on the asymptotic theory of the M-estimators. Broniatowski et al. [8] studied the asymptotic properties and robustness of the MRPEs. The next result recalls the asymptotic distribution of the MRPEs.

Theorem 1. *Let $\boldsymbol{\theta}_0$ be the true unknown value of $\boldsymbol{\theta}$. Then,*

$$\sqrt{n}(\hat{\boldsymbol{\theta}}_\tau - \boldsymbol{\theta}_0) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}(\mathbf{0}_p, \mathbf{V}_\tau(\boldsymbol{\theta}_0)) \quad (11)$$

where

$$\mathbf{V}_\tau(\boldsymbol{\theta}) = \mathbf{S}_\tau(\boldsymbol{\theta})^{-1} \mathbf{K}_\tau(\boldsymbol{\theta}) \mathbf{S}_\tau(\boldsymbol{\theta})^{-1} \quad (12)$$

with

$$\mathbf{S}_\tau(\boldsymbol{\theta}) = -\mathbb{E} \left[\frac{\partial \boldsymbol{\Psi}_\tau(X; \boldsymbol{\theta})^T}{\partial \boldsymbol{\theta}} \right], \quad (13)$$

$$\mathbf{K}_\tau(\boldsymbol{\theta}) = \mathbb{E} \left[\boldsymbol{\Psi}_\tau(X; \boldsymbol{\theta}) \boldsymbol{\Psi}_\tau^T(X; \boldsymbol{\theta}) \right]. \quad (14)$$

Castilla et al. [15] introduced useful notation for the computation of $\mathbf{V}_\tau(\boldsymbol{\theta})$.

$$\mathbf{S}_\tau(\boldsymbol{\theta}) = \mathbf{J}_\tau(\boldsymbol{\theta}) - \frac{1}{\kappa_\tau(\boldsymbol{\theta})} \boldsymbol{\xi}_\tau(\boldsymbol{\theta}) \boldsymbol{\xi}_\tau(\boldsymbol{\theta})^T, \quad (15)$$

$$\mathbf{K}_\tau(\boldsymbol{\theta}) = \mathbf{J}_{2\tau}(\boldsymbol{\theta}) + \frac{1}{\kappa_\tau(\boldsymbol{\theta})} \left(\frac{\kappa_{2\tau}(\boldsymbol{\theta})}{\kappa_\tau(\boldsymbol{\theta})} \boldsymbol{\xi}_\tau(\boldsymbol{\theta}) \boldsymbol{\xi}_\tau(\boldsymbol{\theta})^T - \boldsymbol{\xi}_\tau(\boldsymbol{\theta}) \boldsymbol{\xi}_{2\tau}(\boldsymbol{\theta})^T - \boldsymbol{\xi}_{2\tau}(\boldsymbol{\theta}) \boldsymbol{\xi}_\tau(\boldsymbol{\theta})^T \right), \quad (16)$$

where

$$\mathbf{J}_\tau(\boldsymbol{\theta}) = \int f_\theta(x)^{\tau+1} \mathbf{u}_\theta(x) \mathbf{u}_\theta(x)^T dx, \quad (17)$$

and $\kappa_\tau(\boldsymbol{\theta})$ and $\boldsymbol{\xi}_\tau(\boldsymbol{\theta})$ are as in (9) and (10), respectively.

Toma and Leoni-Aubin [16] defined new robust and efficient measures based on the RP. Later, Toma et al. [17] considered the MRPE for general parametric models and developed a model selection criterion for regression models. Broniatowski et al. [8] applied the method to the multiple regression model (MRM) with random covariates. Subsequently, Castilla et al. [18] developed Wald-type tests based on MRPE for the MRM, and Castilla et al. [19] studied the MRPE for the MRM in the ultra-high dimensional set-up. Further, Jaenada and Pardo [20,21] considered the MRPE and Wald-type test statistics for generalized linear models (GLM). Despite Wald-type test statistics, there exist others relevant test statistics having an important role in the statistical literature: the likelihood-ratio and Rao (or score) tests, which are based on restricted estimators, usually the RMLE. Then, it makes sense to develop robust versions of these popular statistics based on the RMRPE.

3. The Restricted Minimum Rényi Pseudodistance Estimator: Asymptotic Distribution and Influence Function of RMRPE

In this section, we introduce the RMRPE and we derive its asymptotic distribution. Moreover, we study its robustness properties through its influence function (IF).

Definition 2. The RMRPE functional $\tilde{T}_\tau(G)$ evaluated at the distribution G is defined by

$$R_\tau(g, f_{\tilde{T}_\tau(G)}) = \min_{\theta \in \Theta_0} R_\tau(g, f_\theta),$$

given that such a minimum exists.

Accordingly, given random sample X_1, \dots, X_n from the distribution G , the RMRPE of θ is defined as

$$\tilde{\theta}_\tau = \arg \sup_{\theta \in \Theta_0} \sum_{i=1}^n \frac{f_\theta(X_i)^\tau}{C_\tau(\theta)}.$$

Next, the result states the asymptotic distribution of the RMRPE, $\tilde{\theta}_\tau = \tilde{T}_\tau(G)$.

Theorem 2. Suppose that the true distribution satisfies the conditions of the model and let us denote by $\theta_0 \in \Theta_0$ the true parameter. Then, the RMRPE $\tilde{\theta}_\tau$ of θ obtained under the constraints $g(\theta) = \mathbf{0}_r$ has distribution

$$n^{1/2}(\tilde{\theta}_\tau - \theta_0) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}(\mathbf{0}_p, \Sigma_\tau(\theta_0))$$

where

$$\begin{aligned} \Sigma_\tau(\theta_0) &= P_\tau^*(\theta_0)K_\tau(\theta_0)P_\tau^*(\theta_0)^T, \\ P_\tau^*(\theta_0) &= S_\tau(\theta_0)^{-1} - Q_\tau(\theta_0)G(\theta_0)^T S_\tau(\theta_0)^{-1}, \end{aligned} \tag{18}$$

$$Q_\tau(\theta_0) = S_\tau(\theta_0)^{-1}G(\theta_0) \left[G(\theta_0)^T S_\tau(\theta_0)^{-1}G(\theta_0) \right]^{-1}. \tag{19}$$

and $S_\tau(\theta_0)$ is defined in (13), evaluated at $\theta = \theta_0$.

Proof. See Appendix A. \square

To analyze the robustness of an estimator, Hampel et al. [22] introduced the concept of the influence function (IF). Since then, the IF has been widely used in statistical literature to measure robustness in different statistical contexts. Intuitively, the IF describes the effect of an infinitesimal contamination of the model on the estimate. Then, IFs associated to locally robust (B-robust) estimators should be bounded. Let us now obtain the IF of RMRPE and analyze its boundedness to assess the robustness of the proposed estimators. We consider the contaminated model $g_\varepsilon(x) = (1 - \varepsilon)f_\theta(x) + \varepsilon\Delta_x$, with Δ_x the indicator function in x , and we denote $\tilde{\theta}_{\tau,\varepsilon} = \tilde{T}_\tau(G_\varepsilon)$, being G_ε the distribution function associated to g_ε . By definition, $\tilde{\theta}_{\tau,\varepsilon}$ is the minimizer of $R_\tau(g, f_\theta)$ subject to $g(\tilde{\theta}_{\tau,\varepsilon}) = \mathbf{0}$. Following the same steps as in Theorem 5 in Broniatowski et al. [8], it can be seen that the influence function of \tilde{T}_τ in f_θ is given by

$$IF(x, \tilde{T}_\tau, \theta) = M_\tau(\theta)^{-1} [f_\theta(x)^\tau \mathbf{u}_\theta(x) - c_\tau(\theta) f_\theta(x)^\tau], \tag{20}$$

where $c_\tau(\theta)$ was defined in (8) and

$$\begin{aligned} M_\tau(\theta) &= \frac{1}{\int f_\theta(x)^{\tau+1} dx} \left[\int f_\theta(x)^{\tau+1} dx \int f_\theta(x)^{\tau+1} \mathbf{u}_\theta(x) \mathbf{u}_\theta(x)^T dx \right. \\ &\quad \left. - \left(\int f_\theta(x)^{\tau+1} \mathbf{u}_\theta(x) dx \right) \left(\int f_\theta(x)^{\tau+1} \mathbf{u}_\theta(x) dx \right)^T \right], \end{aligned}$$

with the additional condition that $g(\tilde{\theta}_{\tau,\varepsilon}) = \mathbf{0}$. Note that expression (20) corresponds to the IF of the unrestricted MRPE. Differentiating this last equation gives, at $\varepsilon = 0$,

$$G(\theta)^T IF(x, \tilde{T}_\tau, \theta) = 0. \tag{21}$$

Based on (20) and (21) we have

$$\begin{pmatrix} \mathbf{M}_\tau(\boldsymbol{\theta}) \\ \mathbf{G}(\boldsymbol{\theta})^T \end{pmatrix} IF(x, \tilde{T}_\tau, \boldsymbol{\theta}) = \begin{pmatrix} [f_\theta(x)^\tau \mathbf{u}_\theta(x) - \mathbf{c}_\tau(\boldsymbol{\theta}) f_\theta(x)^\tau] \\ \mathbf{0} \end{pmatrix}.$$

Therefore,

$$\begin{pmatrix} \mathbf{M}_\tau(\boldsymbol{\theta})^T & \mathbf{G}(\boldsymbol{\theta}) \end{pmatrix} \begin{pmatrix} \mathbf{M}_\tau(\boldsymbol{\theta}) \\ \mathbf{G}(\boldsymbol{\theta})^T \end{pmatrix} IF(x, \tilde{T}_\tau, \boldsymbol{\theta}) = \mathbf{M}_\tau(\boldsymbol{\theta})^T [f_\theta(x)^\tau \mathbf{u}_\theta(x) - \mathbf{c}_\tau(\boldsymbol{\theta}) f_\theta(x)^\tau]$$

and

$$IF(x, \tilde{T}_\tau, \boldsymbol{\theta}) = \left(\mathbf{M}_\tau(\boldsymbol{\theta})^T \mathbf{M}_\tau(\boldsymbol{\theta}) + \mathbf{G}(\boldsymbol{\theta}) \mathbf{G}(\boldsymbol{\theta})^T \right)^{-1} \mathbf{M}_\tau(\boldsymbol{\theta})^T [f_\theta(x)^\tau \mathbf{u}_\theta(x) - \mathbf{c}_\tau(\boldsymbol{\theta}) f_\theta(x)^\tau]. \tag{22}$$

Note that matrices $\mathbf{M}_\tau(\boldsymbol{\theta})$ and $\mathbf{G}(\boldsymbol{\theta})$ involved in the expression (22) are defined except for the model and tuning parameters $\boldsymbol{\theta}$ and τ , and so the boundedness of the IF of the RMRPE depends, therefore, on the boundedness of the factor

$$[f_\theta(x)^\tau \mathbf{u}_\theta(x) - \mathbf{c}_\tau(\boldsymbol{\theta}) f_\theta(x)^\tau].$$

Therefore, the boundedness of the IF of the RMRPE depends directly on the boundedness of IF of the MRPE, stated in (20). The IF of the MRPE has been widely studied for general statistical models, concluding that the MRPEs are robust for positive values of τ , and that such robustness increases with the tuning parameter. A whole discussion can be found in the work of Broniatowski et al. [8]. Hence, the same properties hold for RMRPEs.

4. Robust Test Statistics Based on RMRPEs

In this section, we develop two statistics based on the RMRPEs for testing composite null hypothesis, and their asymptotic distributions are obtained. Both procedures are particularized to standard deviation testing (with unknown mean) under normal populations, and explicit expressions of the test statistics are obtained.

4.1. Testing Based on Divergence Measures

In this section, we present the family of Rényi’s pseudodistance test statistics (RPTS) for testing the null hypothesis given in (3). This family of test statistics is given by

$$T_\gamma(\hat{\boldsymbol{\theta}}_\tau, \tilde{\boldsymbol{\theta}}_\tau) = 2nR_\gamma(f_{\hat{\boldsymbol{\theta}}_\tau}, f_{\tilde{\boldsymbol{\theta}}_\tau}). \tag{23}$$

The RPTS, $T_\gamma(\hat{\boldsymbol{\theta}}_\tau, \tilde{\boldsymbol{\theta}}_\tau)$, can be understood as a measure between the best unrestricted estimator of the model parameter, and the best estimator satisfying the null hypothesis. Large values of the RPTS indicate that the model densities associated with the restricted and unrestricted estimators are far away one from the other, and so the null hypothesis is not supported by the observed data. Hence, we should reject H_0 for large enough $T_\gamma(\hat{\boldsymbol{\theta}}_\tau, \tilde{\boldsymbol{\theta}}_\tau)$. We can observe that the family of RPTS defined in (23) depends on two tuning parameters, τ and γ . The first is used for estimating the unknown parameters, while the second is applied to obtain the family of test statistics. The following theorem presents the asymptotic distribution of the family of RPTS defined in (23).

Theorem 3. *The asymptotic distribution of $T_\gamma(\hat{\boldsymbol{\theta}}_\tau, \tilde{\boldsymbol{\theta}}_\tau)$ defined in (23) coincides, under the null hypothesis H_0 given in (3), with the distribution of the random variable*

$$\sum_{i=1}^r \lambda_i^{\tau, \gamma}(\boldsymbol{\theta}_0) Z_i^2,$$

where Z_1, \dots, Z_r are independent standard normal variables, $\lambda_1^{\tau, \gamma}(\boldsymbol{\theta}_0), \dots, \lambda_r^{\tau, \gamma}(\boldsymbol{\theta}_0)$ are the nonzero eigenvalues of $\mathbf{M}_{\gamma, \tau}(\boldsymbol{\theta}_0) = \mathbf{A}_\gamma(\boldsymbol{\theta}_0) \mathbf{B}_\tau(\boldsymbol{\theta}_0) \mathbf{K}_\tau(\boldsymbol{\theta}_0) \mathbf{B}_\tau(\boldsymbol{\theta}_0)$ and $k = r$. The matrices $\mathbf{A}_\gamma(\boldsymbol{\theta}_0)$ and $\mathbf{B}_\tau(\boldsymbol{\theta}_0)$ are given by,

$$A_\gamma(\theta_0) = \frac{S_\gamma(\theta_0)}{\kappa_\tau(\theta_0)}, \tag{24}$$

$$B_\tau(\theta_0) = Q_\tau(\theta_0)G(\theta_0)^T S_\tau(\theta_0)^{-1}. \tag{25}$$

Proof. See Appendix A. □

Rényi’s Pseudodistance Test Statistics for Normal Populations

Under the $\mathcal{N}(\mu, \sigma^2)$ model, consider the problem of testing

$$H_0 : \sigma = \sigma_0 \text{ versus } H_1 : \sigma \neq \sigma_0 \tag{26}$$

where μ is an unknown nuisance parameter. In this case, the unrestricted and null parameter spaces are given by $\Theta = \{(\mu, \sigma^2) \in \mathbb{R}^2 | \mu \in \mathbb{R}, \sigma^2 \in \mathbb{R}^+\}$ and $\Theta_0 = \{(\mu, \sigma) \in \mathbb{R}^2 | \sigma = \sigma_0, \mu \in \mathbb{R}\}$, respectively. If we consider the function $g(\theta) = \sigma - \sigma_0$, with $\theta = (\mu, \sigma)^T$, the null hypothesis H_0 can be written as

$$H_0 : g(\theta) = 0$$

and we are in the situation considered in (26). We can observe that in our case $G(\theta) = (0, 1)^T$. Based on (6) and taking into account the fact that $f_\theta(x)$ is the normal density with mean μ and variance σ^2 , the MRPE $\hat{\theta}_\tau = (\hat{\mu}_\tau, \hat{\sigma}_\tau)^T$ of $\theta = (\mu, \sigma)^T$ is the solution of the system of nonlinear equations

$$\begin{cases} \sum_{i=1}^n (X_i - \mu) \exp\left\{-\frac{\tau}{2}\left(\frac{X_i - \mu}{\sigma}\right)^2\right\} = 0 \\ \sum_{i=1}^n \left\{\left(\frac{X_i - \mu}{\sigma}\right)^2 - \frac{1}{1+\tau}\right\} \exp\left\{-\frac{\tau}{2}\left(\frac{X_i - \mu}{\sigma}\right)^2\right\} = 0 \end{cases}$$

while the RMRPE $\tilde{\theta}_\beta = (\tilde{\mu}_\tau, \sigma_0)^T$, when $\sigma = \sigma_0$ is the solution of the nonlinear equation

$$\sum_{i=1}^n \left\{\left(\frac{X_i - \mu}{\sigma_0}\right)^2 - \frac{1}{1+\tau}\right\} \exp\left\{-\frac{\tau}{2}\left(\frac{X_i - \mu}{\sigma_0}\right)^2\right\} = 0.$$

