Journal of Computational and Applied Mathematics 224 (2009) 193–203

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

journal homepage: www.elsevier.com/locate/cam

# Confidence sets and coverage probabilities based on preliminary estimators in logistic regression models

M.L. Menéndez [a], L. Pardo [b], M.C. Pardo [b,*]

[a] Department of Applied Mathematics, E.T.S.A.M., Technical University of Madrid, Spain
[b] Department of Statistics and O.R. I, Complutense University of Madrid, Spain

ARTICLE INFO

*Article history:*
Received 28 October 2007
Received in revised form 11 April 2008

*MSC:*
62F25
62F30
62F12
62J99

*Keywords:*
Preliminary test estimator
Asymptotic coverage probability
Minimum phi-divergence estimator
Recentered confidence set

ABSTRACT

In this paper we present recentered confidence sets for the parameters of a logistic regression model based on preliminary minimum $\phi$-divergence estimators. Asymptotic coverage probabilities are given as well as a simulation study in order to analyze the coverage probabilities for small and moderate sample sizes.



## 1. Introduction

Let $Y_i$, $i = 1, \ldots, n$, be independent binomial random variables with parameters $\pi_i$ and $n_i$, $i = 1, \ldots, n$. We shall assume that the parameters $\pi_i = \Pr(Y_i = 1)$, $i = 1, \ldots, n$, depend on the unknown parameters $\boldsymbol{\beta} = (\beta_0, \ldots, \beta_k)^{\mathrm{T}}$, $\beta_i \in (-\infty, \infty)$ and explanatory variables $\boldsymbol{x}_i^{\mathrm{T}} = (x_{i0}, \ldots, x_{ik})$, $x_{i0} = 1$, $i = 1, \ldots, n$ through the linear predictor

$$\operatorname{logit}(\pi_i) = \sum_{j=0}^{k} x_{ij}\beta_j, \quad i = 1, \ldots, n \tag{1}$$

where $\operatorname{logit}(p) = \log\left(p/(1-p)\right)$. In the following we shall denote the binomial parameter $\pi_i$ by $\pi_i \equiv \pi\left(\boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{\beta}\right)$ and by $\boldsymbol{X}$ the $n \times (k+1)$ matrix with rows $\boldsymbol{x}_i$, $i = 1, \ldots, n$. We also assume that $\operatorname{rank}(\boldsymbol{X}) = k+1$.

In [4] a preliminary test estimator for $\boldsymbol{\beta}$,$\widehat{\boldsymbol{\beta}}^{\mathrm{Pre}}_{\phi_1,\phi_2}$ (see (8) in Section 2) was considered. This estimator is based on the restricted $\widehat{\boldsymbol{\beta}}^{H_0}_{\phi_2}$ (see (7) in Section 2) and the unrestricted $\widehat{\boldsymbol{\beta}}_{\phi_2}$ (see (2) in Section 2) minimum $\phi_2$-divergence estimators of $\boldsymbol{\beta}$. An important problem for the point estimation of $\boldsymbol{\beta}$ is to provide associated confidence sets. In this paper we consider asymptotic recentered confidence sets for $\boldsymbol{\beta}$ based on $\widehat{\boldsymbol{\beta}}^{\mathrm{Pre}}_{\phi_1,\phi_2}$, $\widehat{\boldsymbol{\beta}}^{H_0}_{\phi_2}$ and $\widehat{\boldsymbol{\beta}}_{\phi_2}$ and we study their coverage probabilities.

In Section 2 we present some notation as well as some preliminary results that will be necessary in the paper. Section 3 is devoted to the definition of recentered confidence sets as well as an analytical study of their asymptotic coverage probabilities. Finally, in Section 4 a simulation study is carried out in order to analyze the coverage probabilities for small and moderate sample sizes and different choices on the functions $\phi_1$ and $\phi_2$.

* Corresponding address: Department of Statistics and O.R. I, Complutense University of Madrid, Plaza de Ciencias, 3, 28040 Madrid, Spain. Tel.: +34 91 3944473; fax: +34 91 3944606.
*E-mail address:* mcapardo@mat.ucm.es (M.C. Pardo).