After some algebra (see the Appendix A) we obtain that the RPTS for testing (26) under normal populations can be expressed as

$$\begin{aligned} T_\gamma(\hat{\theta}_\tau, \tilde{\theta}_\tau) &= 2nR_\gamma\left(\mathcal{N}(\hat{\mu}_\tau, \hat{\sigma}_\tau^2), \mathcal{N}(\tilde{\mu}_\tau, \sigma_0)\right) \tag{27} \\ &= \frac{2n}{\gamma(\gamma+1)} \log \left[\frac{1}{\hat{\sigma}_\tau \sigma_0^\gamma} \left(\frac{\sqrt{\hat{\sigma}_\tau^2 + \gamma \sigma_0^2}}{\sqrt{\gamma+1}} \right)^{\gamma+1} \right] + n \frac{(\hat{\mu}_\tau - \tilde{\mu}_\tau)^2}{(\gamma \sigma_0^2 + \hat{\sigma}_\tau^2)} \end{aligned}$$

Based in (27), and taking into account that the eigenvalue of the matrix $A_\gamma(\theta)B_\tau(\theta)K_\tau(\theta)B_\tau(\theta)$ is given by (see Appendix A)

$$l_{\tau,\gamma}(\sigma) = \frac{1}{2} \frac{(\tau+1)^3}{(\gamma+1)^2(2\tau+1)^{\frac{5}{2}}} (3\tau^2 + 4\tau + 2),$$

we apply Theorem 3 such that

$$l_{\tau,\gamma}(\sigma_0)^{-1} \left(\frac{2n}{\gamma(\gamma+1)} \log \left[\frac{1}{\hat{\sigma}_\tau \sigma_0^\gamma} \left(\frac{\sqrt{\hat{\sigma}_\tau^2 + \gamma \sigma_0^2}}{\sqrt{\gamma+1}} \right)^{\gamma+1} \right] + n \frac{(\hat{\mu}_\tau - \tilde{\mu}_\tau)^2}{(\gamma \sigma_0^2 + \hat{\sigma}_\tau^2)} \right) \xrightarrow[n \rightarrow \infty]{L} \chi_1^2.$$

Note that the RPTS is indexed by two tuning parameters, γ and τ , the first controlling the robustness of the pseudodistance and the second controlling the robustness on the estimation. For simplicity, we use $\gamma = \tau$ for the normal population application.

Remark 1. For $\tau = \gamma = 0$, the RPTS coincides with the asymptotic likelihood ratio test for testing (26). Indeed, for $\tau = 0$, we have that the MLE and RMLE are given, respectively, by

$$\hat{\theta} = (\bar{X}, \hat{\sigma}_n^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2) \text{ and } \tilde{\theta} = (\bar{X}, \sigma_0^2).$$

Now, the expression of the Kullback–Leibler divergence (the RP for $\gamma = 0$) between two normal densities, $\mathcal{N}(\mu_1, \sigma_1)$ and $\mathcal{N}(\mu_2, \sigma_2)$, is given by

$$\lim_{\gamma \rightarrow 0} R_\gamma(\mathcal{N}(\mu_1, \sigma_1), \mathcal{N}(\mu_2, \sigma_2)) = \frac{\sigma_2^2 - \sigma_1^2}{2\sigma_1^2} + \ln \frac{\sigma_1}{\sigma_2} + \frac{1}{2} \frac{(\mu_1 - \mu_2)^2}{\sigma_1^2}. \tag{28}$$

and thus the RPTS for $\gamma = \tau = 0$ is

$$T_0(\hat{\theta}, \tilde{\theta}) = n \frac{\sigma_0^2}{\hat{\sigma}_n^2} - n + 2n \ln \frac{\hat{\sigma}_n}{\sigma_0}.$$

On the other hand, the likelihood ratio for testing (26) is given by

$$\lambda(X_1, \dots, X_n) = \left(\frac{\hat{\sigma}_n}{\sigma_0}\right)^{n/2} e^{-n \frac{\hat{\sigma}_n^2}{2\sigma_0^2}} e^{n/2},$$

and so, both expressions are related through

$$-2 \ln \lambda(X_1, \dots, X_n) = T_0(\hat{\theta}, \tilde{\theta}).$$

4.2. Rao’s-Type Tests Based on RMRPE

Rao test statistics are one of the most popular score test statistics for testing a simple and composite null hypothesis in general statistical models. For the simple null hypothesis testing, it requires no parameter estimation, but for composite ones, the classical Rao test is based on the likelihood score function associated with the restricted MLE (see Rao [23]). Basu et al. [24] generalized Rao’s procedure by using score functions associated with RMDPDEs, bringing in a considerable gain of robustness of the Rao-type test obtained. In this section, we develop Rao-type test statistics based on the score function associated to RMRPEs.

Let us consider the τ –score function associated to the RMRPE,

$$\psi_\tau(x; \theta) = f_\theta(x)^\tau (\mathbf{u}_\theta(x) - \mathbf{c}_\tau(\theta)),$$

so the estimating equations for the MRPE are given by

$$\sum_{i=1}^n \psi_\tau(x_i; \theta) = \mathbf{0}_p.$$

Then, the τ -score statistic can be defined as

$$\Psi_\tau(\theta) = \sum_{i=1}^n \psi_\tau(x_i; \theta) = \left(\sum_{i=1}^n \psi_\tau^1(x_i; \theta), \dots, \sum_{i=1}^n \psi_\tau^k(x_i; \theta) \right)^T.$$

However, taking expectations in the corresponding quantities, it is not difficult to show that

$$E \left[\left(\frac{\tau}{C_\tau(\boldsymbol{\theta})} f_{\boldsymbol{\theta}}(X)^\tau (\mathbf{u}_{\boldsymbol{\theta}}(X) - \mathbf{c}_\tau(\boldsymbol{\theta})) \right)_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} \right] = \mathbf{0}_p$$

$$E \left[\left(f_{\boldsymbol{\theta}}(X)^{2\tau} (\mathbf{u}_{\boldsymbol{\theta}}(X) - \mathbf{c}_\tau(\boldsymbol{\theta})) (\mathbf{u}_{\boldsymbol{\theta}}(X) - \mathbf{c}_\tau(\boldsymbol{\theta}))^T \right)_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} \right] = \mathbf{K}_\tau(\boldsymbol{\theta}_0),$$

where $\mathbf{K}_\tau(\boldsymbol{\theta})$ is defined in (16), and so, by the central limit theorem, the τ -score statistic is asymptotically normal,

$$\frac{1}{\sqrt{n}} \boldsymbol{\Psi}_\tau(\boldsymbol{\theta}) \xrightarrow[n \rightarrow \infty]{L} \mathcal{N}(\mathbf{0}_p, \mathbf{K}_\tau(\boldsymbol{\theta})). \tag{29}$$

The previous convergence motivates the definition of the Rao-type test statistics.

4.2.1. Rao-Type Test Statistics for Testing Simple Null Hypothesis

We first consider the simple null hypothesis test

$$H_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0 \text{ vs. } H_1 : \boldsymbol{\theta} \neq \boldsymbol{\theta}_0. \tag{30}$$

Then, the Rao-type test statistics $R_\tau(\boldsymbol{\theta}_0)$ for testing (30) is defined as

$$R_\tau(\boldsymbol{\theta}_0) = \frac{1}{n} \boldsymbol{\Psi}_\tau(\boldsymbol{\theta}_0)^T \mathbf{K}_\tau(\boldsymbol{\theta}_0)^{-1} \boldsymbol{\Psi}_\tau(\boldsymbol{\theta}_0).$$

Note that here the last test statistics depend on τ through the matrices $\boldsymbol{\Psi}_\tau(\boldsymbol{\theta}_0)$ and $\mathbf{K}_\tau(\boldsymbol{\theta}_0)$ involved in the definition, and again, the robustness of the statistics increases with τ . Moreover, the last matrix may have an explicit expression for certain statistical models, but otherwise it would have to be estimated from the sample.

Further, from (29), we have that, under the null hypothesis,

$$R_\tau(\boldsymbol{\theta}_0) \xrightarrow[n \rightarrow \infty]{L} \chi_p^2$$

with p being the dimension of the parameter space. Then, the null hypothesis is rejected if $R_\tau(\boldsymbol{\theta}_0) > \chi_{p,\alpha}^2$, where $\chi_{p,\alpha}^2$ denotes the upper α -quantile of a chi-square distribution with p degrees of freedom.

4.2.2. Rao-Type Test Statistics for Testing Composite Null Hypothesis

Next, let us consider composite null hypothesis of the form

$$H_0 : \mathbf{g}(\boldsymbol{\theta}) = \mathbf{0}_r \text{ vs. } H_1 : \mathbf{g}(\boldsymbol{\theta}) \neq \mathbf{0}_r, \tag{31}$$

where the function $\mathbf{g} : \mathbb{R}^p \rightarrow \mathbb{R}^r$ is a differentiable vector-valued function. Then, any vector $\boldsymbol{\theta}$ satisfying the null hypothesis belongs to a restricted parameter space given in (3). The generalized Rao-type test statistic associated to the RMRPE with tuning parameter τ , $\tilde{\boldsymbol{\theta}}_\tau$, for testing (31) is given by

$$R_\tau(\tilde{\boldsymbol{\theta}}_\tau) = \frac{1}{n} \boldsymbol{\Psi}_\tau(\tilde{\boldsymbol{\theta}}_\tau)^T \mathbf{Q}_\tau(\tilde{\boldsymbol{\theta}}_\tau) \left[\mathbf{Q}_\tau(\tilde{\boldsymbol{\theta}}_\tau)^T \mathbf{K}_\tau(\tilde{\boldsymbol{\theta}}_\tau) \mathbf{Q}_\tau(\tilde{\boldsymbol{\theta}}_\tau) \right]^{-1} \mathbf{Q}_\tau(\tilde{\boldsymbol{\theta}}_\tau)^T \boldsymbol{\Psi}_\tau(\tilde{\boldsymbol{\theta}}_\tau). \tag{32}$$

Using similar arguments to Basu et al. [24], it is possible to show that, under general regularity conditions, the Rao-type test statistics $R_\tau(\tilde{\boldsymbol{\theta}}_\tau)$ have an asymptotic chi-square distribution with r degrees of freedom under the null hypothesis given in (31). Therefore, the rejection region of the test is given by

$$\{X_1, \dots, X_n : R_\tau(\tilde{\boldsymbol{\theta}}_\tau) > \chi_{r,\alpha}^2\}.$$

Again, the tuning parameter τ controls the trade-off between efficiency and robustness of the test. Indeed, for $\tau = 0$, the generalized Rao type test statistic $R_{\tau=0}(\tilde{\boldsymbol{\theta}}_0)$ coincides with the classical Rao test for composite null hypothesis.

4.2.3. Rao Test for Normal Populations

Consider the test defined in (26) for testing the standard deviation value of a normal population with unknown mean. The explicit expression of the main matrices involved in the definition (32) for such testing procedure and assumed parametric model is given by

$$\begin{aligned}\psi_{\tau}(X; (\mu, \sigma)) &= \left(\frac{X - \mu}{\sigma^2} \frac{1}{(\sigma\sqrt{2\pi})^{\tau}} e^{-\frac{\tau}{2} \left(\frac{X - \mu}{\sigma}\right)^2}, \left(\left(\frac{X - \mu}{\sigma}\right)^2 - \frac{1}{1 + \tau} \right) \frac{1}{\sigma} \frac{1}{(\sigma\sqrt{2\pi})^{\tau}} e^{-\frac{\tau}{2} \left(\frac{X - \mu}{\sigma}\right)^2} \right)^T, \\ \mathbf{K}_{\tau}((\mu, \sigma)) &= \frac{1}{\sigma^2} \frac{1}{(\sigma\sqrt{2\pi})^{2\tau}} \frac{1}{(1 + 2\tau)^{3/2}} \begin{pmatrix} 1 & 0 \\ 0 & \frac{3\tau^2 + 2 + 4\tau}{(1 + \tau)^2(1 + 2\tau)} \end{pmatrix}, \\ \mathbf{Q}_{\tau}((\mu, \sigma)) &= \begin{pmatrix} 0 \\ 1 \end{pmatrix}.\end{aligned}$$

The step-by-step calculation of such values are detailed in the Appendix A. Then, the Rao-type test for composite null hypothesis of the form (31) is given by

$$R_{\tau}(\tilde{\mu}) = \frac{1}{n} \frac{(1 + 2\tau)^{3/2}(1 + \tau)^2(1 + 2\tau)}{3\tau^2 + 4\tau + 2} \left[\sum_{i=1}^n \left(\left(\frac{x_i - \tilde{\mu}}{\sigma_0} \right)^2 - \frac{1}{\tau + 1} \right) e^{-\frac{\tau}{2} \left(\frac{x_i - \tilde{\mu}}{\sigma_0} \right)^2} \right]^2$$

where $(\tilde{\mu}_{\tau}, \sigma_0)$ denotes the RMRPE with tuning parameter τ . Note that, for $\tau = 0$, $\tilde{\mu}_{\tau=0} = \bar{X}$. Then, the Rao-type test statistic based on RMRPE with $\tau = 0$ (the restricted MLE) coincides with the classical Rao test.

5. Simulation Study: Application to Normal Populations

In this section, we empirically analyze the performance of the proposed estimators under the normal parametric model and RPTS and Rao-type test statistics for the problem of testing (26) in terms of efficiency and robustness. We examine the accuracy of the RMRPEs, and we further examine the robustness properties of both families of estimators under different contamination scenarios. Further, we investigate the empirical level and power of the proposed test statistics under different sample sizes and contamination scenarios.

Let us consider a univariate normal model with true parameter value $\theta_0 = (\mu = 0, \sigma = 1)$, and the problem of testing

$$H_0 : \sigma = 1 \text{ vs. } H_1 : \sigma \neq 1. \quad (33)$$

The restricted parameter space is then given by

$$\Theta_0 = \{(\mu, 1) : \mu \in \mathbb{R}\}.$$

In order to evaluate the robustness properties of the estimators and test statistics, we introduce contamination in data by replacing a $\varepsilon\%$ of the observations by a contaminated sample, where ε denotes the contamination level. We generate five different scenarios of contamination:

- Pure data.
- Scenario 1: Slightly contaminated data. We replace a $\varepsilon\%$ of the samples by a contaminated sample from a normal distribution, $\mathcal{N}(0, \sqrt{3})$.
- Scenario 2: Heavily contaminated data. We replace a $\varepsilon\%$ of the samples by a contaminated sample from a normal distribution, $\mathcal{N}(0, \sqrt{5})$.

Further, in order to evaluate the power of the test, we consider an alternative true parameter value $\theta_1 = (0, 0.7)$ which does not satisfy the null hypothesis (33) (or equivalently the restrictions of the parameter space). In this scenario, contaminated parameters are set $\theta_1 = (0, 1.2)$ for slightly and $\theta_1 = (0, 1.5)$ for heavily contamination.