0377-0427/$ – see front matter 
doi:10.1016/j.cam.2008.04.023

## 2. Background and notation

We denote by $y_i$ the number of "successes" associated with the binomial random variable $Y_i$, $i = 1, \ldots, n$. Minimum $\phi_2$-divergence estimator ($M\phi_2E$) of $\boldsymbol{\beta}$, $\widehat{\boldsymbol{\beta}}_{\phi_2} = \widehat{\boldsymbol{\beta}}_{\phi_2}(Y_1, \ldots, Y_n)$ is defined as

$$\widehat{\boldsymbol{\beta}}_{\phi_2} = \arg\min_{\boldsymbol{\beta}\in\Theta} \sum_{i=1}^{n} n_i D_{\phi_2}\left(\widehat{\boldsymbol{p}}_i, \boldsymbol{\pi}_i(\boldsymbol{\beta})\right) \tag{2}$$

where

$$\widehat{\boldsymbol{p}}_i = \left(\frac{y_i}{n_i}, \frac{n_i - y_i}{n_i}\right)^{\mathrm{T}} \quad \text{and} \quad \boldsymbol{\pi}_i(\boldsymbol{\beta}) = \left(\pi\left(\boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{\beta}\right), 1 - \pi\left(\boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{\beta}\right)\right)^{\mathrm{T}}, \quad i = 1, \ldots, n, \tag{3}$$

$\Theta = \{\boldsymbol{\beta} = (\beta_0, \beta_1, \ldots, \beta_k) : \beta_j \in (-\infty, +\infty), j = 0, \ldots, k\}$ and $D_{\phi_2}(\widehat{\boldsymbol{p}}_i, \boldsymbol{\pi}_i(\boldsymbol{\beta}))$ is the $\phi_2$-divergence measure between the probability vectors $\widehat{\boldsymbol{p}}_i$ and $\boldsymbol{\pi}_i(\boldsymbol{\beta})$, given by

$$D_{\phi_2}\left(\widehat{\boldsymbol{p}}_i, \boldsymbol{\pi}_i(\boldsymbol{\beta})\right) \equiv \pi\left(\boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{\beta}\right)\phi_2\left(\frac{y_i}{\pi\left(\boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{\beta}\right)n_i}\right) + \left(1 - \pi\left(\boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{\beta}\right)\right)\phi_2\left(\frac{n_i - y_i}{\left(1 - \pi\left(\boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{\beta}\right)\right)n_i}\right), \tag{4}$$

$\phi_2 \in \Phi$, $\Phi$ is the class of all convex functions $\phi_2(x)$, $x > 0$, such that at $x = 1$, $\phi_2(1) = \phi_2'(1) = 0$, $\phi_2''(1) > 0$. In (4) we shall assume the conventions $0\phi_2(0/0) = 0$ and $0\phi_2(p/0) = p\lim_{u\to\infty}\phi_2(u)/u$. For a systematic study of $\phi_2$-divergences see Pardo [6].

For $\phi_2(x) = x\log x - x + 1$ we obtain in (4) the Kullback–Leibler divergence,

$$D_{Kull}\left(\widehat{\boldsymbol{p}}_i, \boldsymbol{\pi}_i(\boldsymbol{\beta})\right) = y_i \log\frac{y_i}{\pi\left(\boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{\beta}\right)n_i} + (n_i - y_i)\log\frac{(n_i - y_i)}{\left(1 - \pi\left(\boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{\beta}\right)n_i\right)}$$

and it is immediately seen that

$$\sum_{i=1}^{n} n_i D_{Kull}\left(\widehat{\boldsymbol{p}}_i, \boldsymbol{\pi}_i(\boldsymbol{\beta})\right) = -l(\boldsymbol{\beta}) + k,$$

where $l(\boldsymbol{\beta})$ is the loglikelihood function defined by

$$l(\boldsymbol{\beta}) = \sum_{i=1}^{n}\log\left(\pi\left(\boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{\beta}\right)^{y_i}\left(1 - \pi\left(\boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{\beta}\right)\right)^{n_i - y_i}\right).$$

Therefore, the maximum likelihood estimator defined by $\widehat{\boldsymbol{\beta}} = \arg\max_{\boldsymbol{\beta}\in\Theta} l(\boldsymbol{\beta})$ can also be defined by

$$\widehat{\boldsymbol{\beta}} = \arg\min_{\boldsymbol{\beta}\in\Theta}\sum_{i=1}^{n} n_i D_{Kull}\left(\widehat{\boldsymbol{p}}_i, \boldsymbol{\pi}_i(\boldsymbol{\beta})\right)$$

and the minimum $\phi_2$-divergence estimator defined in (2) is a natural extension of the maximum likelihood estimator.

We denote $N = \sum_{i=1}^{n} n_i$,

$$\boldsymbol{W}_N(\boldsymbol{\beta}) = \operatorname{diag}\left(\left(\boldsymbol{C}_i(\boldsymbol{\beta})\right)_{i=1,\ldots,n}^{\mathrm{T}}\right)\operatorname{diag}\left(\left(\boldsymbol{C}_i(\boldsymbol{\beta})\right)_{i=1,\ldots,n}\right)$$

with

$$\boldsymbol{C}_i(\boldsymbol{\beta}) = \left(\frac{n_i}{N}\pi(\boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{\beta})\left(1 - \pi(\boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{\beta})\right)\right)^{1/2}\begin{pmatrix}\left(1 - \pi(\boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{\beta})\right)^{1/2} \\ -\pi(\boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{\beta})^{1/2}\end{pmatrix}, \quad i = 1, \ldots, n. \tag{5}$$

In the following we shall assume $\lambda_i = \lim_{N\to\infty} n_i/N$, $i = 1, \ldots, n$. Under the assumption that $\pi$ has continuous second partial derivatives in a neighborhood of the true value of the parameter $\boldsymbol{\beta}_0$, and $\phi_2 \in \Phi$ is twice differentiable at $x > 0$, $\widehat{\boldsymbol{\beta}}_{\phi_2}$ verifies

$$\sqrt{N}\left(\widehat{\boldsymbol{\beta}}_{\phi_2} - \boldsymbol{\beta}_0\right) \underset{N\to\infty}{\overset{L}{\to}} \mathcal{N}\left(\boldsymbol{0}, \left(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}(\boldsymbol{\beta}_0)\boldsymbol{X}\right)^{-1}\right), \tag{6}$$

where $\boldsymbol{W}(\boldsymbol{\beta}_0) = \lim_{N\to\infty}\boldsymbol{W}_N(\boldsymbol{\beta}_0)$. For more properties about $\widehat{\boldsymbol{\beta}}_{\phi_2}$ see Pardo et al. [5].

Now we assume that we have the additional information that $\boldsymbol{\beta} \in \Theta_0 = \{\boldsymbol{\beta} \in \Theta / \boldsymbol{K}^{\mathrm{T}}\boldsymbol{\beta} = \boldsymbol{m}\}$, where $\boldsymbol{K}^{\mathrm{T}}$ is any matrix of $r$ rows and $k+1$ columns and $\boldsymbol{m}$ is a vector of order $r$ of specified constants. The minimum $\phi_2$-divergence estimator restricted to $\Theta_0$ is given by

$$\widehat{\boldsymbol{\beta}}_{\phi_2}^{H_0} \equiv \arg\min_{\boldsymbol{\beta}\in\Theta_0}\sum_{i=1}^{n} n_i D_{\phi_2}\left(\widehat{\boldsymbol{p}}_i, \boldsymbol{\pi}_i(\boldsymbol{\beta})\right). \tag{7}$$