Figure 1 shows the root mean square error (RMSE) of the RMRPE of the scale parameter σ , for different values of the tuning parameter $\tau = 0, 0.2, 0.4, 0.6$ and $\tau = 0.8$ over $R = 10,000$ replications. As expected, large values of the tuning parameter produce more robust estimators, which is particularly advantageous for the heavily contaminated scenario. Furthermore, even when introducing very low levels of contamination in data, $\varepsilon = 5\%$, the RMRPE with moderate value of the tuning parameter outperforms the classical M.F. without a significant loss of efficiency in the absence of contamination.

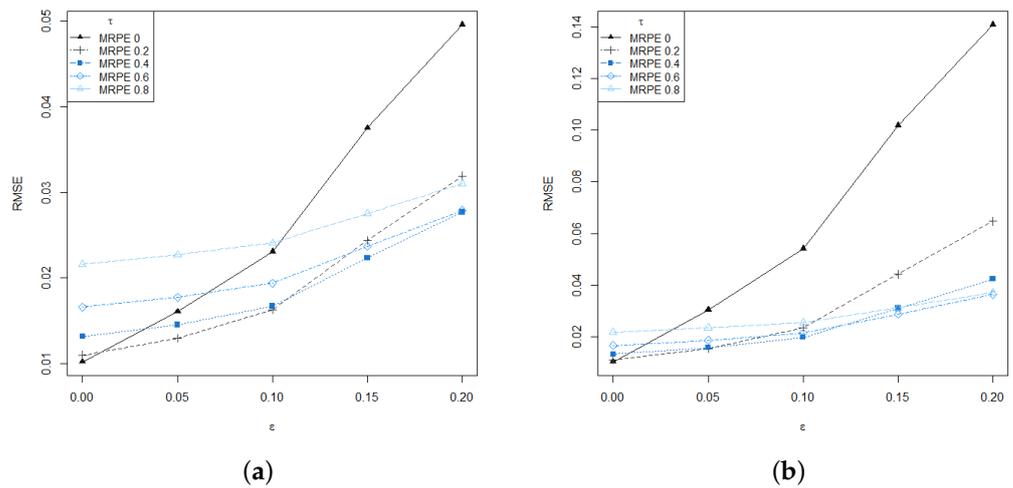


Figure 1. RMSE of the RMRPE under increasing contamination levels (slightly contaminated at left and heavily contaminated at right) for different values of the tuning parameter τ over $R = 10,000$ replications. (a) Scenario 1, (b) Scenario 2.

On the other hand, Figure 2 presents the empirical level and power of both RPTS and Rao-type test statistics based on RMRPEs for different values of the tuning parameter, $\tau = 0, 0.2, 0.4, 0.6, 0.8$, under increasing contamination levels. The empirical level and power are computed as the mean number of rejections over $R = 10,000$ replications. The empirical level produced by the classical ratio and Rao-type tests rapidly increases and separates from levels obtained with any robust test. Regarding the empirical power, all robust tests with moderate and large values of the tuning parameter outperform the classical estimator within their family under contaminated scenarios, but Rao-type test statistics based on RMRPEs are more conservative than RPTSs, thus exhibiting lower levels and powers. Then, the proposed test statistics provides an appealing alternative to classical likelihood ratio and Rao tests, with a small loss of efficiency in favor of a clear gain in terms of robustness.

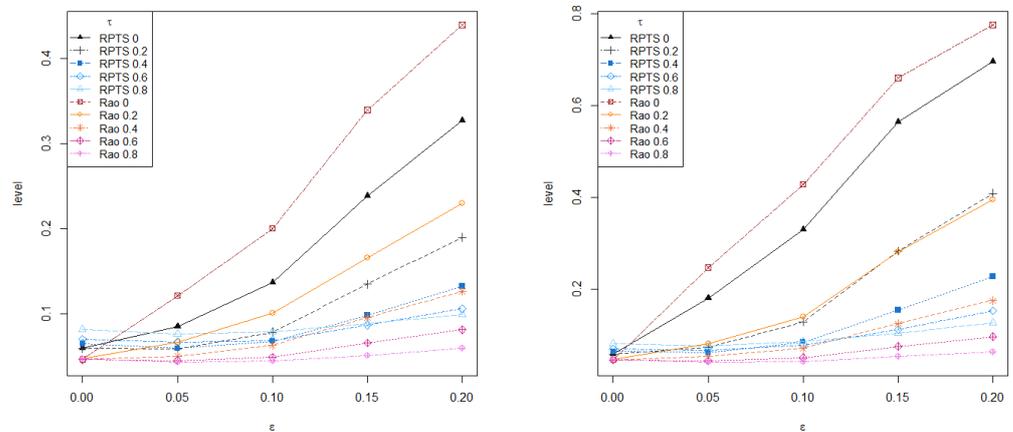


Figure 2. Cont.

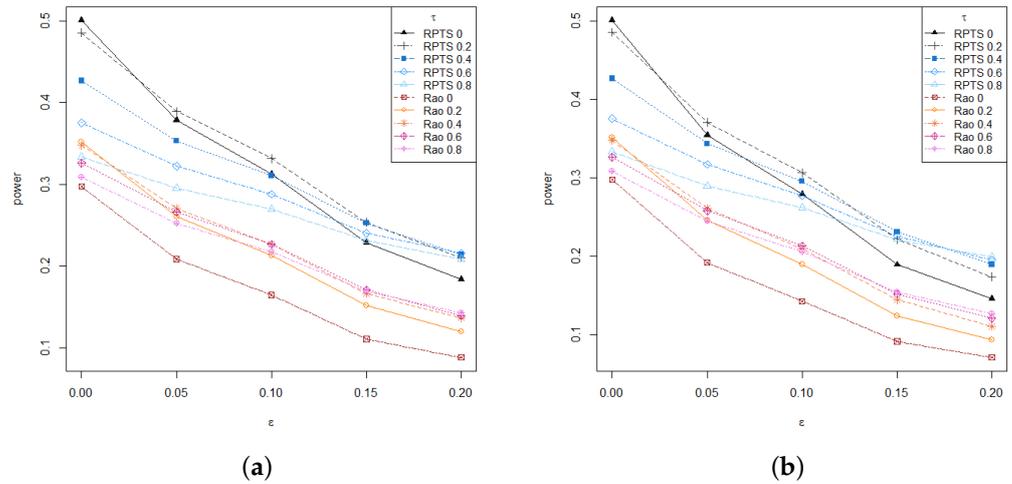


Figure 2. Empirical level and power under increasing contamination (slightly contaminated at left and heavily contaminated at right) over $R = 10,000$ repetitions. (a) Scenario 1, (b) Scenario 2.

On the other hand, the sample size could play a crucial role in the performance of the tests, even more accentuated when there exists data contamination. Figure 3 shows the sample size effect on the performance of the tests in terms of empirical level, under a 10% of contamination level in data. As discussed, Rao-type test statistics based on RMRPEs is more conservative and so tests based on RMRPEs with positive values of the tuning parameter produce lower empirical levels. Here, it outperforms the poor performance of the classical Rao-type test statistics with respect to any other. Moreover, when the sample size increases, the performance gap between non-robust and robust methods is widening.

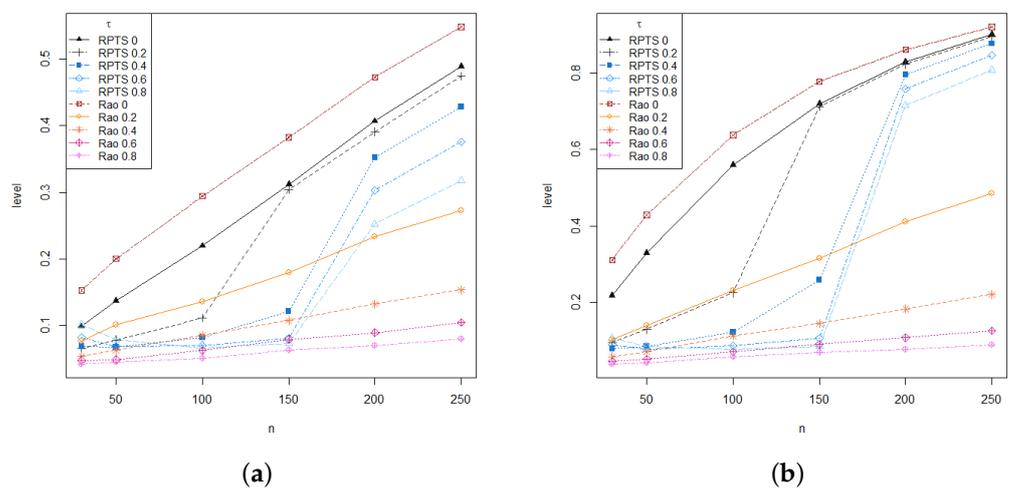


Figure 3. Empirical level under increasing sample sizes for 10% of contamination level (slightly contaminated at left and heavily contaminated at right) over $R = 10,000$ repetitions. (a) Scenario 1, (b) Scenario 2.

Following the discussions in the preceding sections, larger values of the tuning parameter produce more robust but less efficient estimators. Therefore, the optimal value of τ should obtain the best trade-off between efficiency and robustness. Warwick and Jones [25] first introduced a useful data-based procedure for the choice of the tuning parameter for the MDPDE based on minimizing the asymptotic MSE of the estimator. However, this method depends on the choice of a pilot estimator, and Basak et al. [26] improved the method by removing the dependency on an initial estimator. The proposed algorithm was developed

ad hoc for the MDPDE, but it can be easily adapted to the MRPE and RMRPE by simply substituting the expression of the variance of the MDPDE by the variance of the MRPPE or the RMRPE, respectively.

6. Real Data Application

Finally, we illustrate the outperformance of the proposed test statistics in two real data applications, where the gathered information contains some outlying observations. Both real dataset are modeled under the normal model, and hypothesis tests on the standard deviation of the population are performed.

6.1. Telephone-Fault Data

We consider the data on telephone line faults presented and analyzed by Welch [27] and Simpson [4]. The dataset consist of $n = 14$ ordered differences between the inverse test rates and the inverse control rates in matched pairs of areas,

$$-988, -135, -78, 3, 59, 83, 93, 110, 189, 197, 204, 229, 289, 310.$$

Basu et al. ([24,28]) modeled these differences as a normal random variable and pointed out that the first observation is a clear outlier, as its value is distant from the rest of the data. They tested simple and composite null hypotheses for the mean under the normal model, as well as a simple null hypothesis assuming a known mean. Here, we propose to test for the standard deviation of the normal distribution. Note that, computing the MLE of the sample with full and clean data (after removing the outlying observation), we obtain $(\hat{\mu}, \hat{\sigma}) = (40.36, 323.08)$, and $(\hat{\mu}, \hat{\sigma}) = (119.46, 134.82)$, respectively. Accordingly, the outlier clearly influences the model parameter estimates, playing a crucial role on the rejection of any null hypothesis. We consider the composite null hypothesis

$$H_0 : \sigma = 135 \text{ vs. } H_1 : \sigma \neq 135, \tag{34}$$

where the value $\sigma = 135$ has been chosen according to the estimation with clean data.

Figure 4 presents the RPTS (top) and Rao (bottom) test statistics (left) and p -values (right) for the telephone data against increasing tuning parameters. While it is clearly seen that both classical tests fail to not reject the null hypothesis when fitting the model with the original data, the decision turns around sharply as the tuning parameter τ crosses and goes beyond 0.2 for the RPTS and 0.15 for Rao-type test statistics based on MRPEs. On the other hand, the decision of not rejecting is agreed by all statistics when fitting the model with clean data. This example illustrates the great applicability of the robust methods, which are not too affected by a such outlying observation, and the good performance of the proposed statistics under contaminated observations, which stay stable.

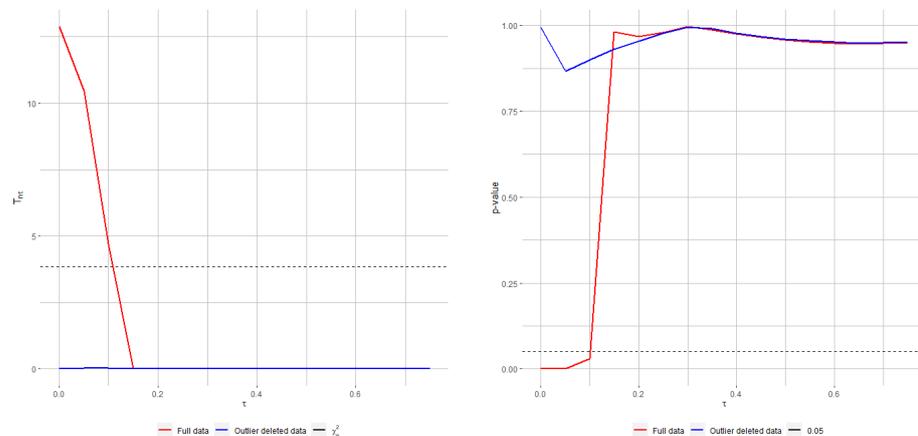


Figure 4. Cont.

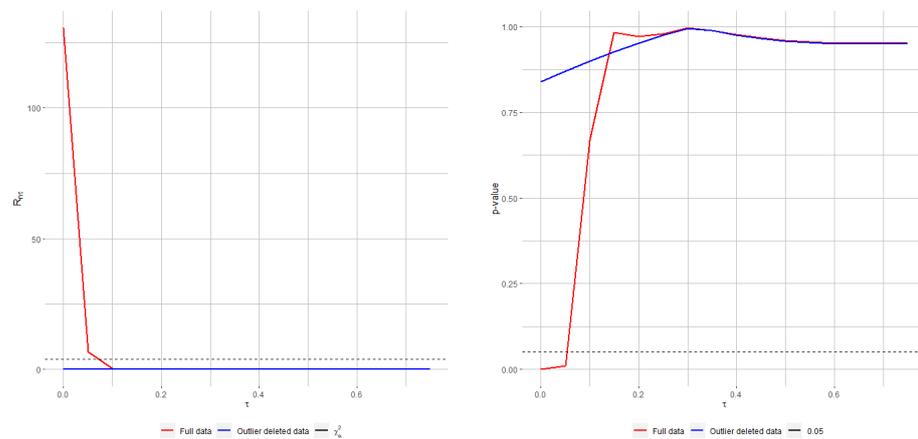


Figure 4. RPTS (top) and Rao-type test statistics (bottom), jointly with their associated p -values (right), for testing (34) with original and cleaned (after outliers removal) telephone-fault data.

6.2. Darwin’s Plant Fertilization Data

Darwin [29] performed an experiment to determine whether self-fertilized plants and cross-fertilized plants have different growth rates. He sowed in pots pairs of *Zea mays* plants, one self-fertilized and the other cross-fertilized, and after a specific time period, the height of each plant was measured. A particular sample of $n = 15$ pairs of plants led to the following paired differences (cross-fertilized minus self-fertilized).

$$-67, -48, 6, 8, 14, 16, 23, 24, 28, 29, 41, 49, 56, 60, 75$$

A parametric approach to analyze the data as a random sample from a normal distribution with unknown mean and standard deviation was developed by Basu et al. [24]. Here, there is not any huge outlying observation, but the first two observations seem to be distant from the rest of the sample, influencing the model parameter estimates and test decisions. Indeed, the MLE, computing with original data, is $(\hat{\mu}, \hat{\sigma}) = (20.93, 37.74)$, while the MLE, when removing the two first observations, switches to $(\hat{\mu}, \hat{\sigma}) = (33, 21.54)$. Therefore, removing influential observations may alter the decision of a test. According to these results, we consider the testing problem

$$H_0 : \sigma = 23 \text{ vs. } H_1 : \sigma \neq 23. \tag{35}$$

Figure 5 shows the test statistics (left) and corresponding p -values (right) for the two families of statistics considered, the RPTS (top) and Rao-type test statistics (bottom) against the tuning parameter value τ . Again, test statistics based on RMRPE with large enough tuning parameters do not reject the null hypothesis, unlike tests based on low values of $\tau = 0$, including the RMLE. The disagreement departs when using the clean data, as all tests agree on not rejecting the null hypothesis.