We refer to it as the restricted minimum $\phi_2$-divergence estimator (RM$\phi_2$E) of $\boldsymbol{\beta} \in \Theta_0$. The RM$\phi_2$E verifies

$$\sqrt{N}\left(\widehat{\boldsymbol{\beta}}_{\phi_2}^{H_0} - \boldsymbol{\beta}_0\right) \underset{N\to\infty}{\overset{L}{\to}} \mathcal{N}\left(\boldsymbol{0}, \boldsymbol{H}^*(\boldsymbol{\beta}_0)\left(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}(\boldsymbol{\beta}_0)\boldsymbol{X}\right)^{-1}\right),$$

where $\boldsymbol{H}^*(\boldsymbol{\beta}_0) = \boldsymbol{I} - \left(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}(\boldsymbol{\beta}_0)\boldsymbol{X}\right)^{-1}\boldsymbol{K}\left(\boldsymbol{K}^{\mathrm{T}}\left(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}(\boldsymbol{\beta}_0)\boldsymbol{X}\right)^{-1}\boldsymbol{K}\right)^{-1}\boldsymbol{K}^{\mathrm{T}}$.

If we consider $\phi_2(x) = x\log x - x + 1$ in (7) we obtain the classical restricted maximum likelihood estimator.

In [3] in order to test the compatibility of the restricted and the unrestricted minimum $\phi_2$-divergence estimators $\widehat{\boldsymbol{\beta}}_{\phi_2}$ and $\widehat{\boldsymbol{\beta}}^{H_0}_{\phi_2}$, i.e., for testing

$$H_0 : \boldsymbol{K}^{\mathrm{T}}\boldsymbol{\beta} = \boldsymbol{m} \quad \text{versus} \quad H_1 : \boldsymbol{K}^{\mathrm{T}}\boldsymbol{\beta} \neq \boldsymbol{m}$$

the following family of $\phi$-divergence statistics was considered

$$T_N^{\phi_1,\phi_2} = \frac{2}{\phi_1''(1)} \sum_{i=1}^{n} n_i D_{\phi_1}\left(\boldsymbol{\pi}_i(\widehat{\boldsymbol{\beta}}_{\phi_2}), \boldsymbol{\pi}_i(\widehat{\boldsymbol{\beta}}^{H_0}_{\phi_2})\right),$$

where $\boldsymbol{\pi}_i(\widehat{\boldsymbol{\beta}}_{\phi_2})$ and $\boldsymbol{\pi}_i(\widehat{\boldsymbol{\beta}}^{H_0}_{\phi_2})$ are obtained from (3) replacing $\boldsymbol{\beta}$ by $\widehat{\boldsymbol{\beta}}_{\phi_2}$ and $\widehat{\boldsymbol{\beta}}^{H_0}_{\phi_2}$ respectively. We can observe that the statistic $T_N^{\phi_1,\phi_2}$ involves two functions $\phi_1$ and $\phi_2$. The function $\phi_2$ is used to compute the minimum $\phi_2$-divergence estimators $\widehat{\boldsymbol{\beta}}_{\phi_2}$ and $\widehat{\boldsymbol{\beta}}^{H_0}_{\phi_2}$, while $\phi_1$ is used to calculate the "distance" between the two probability vectors.

It is interesting to observe that for $\phi_2(x) = \phi_1(x) = x\log x - x + 1$ we obtain $T_N^{\phi_1,\phi_2} = LR + o_P(1)$, where $LR$ is the likelihood-ratio test.

If we accept $H_0$ we choose the RM$\phi_2$E and if we reject $H_0$ we choose the M$\phi_2$E, i.e., the preliminary minimum ($\phi_1$,$\phi_2$)-divergence estimator,

$$\widehat{\boldsymbol{\beta}}^{\mathrm{Pre}}_{\phi_1,\phi_2} = \widehat{\boldsymbol{\beta}}^{H_0}_{\phi_2} I_{(0,\chi^2_{r,\alpha})}(T_N^{\phi_1,\phi_2}) + \widehat{\boldsymbol{\beta}}_{\phi_2} I_{[\chi^2_{r,\alpha},\infty)}(T_N^{\phi_1,\phi_2})$$

or equivalently

$$\widehat{\boldsymbol{\beta}}^{\mathrm{Pre}}_{\phi_1,\phi_2} = \widehat{\boldsymbol{\beta}}^{H_0}_{\phi_2} + \left(\widehat{\boldsymbol{\beta}}_{\phi_2} - \widehat{\boldsymbol{\beta}}^{H_0}_{\phi_2}\right) I_{[\chi^2_{r,\alpha},\infty)}(T_N^{\phi_1,\phi_2}), \tag{8}$$

where $I_A(y)$ denotes an indicator function taking the value 1 if $y \in A$ and 0 if $y \notin A$. Hence, the preliminary estimator depends on $\phi_1$ and $\phi_2$.

In [4] the asymptotic bias and the asymptotic distributional quadratic risk for $\widehat{\boldsymbol{\beta}}^{\mathrm{Pre}}_{\phi_1,\phi_2}$, $\widehat{\boldsymbol{\beta}}_{\phi_2}$ and $\widehat{\boldsymbol{\beta}}^{H_0}_{\phi_2}$ were studied. A closely related problem is the confidence sets based on the preliminary test estimators. Our interest in this paper is to provide asymptotic recentered confidence sets based on $\widehat{\boldsymbol{\beta}}^{\mathrm{Pre}}_{\phi_1,\phi_2}$, $\widehat{\boldsymbol{\beta}}_{\phi_2}$ and $\widehat{\boldsymbol{\beta}}^{H_0}_{\phi_2}$, for contiguous alternative hypotheses and to obtain the asymptotic expressions for their coverage probabilities. Whereas exact expressions have been studied in the multinomial distributional problem, [1] among others, in logistic regression models it is not possible to obtain exact results. Recentered confidence sets are well documented in [7] for different statistical problems.

## 3. Coverage probabilities: An analytical study

We define the recentered confidence set based on the estimator $\widehat{\boldsymbol{\beta}}^{*}_{\phi}$, where $\widehat{\boldsymbol{\beta}}^{*}_{\phi}$ is equal to $\widehat{\boldsymbol{\beta}}^{\mathrm{Pre}}_{\phi_1,\phi_2}$, $\widehat{\boldsymbol{\beta}}_{\phi_2}$ or $\widehat{\boldsymbol{\beta}}^{H_0}_{\phi_2}$, as

$$C_{\boldsymbol{\beta}}\left(\widehat{\boldsymbol{\beta}}^{*}_{\phi}\right) = \left\{\boldsymbol{\beta} : N\left\|\boldsymbol{\beta} - \widehat{\boldsymbol{\beta}}^{*}_{\phi}\right\|^2_{\mathbf{X}^{\mathrm{T}}\mathbf{W}_N(\widehat{\boldsymbol{\beta}}_{\phi_2})\mathbf{X}} \leq \chi^2_{k+1,\alpha}\right\},$$

where $\|\boldsymbol{Y}\|^2_C = \boldsymbol{Y}^{\mathrm{T}}\boldsymbol{C}\boldsymbol{Y}$.

We are going to see the asymptotic behavior of $\widehat{\boldsymbol{\beta}}^{\mathrm{Pre}}_{\phi_1,\phi_2}$, $\widehat{\boldsymbol{\beta}}_{\phi_2}$ and $\widehat{\boldsymbol{\beta}}^{H_0}_{\phi_2}$ under fixed alternative hypotheses defined by

$$H_1 : \boldsymbol{K}^{\mathrm{T}}\boldsymbol{\beta} = \boldsymbol{m} + \boldsymbol{s}$$

with $\boldsymbol{s} \in \mathbb{R}^r$ and fixed. The main results are presented in the following theorem:

**Theorem 1.** *Under fixed alternative hypotheses $H_1 : \boldsymbol{K}^{\mathrm{T}}\boldsymbol{\beta} = \boldsymbol{m} + \boldsymbol{s}$ with $\boldsymbol{s} \in \mathbb{R}^r$, we have:*

(a) $\sqrt{N}\left(\widehat{\boldsymbol{\beta}}^{\mathrm{Pre}}_{\phi_1,\phi_2} - \boldsymbol{\beta}\right) = \sqrt{N}\left(\widehat{\boldsymbol{\beta}}_{\phi_2} - \boldsymbol{\beta}\right) + o_P(1)$.