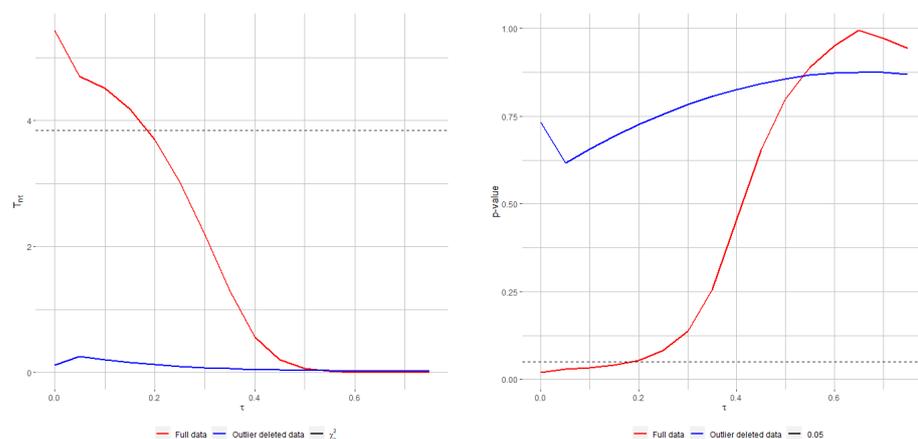


Figure 5. Cont.

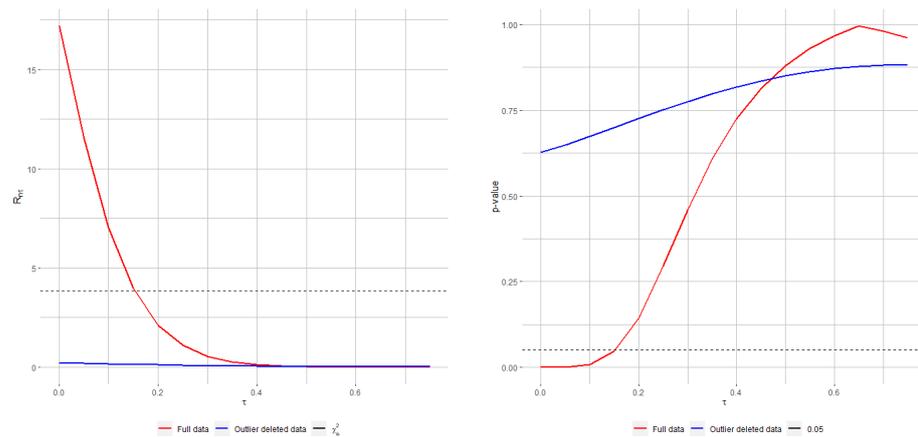


Figure 5. RPTS (top) and Rao-type test statistics (bottom), jointly with their associated p -values (right), for testing (35) with original and cleaned (after outliers removal) Darwing data.

7. Concluding Remarks

In this paper, we presented for the first time the family of RMRPEs. We derived their asymptotic distribution, and proved some suitable properties as consistency under the parameter restriction and robustness against data contamination. Further, based on these RMRPEs, we generalized two important families of statistics, namely RPTS and Rao-type tests, for testing a composite null hypothesis. Moreover, we obtained some explicit expressions of the RMRPEs, RPTS and Rao-type test statistics for testing the variance under a normal population with an unknown mean. It was empirically shown that the proposed RPTS and Rao-type test statistics are robust, unlike classical tests based on the MLE, under normal populations. Indeed, the robustness of the tests is controlled by a tuning parameter τ , and so larger values of τ produce more robust estimators (although less efficient). Finally, some classical numerical examples illustrate the theoretical properties and applicability of the proposed methods.

Author Contributions: Conceptualization, M.J., P.M. and L.P.; methodology, M.J., P.M. and L.P.; software, M.J., P.M. and L.P.; validation, M.J., P.M. and L.P.; formal analysis, M.J., P.M. and L.P.; investigation, M.J., P.M. and L.P.; resources, M.J., P.M. and L.P.; data curation, M.J., P.M. and L.P.; writing—original draft preparation, M.J., P.M. and L.P.; writing—review and editing, M.J., P.M. and L.P.; visualization, M.J., P.M. and L.P.; supervision, M.J., P.M. and L.P.; project administration, M.J., P.M. and L.P.; funding acquisition, M.J., P.M. and L.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research is supported by the Spanish Grants PGC2018-095 194-B-100 (M. Jaenada, P. Miranda and L. Pardo) and FPU 19/01824 (M.Jaenada). The three authors are members of the Interdisciplinary Mathematics Institute (IMI).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: We are very grateful to the referees and associate editor for their helpful comments and suggestions.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

| | |
|-------|---|
| DPD | Density Power Divergence |
| IF | Influence Function |
| KL | Kullback–Leibler |
| LRM | Linear Regression Model |
| MLE | Maximum Likelihood Estimator |
| MDPDE | Minimum Density Power Divergence Estimator |
| MRPE | Minimum Rényi Pseudodistance Estimator |
| RMRPE | Restricted minimum Rényi Pseudodistance Estimator |
| RP | Rényi Pseudodistance |
| RPTS | Rényi Pseudodistance test statistic |

Appendix A

Appendix A.1. Proof of Theorem 2

We denote

$$h_n(\theta) = \frac{1}{n} \sum_{i=1}^n \frac{f_{\theta}(X_i)^{\tau}}{C_{\tau}(\theta)}.$$

Differentiating both sides of the equality, we have

$$\frac{\partial h_n(\theta)}{\partial \theta} = \frac{\tau}{C_{\tau}(\theta)} \frac{1}{n} \sum_{i=1}^n f_{\theta}(X_i)^{\tau} (\mathbf{u}_{\theta}(X_i) - \mathbf{c}_{\tau}(\theta)).$$

Now we establish that

$$\left(\frac{\partial^2 h_n(\theta)}{\partial \theta \partial \theta^T} \right)_{\theta=\theta_0} \xrightarrow[n \rightarrow \infty]{P} -\frac{\tau}{C_{\tau}(\theta_0)} \mathbf{S}_{\tau}(\theta_0).$$

We have

$$\begin{aligned} \frac{\partial^2 h_n(\theta)}{\partial \theta \partial \theta^T} &= \frac{1}{C_{\tau}(\theta)^2} \left\{ \frac{1}{n} \sum_{i=1}^n \left[\left(\tau^2 f_{\theta}(X_i)^{\tau} \mathbf{u}_{\theta}(X_i) \mathbf{u}_{\theta}(X_i)^T + \tau f_{\theta}(X_i)^{\tau} \frac{\partial \mathbf{u}_{\theta}(X_i)}{\partial \theta^T} \right) C_{\tau}(\theta) \right. \right. \\ &\quad \left. \left. - \tau C_{\tau}(\theta) \mathbf{c}_{\tau}(\theta) \tau f_{\theta}(X_i)^{\tau} \mathbf{u}_{\theta}(X_i)^T \right] \right\} \\ &\quad - \frac{1}{C_{\tau}(\theta)^2} \left\{ \frac{1}{n} \sum_{i=1}^n \left[\left(\tau \frac{\partial C_{\tau}(\theta)}{\partial \theta^T} f_{\theta}(X_i)^{\tau} + \tau^2 f_{\theta}(X_i)^{\tau} \mathbf{c}_{\tau}(\theta) \mathbf{u}_{\theta}(X_i)^T \right) C_{\tau}(\theta) \right. \right. \\ &\quad \left. \left. - \tau C_{\tau}(\theta) \mathbf{c}_{\tau}(\theta) \tau C_{\tau}(\theta)^T f_{\theta}(X_i)^{\tau} \right] \right\} \\ &= \frac{1}{C_{\tau}(\theta)} \left\{ \frac{1}{n} \sum_{i=1}^n \left[\tau^2 f_{\theta}(X_i)^{\tau} \mathbf{u}_{\theta}(X_i) \mathbf{u}_{\theta}(X_i)^T + \tau f_{\theta}(X_i)^{\tau} \frac{\partial \mathbf{u}_{\theta}(X_i)}{\partial \theta^T} \right. \right. \\ &\quad \left. \left. - \tau^2 \mathbf{c}_{\tau}(\theta)^T f_{\theta}(X_i)^{\tau} \mathbf{u}_{\theta}(X_i)^T - \tau \frac{\partial C_{\tau}(\theta)}{\partial \theta^T} f_{\theta}(X_i)^{\tau} \right. \right. \\ &\quad \left. \left. - \tau^2 \mathbf{c}_{\tau}(\theta) f_{\theta}(X_i)^{\tau} \mathbf{u}_{\theta}(X_i)^T - \tau^2 \mathbf{c}_{\tau}(\theta) \mathbf{c}_{\tau}(\theta)^T f_{\theta}(X_i)^{\tau} \right] \right\}. \end{aligned}$$

As $n \rightarrow \infty$, we have

$$\left(\frac{\partial^2 h_n(\theta)}{\partial \theta \partial \theta^T} \right)_{\theta=\theta_0} \xrightarrow[n \rightarrow \infty]{P} \mathbf{T}(\theta_0)$$

with $\mathbf{T}(\theta_0)$ being the matrix given by

$$\begin{aligned} \mathbf{T}(\theta_0) &= \frac{1}{C_{\tau}(\theta_0)} \left\{ \tau^2 \int f_{\theta}(x)^{\tau+1} \mathbf{u}_{\theta}(x) \mathbf{u}_{\theta}(x)^T dx + \tau \int f_{\theta}(x)^{\tau+1} \frac{\partial \mathbf{u}_{\theta}(x)}{\partial \theta^T} dx \right. \\ &\quad \left. - \tau^2 \mathbf{c}_{\tau}(\theta_0) \int f_{\theta}(x)^{\tau+1} \mathbf{u}_{\theta}(x)^T dx - \tau \frac{\partial C_{\tau}(\theta_0)}{\partial \theta^T} \int f_{\theta}(x)^{\tau+1} dx \right. \\ &\quad \left. + \tau^2 \mathbf{c}_{\tau}(\theta_0) \int f_{\theta}(x)^{\tau+1} \mathbf{u}_{\theta}(x)^T dx - \tau^2 \mathbf{c}_{\tau}(\theta_0) \mathbf{c}_{\tau}(\theta_0)^T \int f_{\theta}(x)^{\tau+1} dx \right\}. \end{aligned}$$

From the above, after some algebra, we obtain

$$T(\theta_0) = \frac{1}{C_\tau(\theta)} \left\{ \tau^2 \int f_\theta(x)^{\tau+1} \mathbf{u}_\theta(x) \mathbf{u}_\theta(x)^T dx + \tau \int f_\theta(x)^{\tau+1} \frac{\partial \mathbf{u}_\theta(x)}{\partial \theta^T} dx - \tau \frac{\partial \mathbf{c}_\tau(\theta)}{\partial \theta^T} \int f_\theta(x)^{\tau+1} dx - \tau^2 \mathbf{c}_\tau(\theta)^T \mathbf{c}_\tau(\theta) \int f_\theta(x)^{\tau+1} dx \right\}.$$

On the other hand, it not difficult to establish that

$$\frac{\partial \mathbf{c}_\tau(\theta)}{\partial \theta^T} = (\tau + 1) \frac{\int f_\theta(x)^{\tau+1} \mathbf{u}_\theta(x) \mathbf{u}_\theta(x)^T dx}{\int f_\theta(x)^{\tau+1} dx} + \frac{\int f_\theta(x)^{\tau+1} \frac{\partial \mathbf{u}_\theta(x)}{\partial \theta^T} dx}{\int f_\theta(x)^{\tau+1} dx} - (\tau + 1) \frac{\int f_\theta(x)^{\tau+1} \mathbf{u}_\theta(x) dx \int f_\theta(x)^{\tau+1} \mathbf{u}_\theta(x)^T dx}{(\int f_\theta(x)^{\tau+1} dx)^2}.$$

Therefore we have

$$-\tau \frac{\partial \mathbf{c}_\tau(\theta)}{\partial \theta^T} \int f_\theta(x)^{\tau+1} dx = -\tau(\tau + 1) \int f_\theta(x)^{\tau+1} \mathbf{u}_\theta(x) \mathbf{u}_\theta(x)^T dx - \tau \int f_\theta(x)^{\tau+1} \frac{\partial \mathbf{u}_\theta(x)}{\partial \theta^T} dx + \tau(\tau + 1) \frac{\int f_\theta(x)^{\tau+1} \mathbf{u}_\theta(x) dx \int f_\theta(x)^{\tau+1} \mathbf{u}_\theta(x)^T dx}{\int f_\theta(x)^{\tau+1} dx}.$$

Finally,

$$\begin{aligned} T(\theta_0) &= \frac{1}{C_\tau(\theta)} \left\{ \tau^2 \int f_\theta(x)^{\tau+1} \mathbf{u}_\theta(x) \mathbf{u}_\theta(x)^T dx + \tau \int f_\theta(x)^{\tau+1} \frac{\partial \mathbf{u}_\theta(x)}{\partial \theta^T} dx - \tau(\tau + 1) \int f_\theta(x)^{\tau+1} \mathbf{u}_\theta(x) \mathbf{u}_\theta(x)^T dx - \tau \int f_\theta(x)^{\tau+1} \frac{\partial \mathbf{u}_\theta(x)}{\partial \theta^T} dx + \tau(\tau + 1) \frac{\int f_\theta(x)^{\tau+1} \mathbf{u}_\theta(x) dx \int f_\theta(x)^{\tau+1} \mathbf{u}_\theta(x)^T dx}{\int f_\theta(x)^{\tau+1} dx} - \tau^2 \mathbf{c}_\tau(\theta) \mathbf{c}_\tau(\theta)^T \int f_\theta(x)^{\tau+1} dx \right\} \\ &= \frac{1}{C_\tau(\theta)} \left\{ -\tau \int f_\theta(x)^{\tau+1} \mathbf{u}_\theta(x) \mathbf{u}_\theta(x)^T dx + \tau \frac{\int f_\theta(x)^{\tau+1} \mathbf{u}_\theta(x) dx \int f_\theta(x)^{\tau+1} \mathbf{u}_\theta(x)^T dx}{\int f_\theta(x)^{\tau+1} dx} \right\} \\ &= -\frac{\tau}{C_\tau(\theta)} \mathbf{S}(\theta_0). \end{aligned}$$

On the other hand,

$$\sqrt{n} \frac{\partial h_n(\theta)}{\partial \theta} = \frac{\tau}{C_\tau(\theta)} \frac{1}{\sqrt{n}} \sum_{i=1}^n f_\theta(X_i)^\tau (\mathbf{u}_\theta(X_i) - \mathbf{c}_\tau(\theta)) \xrightarrow[n \rightarrow \infty]{L} \mathcal{N} \left(\mathbf{0}_p, \left(\frac{\tau}{C_\tau(\theta)} \right)^2 \mathbf{K}_\tau(\theta_0) \right),$$

as

$$E \left[\left(\frac{\tau}{C_\tau(\theta)} f_\theta(X)^\tau (\mathbf{u}_\theta(X) - \mathbf{c}_\tau(\theta)) \right)_{\theta=\theta_0} \right] = \mathbf{0}_p$$

and

$$\text{Cov} \left[\left(\frac{\tau}{C_\tau(\theta)} f_\theta(X)^\tau (\mathbf{u}_\theta(X) - \mathbf{c}_\tau(\theta)) \right)_{\theta=\theta_0} \right] = \left(\frac{\tau}{C_\tau(\theta_0)} \right)^2 \mathbf{K}_\tau(\theta_0)$$

Then, the RMRPE estimator of θ , $\tilde{\theta}_\tau$, must satisfy

$$\begin{cases} \frac{\partial}{\partial \theta} h_n(\theta) |_{\theta=\tilde{\theta}_\tau} + \mathbf{G}(\tilde{\theta}_\tau) \lambda_n = \mathbf{0}_p, \\ \mathbf{g}(\tilde{\theta}_\tau) = \mathbf{0}_r, \end{cases} \tag{A1}$$

where λ_n is a vector of Lagrangian multipliers. Now we consider $\theta_n = \theta_0 + mn^{-1/2}$, with $\|m\| < k$, for $0 < k < \infty$. We have,

$$\frac{\partial}{\partial \theta} h_n(\theta)|_{\theta=\theta_n} = \frac{\partial}{\partial \theta} h_n(\theta)|_{\theta=\theta_0} + \frac{\partial}{\partial \theta^T} \frac{\partial}{\partial \theta} h_n(\theta)|_{\theta=\theta_0} (\theta_n - \theta_0) + o(\|\theta_n - \theta_0\|^2)$$

and

$$n^{1/2} \frac{\partial}{\partial \theta} h_n(\theta) \Big|_{\theta=\theta_n} = n^{1/2} \frac{\partial}{\partial \theta} h_n(\theta)|_{\theta=\theta_0} - \frac{\partial}{\partial \theta^T} \frac{\partial}{\partial \theta} h_n(\theta)|_{\theta=\theta_0} n^{1/2} (\theta_n - \theta_0) + o(n^{1/2} \|\theta_n - \theta_0\|^2). \tag{A2}$$

However,

$$o(n^{1/2} \|\theta_n - \theta_0\|^2) = o(n^{1/2} \|m\|^2/n) = o(n^{-1/2} \|m\|^2) = o(O_p(1)) = o_p(1).$$

Since

$$\lim_{n \rightarrow \infty} \frac{\partial}{\partial \theta^T} \frac{\partial}{\partial \theta} h_n(\theta)|_{\theta=\theta_0} = -\frac{\tau}{C_\tau(\theta)} S_\tau(\theta_0)$$

we obtain

$$n^{1/2} \frac{\partial}{\partial \theta} h_n(\theta) \Big|_{\theta=\theta_n} = n^{1/2} \frac{\partial}{\partial \theta} h_n(\theta)|_{\theta=\theta_0} + \frac{\tau}{C_\tau(\theta)} S_\tau(\theta_0) n^{1/2} (\theta_n - \theta_0) + o_p(1). \tag{A3}$$

Now, we know that

$$n^{1/2} g(\theta_n) = G(\theta_0)^T n^{1/2} (\theta_n - \theta_0) + o_p(1). \tag{A4}$$

Further, the RMRPE $\tilde{\theta}_\tau$ must satisfy the conditions in (A1), and in view of (A3) and (A4) we have

$$n^{1/2} \frac{\partial}{\partial \theta} h_n(\theta)|_{\theta=\theta_0} + \frac{\tau}{C_\tau(\theta)} S_\tau(\theta_0) n^{1/2} (\tilde{\theta}_\tau - \theta_0) + G(\theta_0) n^{1/2} \lambda_n + o_p(1) = \mathbf{0}_p. \tag{A5}$$

From (A4) it follows that

$$G(\theta_0)^T n^{1/2} (\tilde{\theta}_\tau - \theta_0) + o_p(1) = \mathbf{0}_r. \tag{A6}$$

Now we can express equations (A5) and (A6) in matrix form as

$$\begin{pmatrix} \frac{\tau}{C_\tau(\theta_0)} S_\tau(\theta_0) & G(\theta_0) \\ G(\theta_0)^T & \mathbf{0}_{r \times r} \end{pmatrix} \begin{pmatrix} n^{1/2} (\tilde{\theta}_\tau - \theta_0) \\ n^{1/2} \lambda_n \end{pmatrix} = \begin{pmatrix} -n^{1/2} \frac{\partial}{\partial \theta} h_n(\theta)|_{\theta=\theta_0} \\ \mathbf{0}_r \end{pmatrix} + o_p(1).$$

Therefore

$$\begin{pmatrix} n^{1/2} (\tilde{\theta}_\tau - \theta_0) \\ n^{1/2} \lambda_n \end{pmatrix} = \begin{pmatrix} \frac{\tau}{C_\tau(\theta_0)} S_\tau(\theta_0) & G(\theta_0) \\ G(\theta_0)^T & \mathbf{0}_{r \times r} \end{pmatrix}^{-1} \begin{pmatrix} -n^{1/2} \frac{\partial}{\partial \theta} h_n(\theta)|_{\theta=\theta_0} \\ \mathbf{0}_r \end{pmatrix} + o_p(1).$$

However,

$$\begin{pmatrix} \frac{\tau}{C_\tau(\theta_0)} S_\tau(\theta_0) & G(\theta_0) \\ G(\theta_0)^T & \mathbf{0} \end{pmatrix}^{-1} = \begin{pmatrix} \left(\frac{\tau}{C_\tau(\theta_0)}\right)^{-1} P_\tau^*(\theta_0) & Q_\tau(\theta_0) \\ Q_\tau(\theta_0)^T & R_\tau(\theta_0) \end{pmatrix},$$

where $P_\tau^*(\theta_0)$ and $Q_\tau(\theta_0)$ are defined in (18) and (19), respectively. The matrix $R_\tau(\theta_0)$ is the quantity needed to make the right hand side of the above equation equal to the indicated inverse. Then,

$$n^{1/2} (\tilde{\theta}_\tau - \theta_0) = -\left(\frac{\tau}{C_\tau(\theta)}\right)^{-1} P_\tau^*(\theta_0) n^{1/2} \frac{\partial}{\partial \theta} h_n(\theta)|_{\theta=\theta_0} + o_p(1), \tag{A7}$$

and we know

$$n^{1/2} \frac{\partial}{\partial \theta} h_n(\theta) \Big|_{\theta=\theta_0} \xrightarrow{n \rightarrow \infty} \mathcal{N} \left(\mathbf{0}, \left(\frac{\tau}{C_\tau(\theta_0)} \right)^2 \mathbf{K}_\tau(\theta_0) \right). \tag{A8}$$

Now by (A7) and (A8), we have the desired result.

Appendix A.2. Proof of Theorem 3

Consider the expression $R_\gamma(f_\theta, f_{\tilde{\theta}_\tau})$. A Taylor expansion for an arbitrary $\theta \in \Theta$, around $\tilde{\theta}_\tau$ leads to the relation

$$R_\gamma(f_\theta, f_{\tilde{\theta}_\tau}) = R_\gamma(f_{\tilde{\theta}_\tau}, f_{\tilde{\theta}_\tau}) + \left(\frac{\partial R_\gamma(f_\theta, f_{\tilde{\theta}_\tau})}{\partial \theta} \right)_{\theta=\tilde{\theta}_\tau} (\theta - \tilde{\theta}_\tau) + \frac{1}{2} (\theta - \tilde{\theta}_\tau) \left(\frac{\partial^2 R_\gamma(f_\theta, f_{\tilde{\theta}_\tau})}{\partial \theta \partial \theta^T} \right)_{\theta=\tilde{\theta}_\tau} (\theta - \tilde{\theta}_\tau)^T + o(\|\theta - \tilde{\theta}_\tau\|^2).$$

It is clear that $R_\gamma(f_{\tilde{\theta}_\tau}, f_{\tilde{\theta}_\tau}) = 0$ and

$$\frac{\partial R_\gamma(f_\theta, f_{\tilde{\theta}_\tau})}{\partial \theta} = \frac{\partial L_\gamma^1(\theta)}{\partial \theta} - \frac{\partial L_\gamma^2(\theta)}{\partial \theta},$$

being

$$L_\gamma^1(\theta) = \frac{1}{\gamma + 1} \log \left(\int f_\theta(x)^{\gamma+1} dx \right)$$

and

$$L_\gamma^2(\theta) = \frac{1}{\gamma} \log \left(\int f_\theta(x)^\gamma f_{\tilde{\theta}_\tau}(x) dx \right).$$

Then,

$$\frac{\partial L_\gamma^1(\theta)}{\partial \theta} = \frac{\int f_\theta(x)^{\gamma+1} \mathbf{u}_\theta(x) dx}{\int f_\theta(x)^{\gamma+1} dx} \text{ and } \frac{\partial L_\gamma^2(\theta)}{\partial \theta} = \frac{\int f_\theta(x)^\gamma \mathbf{u}_\theta(x) f_{\tilde{\theta}_\tau}(x) dx}{\int f_\theta(x)^\gamma f_{\tilde{\theta}_\tau}(x) dx}.$$

Therefore,

$$\left(\frac{\partial R_\gamma(f_\theta, f_{\tilde{\theta}_\tau})}{\partial \theta} \right)_{\theta=\tilde{\theta}_\tau} = 0.$$

Regarding the second derivatives, we have

$$\begin{aligned} \frac{\partial^2 L_\gamma^1(\theta)}{\partial \theta \partial \theta^T} &= (\gamma + 1) \frac{\int f_\theta(x)^{\gamma+1} \mathbf{u}_\theta(x) \mathbf{u}_\theta(x)^T dx}{\int f_\theta(x)^{\gamma+1} dx} + \frac{\int f_\theta(x)^{\gamma+1} \frac{\partial \mathbf{u}_\theta(x)}{\partial \theta^T} dx}{\int f_\theta(x)^{\gamma+1} dx} \\ &\quad - (\gamma + 1) \frac{\int f_\theta(x)^{\gamma+1} \mathbf{u}_\theta(x) dx \int f_\theta(x)^{\gamma+1} \mathbf{u}_\theta(x)^T dx}{\left(\int f_\theta(x)^{\gamma+1} dx \right)^2} \end{aligned}$$

and

$$\begin{aligned} \frac{\partial^2 L_\gamma^2(\theta)}{\partial \theta \partial \theta^T} &= \gamma \frac{\int f_\theta(x)^\gamma \mathbf{u}_\theta(x) \mathbf{u}_\theta(x)^T f_{\tilde{\theta}_\tau}(x) dx}{\int f_\theta(x)^\gamma f_{\tilde{\theta}_\tau}(x) dx} + \frac{\int f_\theta(x)^\gamma \frac{\partial \mathbf{u}_\theta(x)}{\partial \theta^T} f_{\tilde{\theta}_\tau}(x) dx}{\int f_\theta(x)^\gamma f_{\tilde{\theta}_\tau}(x) dx} \\ &\quad - \gamma \frac{\int f_\theta(x)^\gamma f_{\tilde{\theta}_\tau}(x) \mathbf{u}_\theta(x) dx \int f_\theta(x)^\gamma f_{\tilde{\theta}_\tau}(x) \mathbf{u}_\theta(x)^T dx}{\left(\int f_\theta(x)^\gamma f_{\tilde{\theta}_\tau}(x) dx \right)^2}. \end{aligned}$$

and so

$$\left(\frac{\partial^2 R_\gamma(f_\theta, f_{\tilde{\theta}_\tau})}{\partial \theta \partial \theta^T} \right)_{\theta=\tilde{\theta}_\tau} = \frac{\mathbf{S}_\gamma(\tilde{\theta}_\tau)}{\kappa_\gamma(\tilde{\theta}_\tau)}$$

Therefore,

$$T_\gamma(\widehat{\boldsymbol{\theta}}_\tau, \widetilde{\boldsymbol{\theta}}_\tau) = 2nR_\gamma(f_{\widehat{\boldsymbol{\theta}}_\tau}, f_{\widetilde{\boldsymbol{\theta}}_\tau}) = n^{1/2}(\widehat{\boldsymbol{\theta}}_\tau - \widetilde{\boldsymbol{\theta}}_\tau)^T \frac{\mathbf{S}_\gamma(\widetilde{\boldsymbol{\theta}}_\tau)}{\kappa_\gamma(\widetilde{\boldsymbol{\theta}}_\tau)} n^{1/2}(\widehat{\boldsymbol{\theta}}_\tau - \widetilde{\boldsymbol{\theta}}_\tau) + n \times o\left(\|\widehat{\boldsymbol{\theta}}_\tau - \widetilde{\boldsymbol{\theta}}_\tau\|^2\right).$$

Under $\boldsymbol{\theta}_0 \in \Theta_0$,

$$\frac{\mathbf{S}_\gamma(\widetilde{\boldsymbol{\theta}}_\tau)}{\kappa_\gamma(\widetilde{\boldsymbol{\theta}}_\tau)} \xrightarrow[n \rightarrow \infty]{\mathcal{P}_\gamma} \frac{\mathbf{S}_\gamma(\boldsymbol{\theta}_0)}{\kappa_\gamma(\boldsymbol{\theta}_0)}.$$

Based on $\widehat{\boldsymbol{\theta}}_\tau$ and using by (A4) and (A5), we have that

$$n^{1/2} \frac{\partial}{\partial \boldsymbol{\theta}} h_n(\boldsymbol{\theta})|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} = -\frac{\tau}{C_\tau(\boldsymbol{\theta}_0)} n^{1/2} \mathbf{S}_\tau(\boldsymbol{\theta}_0) (\widehat{\boldsymbol{\theta}}_\tau - \boldsymbol{\theta}_0) + o_p(1),$$

and using (A7), we obtain

$$\begin{aligned} n^{1/2}(\widetilde{\boldsymbol{\theta}}_\tau - \boldsymbol{\theta}_0) &= \mathbf{P}_\tau^*(\boldsymbol{\theta}_0) n^{1/2} \mathbf{S}_\tau(\boldsymbol{\theta}_0) (\widehat{\boldsymbol{\theta}}_\tau - \boldsymbol{\theta}_0) + o_p(1) \\ &= n^{1/2}(\widehat{\boldsymbol{\theta}}_\tau - \boldsymbol{\theta}_0) - \mathbf{Q}_\tau(\boldsymbol{\theta}_0) \mathbf{G}(\boldsymbol{\theta}_0)^T n^{1/2}(\widehat{\boldsymbol{\theta}}_\tau - \boldsymbol{\theta}_0) + o_p(1). \end{aligned}$$

Therefore,

$$n^{1/2}(\widehat{\boldsymbol{\theta}}_\tau - \widetilde{\boldsymbol{\theta}}_\tau) = \mathbf{Q}_\tau(\boldsymbol{\theta}_0) \mathbf{G}(\boldsymbol{\theta}_0)^T n^{1/2}(\widehat{\boldsymbol{\theta}}_\tau - \boldsymbol{\theta}_0) + o_p(1). \tag{A9}$$

On the other hand, we know that

$$n^{1/2}(\widehat{\boldsymbol{\theta}}_\tau - \boldsymbol{\theta}_0) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}(0, \mathbf{S}_\tau(\boldsymbol{\theta}_0)^{-1} \mathbf{K}_\tau(\boldsymbol{\theta}_0) \mathbf{S}_\tau(\boldsymbol{\theta}_0)^{-1}).$$

From equations (19) and (25), we can establish that

$$\begin{aligned} \mathbf{B}_\tau(\boldsymbol{\theta}_0) &= \mathbf{S}_\tau(\boldsymbol{\theta}_0)^{-1} \mathbf{G}(\boldsymbol{\theta}_0) \left[\mathbf{G}(\boldsymbol{\theta}_0)^T \mathbf{S}_\tau(\boldsymbol{\theta}_0)^{-1} \mathbf{G}(\boldsymbol{\theta}_0) \right]^{-1} \mathbf{G}(\boldsymbol{\theta}_0)^T \mathbf{S}_\tau(\boldsymbol{\theta}_0)^{-1} \\ &= \mathbf{Q}_\tau(\boldsymbol{\theta}_0) \mathbf{G}(\boldsymbol{\theta}_0)^{-1} \mathbf{S}_\tau(\boldsymbol{\theta}_0)^{-1}. \end{aligned}$$