(b) $\sqrt{N}\left(\widehat{\boldsymbol{\beta}}^{H_0}_{\phi_2} - \boldsymbol{\beta}\right)$ *has a degenerate asymptotic distribution.*

**Proof.** (a) First we are going to establish that $T_N^{\phi_1,\phi_2} \to \infty$ as $N \to \infty$. On the one hand

$$\begin{aligned}
\sqrt{N}\left(\boldsymbol{K}^{\mathrm{T}}\widehat{\boldsymbol{\beta}}_{\phi_2} - \boldsymbol{m}\right) &= \sqrt{N}\boldsymbol{K}^{\mathrm{T}}\widehat{\boldsymbol{\beta}}_{\phi_2} - \sqrt{N}\boldsymbol{m} - \sqrt{N}\boldsymbol{K}^{\mathrm{T}}\boldsymbol{\beta} + \sqrt{N}\boldsymbol{K}^{\mathrm{T}}\boldsymbol{\beta} \\
&= \sqrt{N}\boldsymbol{K}^{\mathrm{T}}\left(\widehat{\boldsymbol{\beta}}_{\phi_2} - \boldsymbol{\beta}\right) + \sqrt{N}\left(\boldsymbol{K}^{\mathrm{T}}\boldsymbol{\beta} - \boldsymbol{m}\right) \\
&= \sqrt{N}\boldsymbol{K}^{\mathrm{T}}\left(\widehat{\boldsymbol{\beta}}_{\phi_2} - \boldsymbol{\beta}\right) + \sqrt{N}\boldsymbol{s}
\end{aligned}$$

and in [3] we obtain,

$$\begin{aligned} T_N^{\phi_1,\phi_2} &= \sqrt{N}\left(\boldsymbol{K}^{\mathrm{T}}\widehat{\boldsymbol{\beta}}_{\phi_2}-\boldsymbol{m}\right)^{\mathrm{T}}(\boldsymbol{K}^{\mathrm{T}}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}_N(\boldsymbol{\beta})\boldsymbol{X})^{-1}\boldsymbol{K})^{-1}\sqrt{N}\left(\boldsymbol{K}^{\mathrm{T}}\widehat{\boldsymbol{\beta}}_{\phi_2}-\boldsymbol{m}\right)+o_P(1) \\ &= \sqrt{N}\left(\widehat{\boldsymbol{\beta}}_{\phi_2}-\boldsymbol{\beta}\right)^{\mathrm{T}}\boldsymbol{K}(\boldsymbol{K}^{\mathrm{T}}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}_N(\boldsymbol{\beta})\boldsymbol{X})^{-1}\boldsymbol{K})^{-1}\sqrt{N}\boldsymbol{K}^{\mathrm{T}}\left(\widehat{\boldsymbol{\beta}}_{\phi_2}-\boldsymbol{\beta}\right) \\ &\quad+\sqrt{N}\boldsymbol{s}^{\mathrm{T}}(\boldsymbol{K}^{\mathrm{T}}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}_N(\boldsymbol{\beta})\boldsymbol{X})^{-1}\boldsymbol{K})^{-1}\sqrt{N}\boldsymbol{s}+2N\left(\widehat{\boldsymbol{\beta}}_{\phi_2}-\boldsymbol{\beta}\right)^{\mathrm{T}}\boldsymbol{K}(\boldsymbol{K}^{\mathrm{T}}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}_N(\boldsymbol{\beta})\boldsymbol{X})^{-1}\boldsymbol{K})^{-1}\boldsymbol{s}. \end{aligned}$$