Therefore, it follows that

$$n^{1/2}(\widehat{\boldsymbol{\theta}}_\tau - \widetilde{\boldsymbol{\theta}}_\tau) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}(0, \mathbf{B}_\tau(\boldsymbol{\theta}_0) \mathbf{K}_\tau(\boldsymbol{\theta}_0) \mathbf{B}_\tau(\boldsymbol{\theta}_0)^T).$$

Now, observe from the definition that $\mathbf{B}_\tau(\boldsymbol{\theta}_0) = \mathbf{B}_\tau(\boldsymbol{\theta}_0)^T$. Then, the asymptotic distribution of the random variables

$$T_\gamma(\widehat{\boldsymbol{\theta}}_\tau, \widetilde{\boldsymbol{\theta}}_\tau) = 2nR_\gamma(f_{\widehat{\boldsymbol{\theta}}_\tau}, f_{\widetilde{\boldsymbol{\theta}}_\tau})$$

and

$$n^{1/2}(\widehat{\boldsymbol{\theta}}_\tau - \widetilde{\boldsymbol{\theta}}_\tau)^T \frac{\mathbf{S}_\gamma(\boldsymbol{\theta}_0)}{\kappa_\gamma(\boldsymbol{\theta}_0)} n^{1/2}(\widehat{\boldsymbol{\theta}}_\tau - \widetilde{\boldsymbol{\theta}}_\tau)$$

are the same, as we have established that

$$n \times o\left(\|\widehat{\boldsymbol{\theta}}_\tau - \widetilde{\boldsymbol{\theta}}_\tau\|^2\right) = o_p(1).$$

Next, we apply Corollary 2.1 in Dik and Gunst [30], which states: “Let \mathbf{X} be a q -variate normal random variable with mean vector $\mathbf{0}$ and variance-covariance matrix $\boldsymbol{\Sigma}$. Let \mathbf{M} be a real symmetric matrix of order q . Let $k = \text{rank}(\boldsymbol{\Sigma} \mathbf{M} \boldsymbol{\Sigma})$, $k \geq 1$ and let $\lambda_1, \dots, \lambda_k$, be the nonzero eigenvalues of $\mathbf{M} \boldsymbol{\Sigma}$. Then, the distribution of the quadratic form $\mathbf{X}^T \mathbf{M} \mathbf{X}$ coincides with the distribution of the random variable $\sum_{i=1}^k \lambda_i Z_i^2$, where Z_1, \dots, Z_k are independent, each having a standard normal variable”. In our case, the asymptotic distribution of $T_\gamma(\widehat{\boldsymbol{\theta}}_\tau, \widetilde{\boldsymbol{\theta}}_\tau)$ coincides with the distribution of the random variable $\sum_{i=1}^k \lambda_i^{\tau, \gamma}(\boldsymbol{\theta}_0) Z_i^2$ where $\lambda_1^{\tau, \gamma}(\boldsymbol{\theta}_0), \dots, \lambda_k^{\tau, \gamma}(\boldsymbol{\theta}_0)$, are the nonzero eigenvalues of $\mathbf{A}_\gamma(\boldsymbol{\theta}_0) \mathbf{B}_\tau(\boldsymbol{\theta}_0) \mathbf{K}_\tau(\boldsymbol{\theta}_0) \mathbf{B}_\tau(\boldsymbol{\theta}_0)$ and

$$k = \min\{r, \text{rank}(\mathbf{B}_\tau(\boldsymbol{\theta}_0)\mathbf{K}_\tau(\boldsymbol{\theta}_0)\mathbf{B}_\tau(\boldsymbol{\theta}_0)\mathbf{A}_\gamma(\boldsymbol{\theta}_0)\mathbf{B}_\tau(\boldsymbol{\theta}_0)\mathbf{K}_\tau(\boldsymbol{\theta}_0)\mathbf{B}_\tau(\boldsymbol{\theta}_0))\}. \tag{A10}$$

We now establish that $k = r$. The matrix,

$$\mathbf{N}_\tau(\boldsymbol{\theta}_0) = \mathbf{B}_\tau(\boldsymbol{\theta}_0)\mathbf{K}_\tau(\boldsymbol{\theta}_0)\mathbf{B}_\tau(\boldsymbol{\theta}_0)$$

is given by

$$\begin{aligned} \mathbf{N}_\tau(\boldsymbol{\theta}_0) &= \mathbf{S}_\tau(\boldsymbol{\theta}_0)^{-1}\mathbf{G}(\boldsymbol{\theta}_0)\left[\mathbf{G}(\boldsymbol{\theta}_0)^T\mathbf{S}_\tau(\boldsymbol{\theta}_0)^{-1}\mathbf{G}(\boldsymbol{\theta}_0)\right]^{-1}\mathbf{G}(\boldsymbol{\theta}_0)^T\mathbf{S}_\tau(\boldsymbol{\theta}_0)^{-1} \\ &\quad \mathbf{K}_\tau(\boldsymbol{\theta}_0)\mathbf{S}_\tau(\boldsymbol{\theta}_0)^{-1}\mathbf{G}(\boldsymbol{\theta}_0)\left[\mathbf{G}(\boldsymbol{\theta}_0)^T\mathbf{S}_\tau(\boldsymbol{\theta}_0)^{-1}\mathbf{G}(\boldsymbol{\theta}_0)\right]^{-1}\mathbf{G}(\boldsymbol{\theta}_0)^T\mathbf{S}_\tau(\boldsymbol{\theta}_0)^{-1}. \end{aligned}$$

Corollary 14.11.3 in Harville [31] (p. 259) establishes the following: “For any $m \times n$ matrix A and any $m \times m$ symmetric positive definite matrix W , $\text{rank}(A^T W A) = \text{rank}(A)$ ”. Based on this Corollary we have that $\text{rank}(\mathbf{N}_\tau(\boldsymbol{\theta}_0))$ coincides with $\text{rank}(\mathbf{N}_\tau(\boldsymbol{\theta}_0)\mathbf{S}_\gamma(\boldsymbol{\theta}_0)\mathbf{N}_\tau(\boldsymbol{\theta}_0))$.

On the other hand, we know the following additional properties:

- (a) $\text{rank}(\mathbf{AB}) = \text{rank}(A)$ if B is full rank (Corollary b.3.3 in Harville [31] (p. 83)).
- (b) $\text{rank}(\mathbf{AB}) = \text{rank}(\mathbf{BA})$ if dimension of A coincides with dimension of B^T .

Matrix $\mathbf{K}_\tau(\boldsymbol{\theta}_0)$ should be “full rank”; in fact, if $\mathbf{K}_\tau(\boldsymbol{\theta}_0)$ were not full rank, the variance-covariance matrix of $\hat{\boldsymbol{\theta}}_\beta$ and $\tilde{\boldsymbol{\theta}}_\beta$ would not be full rank (there were redundant components in $\boldsymbol{\theta}$ and this is not true).

Therefore, we have

$$\begin{aligned} \text{rank}(\mathbf{N}_\tau(\boldsymbol{\theta}_0)) &= \text{(a)} \text{rank}\left(\mathbf{S}_\tau(\boldsymbol{\theta}_0)^{-\frac{1}{2}}\mathbf{G}(\boldsymbol{\theta}_0)\left[\mathbf{G}(\boldsymbol{\theta}_0)^T\mathbf{S}_\tau(\boldsymbol{\theta}_0)^{-1}\mathbf{G}(\boldsymbol{\theta}_0)\right]^{-1}\right. \\ &\quad \left.\mathbf{G}(\boldsymbol{\theta}_0)^T\mathbf{S}_\tau(\boldsymbol{\theta}_0)^{-1}\mathbf{K}_\tau(\boldsymbol{\theta}_0)\mathbf{S}_\tau(\boldsymbol{\theta}_0)^{-1}\mathbf{G}(\boldsymbol{\theta}_0)\right. \\ &\quad \left.\left[\mathbf{G}(\boldsymbol{\theta}_0)^{-1}\mathbf{S}_\tau(\boldsymbol{\theta}_0)^{-1}\mathbf{G}(\boldsymbol{\theta}_0)\right]^{-1}\mathbf{G}(\boldsymbol{\theta}_0)^T\mathbf{S}_\tau(\boldsymbol{\theta}_0)^{-\frac{1}{2}}\right) \\ &= \text{(b)} \text{rank}\left(\mathbf{G}(\boldsymbol{\theta}_0)^T\mathbf{S}_\tau(\boldsymbol{\theta}_0)^{-1}\mathbf{K}_\tau(\boldsymbol{\theta}_0)\mathbf{S}_\tau(\boldsymbol{\theta}_0)^{-1}\mathbf{G}(\boldsymbol{\theta}_0)\left[\mathbf{G}(\boldsymbol{\theta}_0)^T\mathbf{S}_\tau(\boldsymbol{\theta}_0)^{-1}\mathbf{G}(\boldsymbol{\theta}_0)\right]^{-1}\right) \\ &= \text{(a)} \text{rank}\left(\mathbf{G}(\boldsymbol{\theta}_0)^T\mathbf{S}_\tau(\boldsymbol{\theta}_0)^{-1}\mathbf{K}_\tau(\boldsymbol{\theta}_0)\mathbf{S}_\tau(\boldsymbol{\theta}_0)^{-1}\mathbf{G}(\boldsymbol{\theta}_0)\right) \\ &= \text{Corollary 14.11.3} \text{rank}\left(\mathbf{S}_\tau(\boldsymbol{\theta}_0)^{-1}\mathbf{G}(\boldsymbol{\theta}_0)\right) \\ &= \text{(a)} \text{rank}(\mathbf{G}(\boldsymbol{\theta}_0)) = r. \end{aligned}$$

Appendix A.3. Rényi’s Pseudodistance between Normal Populations

Here, we compute the expression of the RP between densities belonging to the normal model with parameters (μ_1, σ_1) and (μ_2, σ_2) , respectively. The RP between $\mathcal{N}(\mu_1, \sigma_1)$ and $\mathcal{N}(\mu_2, \sigma_2)$ is given by

$$\begin{aligned} R_\gamma(\mathcal{N}(\mu_1, \sigma_1), \mathcal{N}(\mu_2, \sigma_2)) &= \frac{1}{\gamma + 1} \log \int \mathcal{N}(\mu_1, \sigma_1)^{\gamma+1} dx \\ &\quad + \frac{1}{\gamma(\gamma + 1)} \log \int \mathcal{N}(\mu_2, \sigma_2)^{\gamma+1} dx - \frac{1}{\gamma} \log \int \mathcal{N}(\mu_1, \sigma_1)^\gamma \mathcal{N}(\mu_2, \sigma_2) dx \\ &= \frac{1}{\gamma + 1} \log L_1 + \frac{1}{\gamma(\gamma + 1)} \log L_2 - \frac{1}{\gamma} \log L_3. \end{aligned}$$

We first compute

$$\int \mathcal{N}(\mu, \sigma)^\beta dx$$

for the seek of simplicity in later calculations.

$$\begin{aligned} \int \mathcal{N}(\mu, \sigma)^\beta dx &= \int \left(\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \right)^\beta dx \\ &= \frac{1}{\sigma^{\beta-1}(\sqrt{2\pi})^{\beta-1}} \frac{1}{\sqrt{\beta}} \int \frac{1}{\frac{\sigma}{\sqrt{\beta}}\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\frac{\sigma}{\sqrt{\beta}}}\right)^2} dx \\ &= \frac{1}{\sigma^{\beta-1}(\sqrt{2\pi})^{\beta-1}} \frac{1}{\sqrt{\beta}}. \end{aligned}$$

Therefore,

$$L_1 = \frac{1}{\sigma_1^\gamma(\sqrt{2\pi})^\gamma} \frac{1}{\sqrt{\gamma+1}} \text{ and } L_2 = \frac{1}{\sigma_2^\gamma(\sqrt{2\pi})^\gamma} \frac{1}{\sqrt{\gamma+1}}.$$

In relation with L_3 we have,

$$\begin{aligned} L_3 &= \int \mathcal{N}(\mu_1, \sigma_1)^\gamma \mathcal{N}(\mu_2, \sigma_2) dx \\ &= \int \frac{1}{\sigma_1^\gamma(\sqrt{2\pi})^\gamma} e^{-\frac{1}{2}\left(\frac{x-\mu_1}{\frac{\sigma_1}{\sqrt{\gamma}}}\right)^2} \frac{1}{\sigma_2\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu_2}{\sigma_2}\right)^2} \\ &= \frac{1}{\sigma_1^\gamma(\sqrt{2\pi})^\gamma} \frac{1}{\sigma_2\sqrt{2\pi}} \times \\ &\quad \times \int \exp \left\{ -\frac{1}{2} \left[x^2 \left(\frac{1}{\left(\frac{\sigma_1}{\sqrt{\gamma}}\right)^2} + \frac{1}{\sigma_2^2} \right) - 2x \left(\frac{\mu_1}{\left(\frac{\sigma_1}{\sqrt{\gamma}}\right)^2} + \frac{\mu_2}{\sigma_2^2} \right) + \frac{\mu_1^2}{\left(\frac{\sigma_1}{\sqrt{\gamma}}\right)^2} + \frac{\mu_2^2}{\sigma_2^2} \right] \right\} dx \\ &= \frac{1}{\sigma_1^\gamma(\sqrt{2\pi})^\gamma} \frac{1}{\sigma_2\sqrt{2\pi}} \exp \left\{ -\frac{1}{2} \left[\frac{\mu_1^2}{\left(\frac{\sigma_1}{\sqrt{\gamma}}\right)^2} + \frac{\mu_2^2}{\sigma_2^2} \right] \right\} \times \\ &\quad \times \int \exp \left\{ -\frac{1}{2} \left[x^2 \left(\frac{1}{\left(\frac{\sigma_1}{\sqrt{\gamma}}\right)^2} + \frac{1}{\sigma_2^2} \right) - 2x \left(\frac{\mu_1}{\left(\frac{\sigma_1}{\sqrt{\gamma}}\right)^2} + \frac{\mu_2}{\sigma_2^2} \right) \right] \right\} dx \\ &= \frac{1}{\sigma_1^\gamma(\sqrt{2\pi})^\gamma} \frac{1}{\sigma_2\sqrt{2\pi}} \exp \left\{ -\frac{1}{2} \left[\frac{\mu_1^2}{\left(\frac{\sigma_1}{\sqrt{\gamma}}\right)^2} + \frac{\mu_2^2}{\sigma_2^2} \right] \right\} \exp \left\{ \frac{1}{2} \frac{A^2}{B^2} \right\} B\sqrt{2\pi} \times \\ &\quad \times \int \frac{1}{\sqrt{2\pi}B} \exp \left\{ -\frac{1}{2} \left(\frac{x-A}{B} \right)^2 \right\} dx \\ &= \frac{1}{\sigma_1^\gamma(\sqrt{2\pi})^\gamma} \frac{1}{\sigma_2\sqrt{2\pi}} \exp \left\{ -\frac{1}{2} \left[\frac{\mu_1^2}{\left(\frac{\sigma_1}{\sqrt{\gamma}}\right)^2} + \frac{\mu_2^2}{\sigma_2^2} \right] \right\} \exp \left\{ \frac{1}{2} \frac{A^2}{B^2} \right\} B\sqrt{2\pi}. \end{aligned}$$

Now it is necessary to obtain A and B . However, for this, we have,

$$\begin{cases} \frac{1}{B^2} = \frac{1}{\left(\frac{\sigma_1}{\sqrt{\gamma}}\right)^2} + \frac{1}{\sigma_2^2} \\ \frac{A}{B^2} = \left(\frac{\mu_1}{\left(\frac{\sigma_1}{\sqrt{\gamma}}\right)^2} + \frac{\mu_2}{\sigma_2^2} \right) \end{cases}.$$