It is not difficult to see that

$$\sqrt{N}\boldsymbol{K}(\boldsymbol{K}^{\mathrm{T}}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}_N(\boldsymbol{\beta})\boldsymbol{X})^{-1}\boldsymbol{K})^{-1/2}\left(\widehat{\boldsymbol{\beta}}_{\phi_2}-\boldsymbol{\beta}\right)\xrightarrow[N\to\infty]{L}\mathcal{N}(\boldsymbol{0},\boldsymbol{I})$$

and

$$N\boldsymbol{s}^{\mathrm{T}}(\boldsymbol{K}^{\mathrm{T}}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}_N(\boldsymbol{\beta})\boldsymbol{X})^{-1}\boldsymbol{K})^{-1}\boldsymbol{s}\xrightarrow[N\to\infty]{L}\infty$$

$$2N\left(\widehat{\boldsymbol{\beta}}_{\phi_2}-\boldsymbol{\beta}\right)^{\mathrm{T}}\boldsymbol{K}(\boldsymbol{K}^{\mathrm{T}}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}_N(\boldsymbol{\beta})\boldsymbol{X})^{-1}\boldsymbol{K})^{-1}\boldsymbol{s}\xrightarrow[N\to\infty]{L}\infty.$$

Therefore $T_N^{\phi_1,\phi_2}\to\infty$.

In order to establish (a) we consider, based on (8), the quadratic difference

$$\begin{aligned} N\left\|\widehat{\boldsymbol{\beta}}^{\mathrm{Pre}}_{\phi_1,\phi_2}-\widehat{\boldsymbol{\beta}}_{\phi_2}\right\|^2_{\mathbf{X}^{\mathrm{T}}\mathbf{W}_N(\boldsymbol{\beta})\mathbf{X}} &= N\left(\widehat{\boldsymbol{\beta}}_{\phi_2}-\widehat{\boldsymbol{\beta}}^{H_0}_{\phi_2}\right)^{\mathrm{T}}\mathbf{X}^{\mathrm{T}}\mathbf{W}_N(\boldsymbol{\beta})\mathbf{X}\left(\widehat{\boldsymbol{\beta}}_{\phi_2}-\widehat{\boldsymbol{\beta}}^{H_0}_{\phi_2}\right)I_{[0,\chi^2_{r,\alpha})}(T_N^{\phi_1,\phi_2}) \\ &= \left[T_N^{\phi_1,\phi_2}+o_P(1)\right]I_{[0,\chi^2_{r,\alpha})}(T_N^{\phi_1,\phi_2})\le\left[\chi^2_{r,\alpha}+o_P(1)\right]I_{[0,\chi^2_{r,\alpha})}(T_N^{\phi_1,\phi_2}). \end{aligned}$$

Therefore

$$\lim_{N\to\infty}E\left[N\left\|\widehat{\boldsymbol{\beta}}^{\mathrm{Pre}}_{\phi_1,\phi_2}-\widehat{\boldsymbol{\beta}}_{\phi_2}\right\|^2_{\mathbf{X}^{\mathrm{T}}\mathbf{W}_N(\boldsymbol{\beta})\mathbf{X}}\right]\le\lim_{N\to\infty}E\left[\left[\chi^2_{r,\alpha}+o_P(1)\right]I_{[0,\chi^2_{r,\alpha})}(T_N^{\phi_1,\phi_2})\right]=0$$

which means

$$\sqrt{N}\widehat{\boldsymbol{\beta}}^{\mathrm{Pre}}_{\phi_1,\phi_2}-\sqrt{N}\widehat{\boldsymbol{\beta}}_{\phi_2}\underset{N\to\infty}{\overset{q.m.}{\to}}0$$

and $\sqrt{N}\widehat{\boldsymbol{\beta}}^{\mathrm{Pre}}_{\phi_1,\phi_2}-\sqrt{N}\widehat{\boldsymbol{\beta}}_{\phi_2}\underset{N\to\infty}{\overset{P}{\to}}0$. Then,

$$\sqrt{N}\left(\widehat{\boldsymbol{\beta}}^{\mathrm{Pre}}_{\phi_1,\phi_2}-\boldsymbol{\beta}\right)=\sqrt{N}\left(\widehat{\boldsymbol{\beta}}_{\phi_2}-\boldsymbol{\beta}\right)+o_P(1).$$