Then,

$$A \left(\frac{1}{\left(\frac{\sigma_1}{\sqrt{\gamma}}\right)^2} + \frac{1}{\sigma_2^2} \right) = \frac{\mu_1}{\left(\frac{\sigma_1}{\sqrt{\gamma}}\right)^2} + \frac{\mu_2}{\sigma_2^2}$$

and

$$A = \frac{\frac{\mu_1}{\left(\frac{\sigma_1}{\sqrt{\gamma}}\right)^2} + \frac{\mu_2}{\sigma_2^2}}{\frac{1}{\left(\frac{\sigma_1}{\sqrt{\gamma}}\right)^2} + \frac{1}{\sigma_2^2}} = \frac{\frac{\sigma_2^2 \mu_1 + \mu_2 \left(\frac{\sigma_1}{\sqrt{\gamma}}\right)^2}{\sigma_2^2 \left(\frac{\sigma_1}{\sqrt{\gamma}}\right)^2}}{\frac{\sigma_2^2 + \left(\frac{\sigma_1}{\sqrt{\gamma}}\right)^2}{\sigma_2^2 \left(\frac{\sigma_1}{\sqrt{\gamma}}\right)^2}} = \frac{\sigma_2^2 \mu_1 + \mu_2 \frac{\sigma_1^2}{\gamma}}{\sigma_2^2 + \frac{\sigma_1^2}{\gamma}} = \frac{\gamma \sigma_2^2 \mu_1 + \mu_2 \sigma_1^2}{\gamma \sigma_2^2 + \sigma_1^2}.$$

We have,

$$\frac{1}{B^2} = \frac{1}{\left(\frac{\sigma_1}{\sqrt{\gamma}}\right)^2} + \frac{1}{\sigma_2^2} = \frac{\gamma}{\sigma_1^2} + \frac{1}{\sigma_2^2} = \frac{\sigma_2^2 \gamma + \sigma_1^2}{\sigma_1^2 \sigma_2^2}$$

Therefore,

$$B = \frac{\sigma_1 \sigma_2}{\sqrt{\sigma_2^2 \gamma + \sigma_1^2}}.$$

On the other hand,

$$\frac{A^2}{B^2} = \left(\frac{\gamma \sigma_2^2 \mu_1 + \mu_2 \sigma_1^2}{\gamma \sigma_2^2 + \sigma_1^2} \right)^2 \frac{\sigma_2^2 \gamma + \sigma_1^2}{\sigma_1^2 \sigma_2^2} = \frac{(\gamma \sigma_2^2 \mu_1 + \mu_2 \sigma_1^2)^2}{(\gamma \sigma_2^2 + \sigma_1^2) \sigma_1^2 \sigma_2^2}$$

and

$$\begin{aligned} L_3 &= \frac{1}{\sigma_1^\gamma (\sqrt{2\pi})^\gamma} \frac{1}{\sigma_2 \sqrt{2\pi}} \exp \left\{ -\frac{1}{2} \left[\frac{\mu_1^2}{\left(\frac{\sigma_1}{\sqrt{\gamma}}\right)^2} + \frac{\mu_2^2}{\sigma_2^2} \right] \right\} \exp \left\{ \frac{1}{2} \frac{A^2}{B^2} \right\} B \sqrt{2\pi} \\ &= \frac{1}{\sigma_1^\gamma (\sqrt{2\pi})^\gamma} \frac{1}{\sigma_2} \exp \left\{ -\frac{1}{2} \left[\frac{\mu_1^2}{\left(\frac{\sigma_1}{\sqrt{\gamma}}\right)^2} + \frac{\mu_2^2}{\sigma_2^2} \right] \right\} \exp \left\{ \frac{1}{2} \frac{(\gamma \sigma_2^2 \mu_1 + \mu_2 \sigma_1^2)^2}{(\gamma \sigma_2^2 + \sigma_1^2) \sigma_1^2 \sigma_2^2} \right\} \frac{\sigma_1 \sigma_2}{\sqrt{\sigma_2^2 \gamma + \sigma_1^2}} \\ &= \frac{\sigma_1 \sigma_2}{\sqrt{\sigma_2^2 \gamma + \sigma_1^2}} \frac{1}{\sigma_1^\gamma (\sqrt{2\pi})^\gamma} \frac{1}{\sigma_2} \exp \left\{ \frac{1}{2} \left[\frac{(\gamma \sigma_2^2 \mu_1 + \mu_2 \sigma_1^2)^2}{(\gamma \sigma_2^2 + \sigma_1^2) \sigma_1^2 \sigma_2^2} - \frac{\gamma \mu_1^2 \sigma_2^2 + \sigma_1^2 \mu_2^2}{\sigma_2^2 \sigma_1^2} \right] \right\}. \end{aligned}$$

However,

$$\begin{aligned} \frac{(\gamma \sigma_2^2 \mu_1 + \mu_2 \sigma_1^2)^2}{(\gamma \sigma_2^2 + \sigma_1^2) \sigma_1^2 \sigma_2^2} - \frac{\gamma \mu_1^2 \sigma_2^2 + \sigma_1^2 \mu_2^2}{\sigma_2^2 \sigma_1^2} &= \frac{(\gamma \sigma_2^2 \mu_1 + \mu_2 \sigma_1^2)^2 - (\gamma \mu_1^2 \sigma_2^2 + \sigma_1^2 \mu_2^2)(\gamma \sigma_2^2 + \sigma_1^2)}{(\gamma \sigma_2^2 + \sigma_1^2) \sigma_1^2 \sigma_2^2} \\ &= \frac{\gamma^2 \sigma_2^4 \mu_1^2 + \mu_2^2 \sigma_1^4 + 2\gamma \sigma_2^2 \mu_1 \mu_2 \sigma_1^2}{(\gamma \sigma_2^2 + \sigma_1^2) \sigma_1^2 \sigma_2^2} \\ &\quad - \frac{\gamma^2 \mu_1^2 \sigma_2^4 + \gamma \mu_1^2 \sigma_2^2 \sigma_1^2 + \mu_2^2 \gamma \sigma_2^2 \sigma_1^2 + \mu_2^2 \sigma_1^4}{(\gamma \sigma_2^2 + \sigma_1^2) \sigma_1^2 \sigma_2^2} \\ &= \frac{2\gamma \sigma_2^2 \mu_1 \mu_2 \sigma_1^2 - \gamma \mu_1^2 \sigma_2^2 \sigma_1^2 - \mu_2^2 \gamma \sigma_2^2 \sigma_1^2}{(\gamma \sigma_2^2 + \sigma_1^2) \sigma_1^2 \sigma_2^2} \\ &= \frac{\sigma_2^2 \sigma_1^2 \gamma (2\mu_1 \mu_2 - \mu_1^2 - \mu_2^2)}{(\gamma \sigma_2^2 + \sigma_1^2) \sigma_1^2 \sigma_2^2} \\ &= -\frac{\gamma (\mu_1 - \mu_2)^2}{(\gamma \sigma_2^2 + \sigma_1^2)} \end{aligned}$$

Therefore,

$$L_3 = \frac{1}{\sigma_1^{\gamma-1} \sqrt{\sigma_2^2 \gamma + \sigma_1^2}} \frac{1}{(\sqrt{2\pi})^\gamma} \exp\left\{-\frac{1}{2} \frac{\gamma(\mu_1 - \mu_2)^2}{(\gamma\sigma_2^2 + \sigma_1^2)}\right\}.$$

Then,

$$\begin{aligned} R_\gamma(\mathcal{N}(\mu_1, \sigma_1), \mathcal{N}(\mu_2, \sigma_2)) &= \frac{1}{\gamma+1} \ln L_1 + \frac{1}{\gamma(\gamma+1)} \ln L_2 - \frac{1}{\gamma} \ln L_3 \\ &= \frac{1}{\gamma+1} \ln \frac{1}{\sigma_1^\gamma (\sqrt{2\pi})^\gamma} \frac{1}{\sqrt{\gamma+1}} + \frac{1}{\gamma(\gamma+1)} \ln \frac{1}{\sigma_2^\gamma (\sqrt{2\pi})^\gamma} \frac{1}{\sqrt{\gamma+1}} \\ &\quad - \frac{1}{\gamma} \ln \frac{1}{\sigma_1^{\gamma-1} \sqrt{\sigma_2^2 \gamma + \sigma_1^2}} \frac{1}{(\sqrt{2\pi})^\gamma} + \frac{1}{2} \frac{\gamma(\mu_1 - \mu_2)^2}{\gamma(\gamma\sigma_2^2 + \sigma_1^2)} \\ &= \frac{1}{\gamma(\gamma+1)} \left(\ln \frac{\sigma_1^{\gamma-1}}{\sigma_2^\gamma \sqrt{\gamma+1}} \sqrt{\sigma_1^2 + \gamma\sigma_2^2} + \gamma \ln \frac{1}{\sigma_1 \sqrt{\gamma+1}} \sqrt{\sigma_1^2 + \gamma\sigma_2^2} \right) \\ &\quad + \frac{1}{2} \frac{\gamma(\mu_1 - \mu_2)^2}{\gamma(\gamma\sigma_2^2 + \sigma_1^2)} \\ &= \frac{1}{\gamma(\gamma+1)} \ln \frac{1}{\sigma_1 \sigma_2^\gamma} \left(\frac{\sqrt{\sigma_1^2 + \gamma\sigma_2^2}}{\sqrt{\gamma+1}} \right)^{\gamma+1} + \frac{1}{2} \frac{(\mu_1 - \mu_2)^2}{(\gamma\sigma_2^2 + \sigma_1^2)} \end{aligned}$$

For $\gamma \rightarrow 0$ we have,

$$\lim_{\gamma \rightarrow 0} R_\gamma(\mathcal{N}(\mu_1, \sigma_1), \mathcal{N}(\mu_2, \sigma_2)) = \frac{\sigma_2^2 - \sigma_1^2}{2\sigma_1^2} + \ln \frac{\sigma_1}{\sigma_2} + \frac{1}{2} \frac{(\mu_1 - \mu_2)^2}{\sigma_1^2}. \tag{A11}$$

Appendix A.4. Computation of the Nonzero Eigenvalues of $\mathbf{A}_\gamma(\boldsymbol{\theta}_0)\mathbf{B}_\tau(\boldsymbol{\theta}_0)\mathbf{K}_\tau(\boldsymbol{\theta}_0)\mathbf{B}_\tau(\boldsymbol{\theta}_0)$

We know that the matrix $\boldsymbol{\xi}(\boldsymbol{\theta})$ can be expressed as

$$\boldsymbol{\xi}(\boldsymbol{\theta}) = c_\tau(\boldsymbol{\theta})\kappa(\boldsymbol{\theta})$$

with

$$\kappa(\boldsymbol{\theta}) = \int f_\theta(x)^{\tau+1} dx = \frac{1}{\sigma^\tau (\sqrt{2\pi})^\tau \sqrt{1+\tau}}.$$

Then,

$$\boldsymbol{\xi}(\boldsymbol{\theta}) = \frac{1}{\sigma^\tau (\sqrt{2\pi})^\tau \sqrt{1+\tau}} \left(0, -\frac{\tau}{(\tau+1)} \frac{1}{\sigma} \right)^T.$$

Therefore,

$$c_\tau(\boldsymbol{\theta}) = \frac{\boldsymbol{\xi}(\boldsymbol{\theta})}{\kappa(\boldsymbol{\theta})} = \left(0, -\frac{\tau}{(\tau+1)} \frac{1}{\sigma} \right).$$

On the other hand

$$\frac{\partial \log f_{\mu,\sigma}(X_i)}{\partial \mu} = \frac{X_i - \mu}{\sigma^2} \text{ and } \frac{\partial \log f_{\mu,\sigma}(X_i)}{\partial \sigma} = -\frac{1}{\sigma} + \frac{1}{\sigma^3} (X_i - \mu)^2$$

and

$$\mathbf{u}_\theta(X_i) = \left(\frac{X_i - \mu}{\sigma^2}, -\frac{1}{\sigma} + \frac{1}{\sigma^3} (X_i - \mu)^2 \right).$$

Then,

$$\Psi_\tau(X; \theta) = (\Psi_\tau^1(X; \theta), \Psi_\tau^2(X; \theta))$$

is given by

$$\Psi_\tau(X; \theta) = \left(\frac{X - \mu}{\sigma^2} \frac{1}{(\sigma\sqrt{2\pi})^\tau} e^{-\frac{\tau}{2} \left(\frac{X-\mu}{\sigma}\right)^2}, \left(\left(\frac{X - \mu}{\sigma}\right)^2 - \frac{1}{1 + \tau} \right) \frac{1}{\sigma} \frac{1}{(\sigma\sqrt{2\pi})^\tau} e^{-\frac{\tau}{2} \left(\frac{X-\mu}{\sigma}\right)^2} \right)$$

and

$$K_\tau(\theta) = E[\Psi_\tau(X; \theta)\Psi_\tau(X; \theta)^T].$$

Now we obtain the elements of that matrix,

$$\begin{aligned} K_\tau^{11}(\theta) &= E \left[\left(\frac{X - \mu}{\sigma^2} \right)^2 \frac{1}{(\sigma\sqrt{2\pi})^{2\tau}} e^{-\frac{2\tau}{2} \left(\frac{X-\mu}{\sigma}\right)^2} \right] \\ &= \frac{1}{(\sigma\sqrt{2\pi})^{2\tau} (1 + 2\tau)^{3/2} \sigma^2} \end{aligned}$$

$$\begin{aligned} K_\tau^{12}(\theta) &= K_\tau^{21}(\theta) = E \left[\left(\frac{X - \mu}{\sigma^2} \right) \left(\left(\frac{X - \mu}{\sigma}\right)^2 - \frac{1}{1 + \tau} \right) \frac{1}{\sigma} \frac{1}{(\sigma\sqrt{2\pi})^{2\tau}} e^{-\frac{2\tau}{2} \left(\frac{X-\mu}{\sigma}\right)^2} \right] \\ &= 0 \end{aligned}$$

and

$$\begin{aligned} K_\tau^{22}(\theta) &= E \left[\left(\left(\frac{X - \mu}{\sigma}\right)^2 - \frac{1}{1 + \tau} \right)^2 \frac{1}{\sigma^2} \frac{1}{(\sigma\sqrt{2\pi})^{2\tau}} e^{-\frac{2\tau}{2} \left(\frac{X-\mu}{\sigma}\right)^2} \right] \\ &= \frac{1}{\sigma^2} \frac{3\tau^2 + 2 + 4\tau}{(\sigma\sqrt{2\pi})^{2\tau} (1 + 2\tau)^{5/2} (1 + \tau)^2} \end{aligned}$$

and

$$\begin{aligned} K_\tau(\theta) &= \begin{pmatrix} \frac{1}{(\sigma\sqrt{2\pi})^{2\tau} (1+2\tau)^{3/2} \sigma^2} & 0 \\ 0 & \frac{1}{\sigma^2} \frac{3\tau^2+2+4\tau}{(\sigma\sqrt{2\pi})^{2\tau} (1+2\tau)^{5/2} (1+\tau)^2} \end{pmatrix} \\ &= \frac{1}{\sigma^2} \frac{1}{(\sigma\sqrt{2\pi})^{2\tau} (1 + 2\tau)^{3/2}} \begin{pmatrix} 1 & 0 \\ 0 & \frac{3\tau^2+2+4\tau}{(1+\tau)^2 (1+2\tau)} \end{pmatrix}. \end{aligned}$$

Now we obtain the matrix $S_\tau(\theta)$. We have

$$\xi(\theta) = c_\tau(\theta)\kappa(\theta)$$

with

$$\kappa(\theta) = \int f_\theta(x)^{\tau+1} dx = \frac{1}{\sigma^\tau (\sqrt{2\pi})^\tau \sqrt{1 + \tau}}.$$

Then,

$$\xi(\theta) = \frac{1}{\sigma^\tau (\sqrt{2\pi})^\tau \sqrt{1+\tau}} \left(0, -\frac{\tau}{(\tau+1)} \frac{1}{\sigma} \right)^T$$

and

$$\frac{1}{\kappa(\theta)} \xi(\theta) \xi(\theta)^T = \frac{1}{\sigma^{\tau+2} (\sqrt{2\pi})^\tau \sqrt{1+\tau}} \begin{pmatrix} 0 & 0 \\ 0 & \frac{\tau^2}{(\tau+1)^2} \end{pmatrix}.$$

On the other hand

$$J_\tau(\theta) = E \left[\begin{pmatrix} \frac{1}{\sigma^4} (X - \mu)^2 & \frac{1}{\sigma^2} \left(\frac{1}{\sigma} - \frac{1}{\sigma^3} (X - \mu)^2 \right) (X - \mu) \\ \frac{1}{\sigma^2} \left(\frac{1}{\sigma} - \frac{1}{\sigma^3} (X - \mu)^2 \right) (X - \mu) & \left(\frac{1}{\sigma} - \frac{1}{\sigma^3} (X - \mu)^2 \right)^2 \end{pmatrix} \frac{1}{(\sigma\sqrt{2\pi})^\tau} e^{-\frac{\tau}{2} \left(\frac{X-\mu}{\sigma} \right)^2} \right]$$

$$J_\tau^{11}(\theta) = E \left[\frac{1}{\sigma^4} (\mu - X)^2 \frac{1}{(\sigma\sqrt{2\pi})^\tau} e^{-\frac{\tau}{2} \left(\frac{X-\mu}{\sigma} \right)^2} \right] = \frac{1}{\sigma^{\tau+2}} \frac{1}{(\tau+1)^{3/2}} \frac{1}{(\sqrt{2\pi})^\tau}$$

$$J_\tau^{12}(\theta) = J_\tau^{21}(\theta) = 0$$

$$J_\tau^{22}(\theta) = E \left[\left(\frac{1}{\sigma} - \frac{1}{\sigma^3} (\mu - X)^2 \right)^2 \frac{1}{(\sigma\sqrt{2\pi})^\tau} e^{-\frac{\tau}{2} \left(\frac{X-\mu}{\sigma} \right)^2} \right] = \frac{1}{\sigma^{\tau+2}} \frac{1}{(\sqrt{2\pi})^\tau} \frac{1}{\sqrt{1+\tau}} \frac{2+\tau^2}{(1+\tau)^2}$$

Therefore

$$J_\tau(\theta) = \frac{1}{\sigma^{\tau+2}} \frac{1}{(\sqrt{2\pi})^\tau} \frac{1}{\sqrt{1+\tau}} \begin{pmatrix} \frac{1}{1+\tau} & 0 \\ 0 & \frac{2+\tau^2}{(1+\tau)^2} \end{pmatrix}$$

$$\begin{aligned} S_\tau(\theta) &= J_\tau(\theta) - \frac{1}{\kappa(\theta)} \xi(\theta) \xi(\theta)^T \\ &= \frac{1}{\sigma^{\tau+2}} \frac{1}{(\sqrt{2\pi})^\tau} \frac{1}{\sqrt{1+\tau}} \left(\begin{pmatrix} \frac{1}{1+\tau} & 0 \\ 0 & \frac{2+\tau^2}{(1+\tau)^2} \end{pmatrix} - \begin{pmatrix} 0 & 0 \\ 0 & \frac{\tau^2}{(\tau+1)^2} \end{pmatrix} \right) \\ &= \frac{1}{\sigma^{\tau+2}} \frac{1}{(\sqrt{2\pi})^\tau} \frac{1}{\sqrt{1+\tau}} \begin{pmatrix} \frac{1}{1+\tau} & 0 \\ 0 & \frac{2}{(\tau+1)^2} \end{pmatrix} \end{aligned}$$

Now we have,

- The matrix $[G(\theta_0)^T S_\tau(\theta_0)^{-1} G(\theta_0)]^{-1}$ ($G(\theta) = (0, 1)^T$)

$$\begin{aligned} G(\theta_0)^T S_\tau(\theta_0)^{-1} G(\theta_0) &= (0 \ 1) \left(\frac{1}{\sigma^{\tau+2}} \frac{1}{(\sqrt{2\pi})^\tau} \frac{1}{\sqrt{1+\tau}} \begin{pmatrix} \frac{1}{1+\tau} & 0 \\ 0 & \frac{2}{(\tau+1)^2} \end{pmatrix} \right)^{-1} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ &= \frac{1}{2} \sigma^2 \sigma^\tau (\tau+1)^{\frac{5}{2}} (\sqrt{2}\sqrt{\pi})^\tau \end{aligned}$$

- The matrix $Q_\tau(\theta_0) = S_\tau^{-1}(\theta_0) G(\theta_0) [G^T(\theta_0) S_\tau^{-1}(\theta_0) G(\theta_0)]^{-1}$

$$\begin{aligned} \mathbf{Q}_\tau(\boldsymbol{\theta}_0) &= \left(\frac{1}{\sigma^{\tau+2}} \frac{1}{(\sqrt{2\pi})^\tau} \frac{1}{\sqrt{1+\tau}} \begin{pmatrix} \frac{1}{1+\tau} & 0 \\ 0 & \frac{2}{(\tau+1)^2} \end{pmatrix} \right)^{-1} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \left(\frac{1}{2} \sigma^2 \sigma^\tau (\tau+1)^{\frac{5}{2}} (\sqrt{2}\sqrt{\pi})^\tau \right)^{-1} \\ &= \begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{aligned}$$

- The matrix $\mathbf{B}_\tau(\boldsymbol{\theta}_0) = \mathbf{S}_\tau(\boldsymbol{\theta}_0)^{-1} \mathbf{G}(\boldsymbol{\theta}_0) [\mathbf{G}(\boldsymbol{\theta}_0)^T \mathbf{S}_\tau(\boldsymbol{\theta}_0)^{-1} \mathbf{G}(\boldsymbol{\theta}_0)]^{-1} \mathbf{G}(\boldsymbol{\theta}_0)^T \mathbf{S}_\tau(\boldsymbol{\theta}_0)^{-1} = \mathbf{Q}_\tau(\boldsymbol{\theta}_0) \mathbf{G}(\boldsymbol{\theta}_0)^T \mathbf{S}_\tau(\boldsymbol{\theta}_0)^{-1}$

$$\begin{aligned} \mathbf{B}_\tau(\boldsymbol{\theta}_0) &= \mathbf{Q}_\tau(\boldsymbol{\theta}_0) \mathbf{G}^T(\boldsymbol{\theta}_0) \mathbf{S}_\tau^{-1}(\boldsymbol{\theta}_0) = \begin{pmatrix} 0 \\ 1 \end{pmatrix} (0 \quad 1) \left(\frac{1}{\sigma^{\tau+2}} \frac{1}{(\sqrt{2\pi})^\tau} \frac{1}{\sqrt{1+\tau}} \begin{pmatrix} \frac{1}{1+\tau} & 0 \\ 0 & \frac{2}{(\tau+1)^2} \end{pmatrix} \right)^{-1} \\ &= \begin{pmatrix} 0 & 0 \\ 0 & \frac{1}{2} \sigma^2 \sigma^\tau (\tau+1)^{\frac{5}{2}} (\sqrt{2}\sqrt{\pi})^\tau \end{pmatrix} \end{aligned}$$

- The matrix $\mathbf{M}_{\gamma,\tau}(\boldsymbol{\theta}_0) = \frac{\mathbf{S}_\gamma(\boldsymbol{\theta}_0)}{\kappa_\gamma(\boldsymbol{\theta}_0)} \mathbf{B}_\tau(\boldsymbol{\theta}_0) \mathbf{K}_\tau(\boldsymbol{\theta}_0) \mathbf{B}_\tau(\boldsymbol{\theta}_0)$

$$\begin{aligned} \mathbf{M}_{\gamma,\tau}(\boldsymbol{\theta}_0) &= \frac{\sigma^\gamma (\sqrt{2\pi})^\gamma \sqrt{1+\gamma}}{\sigma^{\gamma+2}} \frac{1}{(\sqrt{2\pi})^\gamma} \frac{1}{\sqrt{1+\gamma}} \begin{pmatrix} \frac{1}{1+\gamma} & 0 \\ 0 & \frac{2}{(\gamma+1)^2} \end{pmatrix} \\ &\quad \times \begin{pmatrix} 0 & 0 \\ 0 & \frac{1}{2} \sigma^2 \sigma^\tau (\tau+1)^{\frac{5}{2}} (\sqrt{2}\sqrt{\pi})^\tau \end{pmatrix} \\ &\quad \times \frac{1}{\sigma^2} \frac{1}{(\sigma\sqrt{2\pi})^{2\tau} (1+2\tau)^{3/2}} \begin{pmatrix} 1 & 0 \\ 0 & \frac{3\tau^2+2+4\tau}{(1+\tau)^2(1+2\tau)} \end{pmatrix} \\ &\quad \times \begin{pmatrix} 0 & 0 \\ 0 & \frac{1}{2} \sigma^2 \sigma^\tau (\tau+1)^{\frac{5}{2}} (\sqrt{2}\sqrt{\pi})^\tau \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 \\ 0 & \frac{1}{2} \frac{(\tau+1)^3}{(\gamma+1)^2 (2\tau+1)^{\frac{5}{2}}} (3\tau^2 + 4\tau + 2) \end{pmatrix}. \end{aligned}$$

References

- Beran, R. Minimum Hellinger distance estimates for parametric models. *Ann. Stat.* **1977**, *5*, 445–463. [\[CrossRef\]](#)
- Tamura, R.N.; Boos, D.D. Minimum Hellinger distance estimation for multivariate location and covariance. *J. Am. Stat. Assoc.* **1986**, *81*, 223–229. [\[CrossRef\]](#)
- Simpson, D.G. Minimum Hellinger distance estimation for the analysis of count data. *J. Am. Stat. Assoc.* **1987**, *82*, 802–807. [\[CrossRef\]](#)
- Simpson, D.G. Hellinger deviance tests: Efficiency, breakdown points, and examples. *J. Am. Stat. Assoc.* **1989**, *84*, 107–113. [\[CrossRef\]](#)
- Lindsay, B.G. Efficiency versus robustness: The case for minimum Hellinger distance and related methods. *Ann. Stat.* **1994**, *22*, 1081–1114. [\[CrossRef\]](#)
- Pardo, L. *Statistical Inference Based on Divergence Measures*; Chapman & Hall/CRC: Boca de Raton, FL, USA, 2006.
- Basu, A.; Shioya, H.; Park, C. *Statistical Inference: The minimum Distance Approach*; Chapman & Hall/CRC Press: Boca de Raton, FL, USA, 2011.
- Broniatowski, M.; Toma, A.; Vajda, I. Decomposable pseudodistances and applications in statistical estimation. *J. Stat. Plan. Inference* **2012**, *142*, 2574–2585. [\[CrossRef\]](#)
- Castilla, E.; Jaenada, M.; Pardo, L. Estimation and testing on independent not identically distributed observations based on Rényi's pseudodistances. *IEEE Trans. Inf. Theory* **2022**, *in press*. [\[CrossRef\]](#)
- Silvey, S.D. *Reprinting, Monographs on Statistical Subjects*; Chapman and Hall: London, UK, 1975.

11. Basu, A.; Mandal, A.; Martin, N.; Pardo, L. Testing Composite Hypothesis Based on the Density Power Divergence. *Sankhya B Indian J. Stat.* **2018**, *80*, 222–262. [[CrossRef](#)]
12. Ghosh, A. Influence function analysis of the restricted minimum divergence estimators: A general form. *Electron. J. Stat.* **2015**, *9*, 1017–1040. [[CrossRef](#)]
13. Jones, M.C.; Hjort, N.L.; Harris, I.R.; Basu, A. A comparison of related density-based minimum divergence estimators. *Biometrika* **2001**, *88*, 865–873. [[CrossRef](#)]
14. Fujisawa, H.; Eguchi, S. Robust parameter estimation with a small bias against heavy contamination. *J. Multivariate Anal.* **2008**, *99*, 2053–2081. [[CrossRef](#)]
15. Castilla, E.; Jaenada, M.; Martin, N.; Pardo, L. Robust approach for comparing two dependent normal populations through Wald-type tests based on Rényi's pseudodistance estimators. *arXiv* **2022**, arXiv:2202.00982.
16. Toma, A.; Leoni-Aubin, S. Robust tests based on dual divergence estimators and saddlepoint approximations. *J. Multivariate Anal.* **2010**, *101*, 1143–1155. [[CrossRef](#)]
17. Toma, A.; Karagrigoriou, A.; Trentou, P. Robust model selection criteria based on pseudodistances. *Entropy* **2020**, *22*, 304. [[CrossRef](#)] [[PubMed](#)]
18. Castilla, E.; Martin, N.; Muñoz, S.; Pardo, L. Robust Wald-type tests based on Minimum Rényi Pseudodistance Estimators for the Multiple Regression Model. *J. Stat. Comput. Simul.* **2020**, *14*, 2592–2613. [[CrossRef](#)]
19. Castilla, E.; Ghosh, A.; Jaenada, M.; Pardo, L. On regularization methods based on Rényi's pseudodistances for sparse high-dimensional linear regression models. *arXiv* **2022**, arXiv:2202.00982.
20. Jaenada, M.; Pardo, L. The minimum Rényi's Pseudodistances estimators for Generalized Linear Models. In *Data Analysis and Related Applications: Theory and Practice*; Proceeding of the ASMDA; Wiley: Athens, Greece, 2021.
21. Jaenada, M.; Pardo, L. Robust Statistical Inference in Generalized Linear Models Based on Minimum Rényi's Pseudodistance Estimators. *Entropy* **2022**, *24*, 123. [[CrossRef](#)]
22. Hampel, F.R.; Ronchetti, E.; Rousseauw, P.J.; Stahel, W. *Robust Statistics: The Approach Based on Influence Functions*; John Wiley & Sons: Hoboken, NJ, USA, 1986.
23. Rao, C.R. Score test: Historical review and recent developments. In *Advances in Ranking and Selection, Multiple Comparisons, and Reliability*; Birkhäuser: Boston, MA, USA, 2005; pp. 3–20.
24. Basu, A.; Ghosh, A.; Martin, N.; Pardo, L. A Robust Generalization of the Rao Test. *J. Bus. Econ. Stat.* **2021**, *40*, 868–879. [[CrossRef](#)]
25. Warwick, J.; Jones, M.C. Choosing a robustness tuning parameter. *J. Stat. Comput. Simul.* **2005**, *75*, 581–588. [[CrossRef](#)]
26. Basak, S.; Basu, A.; Jones, M.C. On the optimal density power divergence tuning parameter. *J. Appl. Stat.* **2021**, *48*, 536–556. [[CrossRef](#)]
27. Welch, W.J. Rerandomizing the median in matched-pairs designs. *Biometrika* **1987**, *74*, 609–614. [[CrossRef](#)]
28. Basu, A.; Mandal, A.; Martin, N.; Pardo, L. Testing statistical hypotheses based on the density power divergence. *Ann. Inst. Stat. Math.* **2013**, *65*, 319–348. [[CrossRef](#)]
29. Darwin, C. *The Effects of Cross and Self Fertilisation in the Vegetable Kingdom*; AMS Press Inc.: New York, NY, USA, 1877.
30. Dik, J.J.; de Gunst, M.C.M. The Distribution of General Quadratic Forms in Norms. *Stat. Neerl.* **1985**, *39*, 14–26. [[CrossRef](#)]
31. Harville, D.A. *Matrix Algebra from a Statistician's Perspective*; Springer: New York, NY, USA, 2008.