(b) Based on Pardo et al. [5]

$$\widehat{\boldsymbol{\beta}}_{\phi_2}=\boldsymbol{\beta}_0+\left(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}_N(\boldsymbol{\beta})\boldsymbol{X}\right)^{-1}\boldsymbol{X}^{\mathrm{T}}\mathrm{diag}\left((\boldsymbol{C}_i(\boldsymbol{\beta}_0))^{\mathrm{T}}_{i=1,\ldots,n}\right)\mathrm{diag}\left(\boldsymbol{p}\left(\boldsymbol{\beta}^0\right)^{-1/2}\right)\left(\widehat{\boldsymbol{p}}-\boldsymbol{p}\left(\boldsymbol{\beta}^0\right)\right)+o_P(N^{-1/2}) \tag{9}$$

and based on Menéndez et al. [3],

$$\widehat{\boldsymbol{\beta}}^{H_0}_{\phi_2}=\boldsymbol{\beta}_0+\boldsymbol{H}_N(\boldsymbol{\beta}_0)\left(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}_N(\boldsymbol{\beta}_0)\boldsymbol{X}\right)^{-1}\boldsymbol{X}^{\mathrm{T}}\mathrm{diag}\left((\boldsymbol{C}_i(\boldsymbol{\beta}_0))^{\mathrm{T}}_{i=1,\ldots,n}\right)\mathrm{diag}\left(\boldsymbol{p}\left(\boldsymbol{\beta}^0\right)^{-1/2}\right)(\widehat{\boldsymbol{p}}-\boldsymbol{p}(\boldsymbol{\beta}_0))+o_P(N^{-1/2}).$$

Therefore,

$$\widehat{\boldsymbol{\beta}}^{H_0}_{\phi_2}-\boldsymbol{\beta}=\widehat{\boldsymbol{\beta}}_{\phi_2}-\boldsymbol{\beta}-\left(\mathbf{X}^{\mathrm{T}}\mathbf{W}_N(\boldsymbol{\beta})\mathbf{X}\right)^{-1}\boldsymbol{K}(\boldsymbol{K}^{\mathrm{T}}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}_N(\boldsymbol{\beta})\boldsymbol{X})^{-1}\boldsymbol{K})^{-1}\left(\boldsymbol{K}^{\mathrm{T}}\widehat{\boldsymbol{\beta}}_{\phi_2}-\boldsymbol{m}\right)+o_P(N^{-1/2}).$$

Now taking into account that $\boldsymbol{m}=\boldsymbol{K}^{\mathrm{T}}\boldsymbol{\beta}-\boldsymbol{s}$ we have

$$\begin{aligned} \sqrt{N}\left(\widehat{\boldsymbol{\beta}}^{H_0}_{\phi_2}-\boldsymbol{\beta}\right) &= \sqrt{N}\left(\widehat{\boldsymbol{\beta}}_{\phi_2}-\boldsymbol{\beta}\right)-\left(\mathbf{X}^{\mathrm{T}}\mathbf{W}_N(\boldsymbol{\beta})\mathbf{X}\right)^{-1}\boldsymbol{K}(\boldsymbol{K}^{\mathrm{T}}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}_N(\boldsymbol{\beta})\boldsymbol{X})^{-1}\boldsymbol{K})^{-1}\sqrt{N}\left(\boldsymbol{K}^{\mathrm{T}}\widehat{\boldsymbol{\beta}}_{\phi_2}-\boldsymbol{K}^{\mathrm{T}}\boldsymbol{\beta}+\boldsymbol{s}\right) \\ &= \sqrt{N}\left(\widehat{\boldsymbol{\beta}}_{\phi_2}-\boldsymbol{\beta}\right)-\left(\mathbf{X}^{\mathrm{T}}\mathbf{W}_N(\boldsymbol{\beta})\mathbf{X}\right)^{-1}\boldsymbol{K}(\boldsymbol{K}^{\mathrm{T}}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}_N(\boldsymbol{\beta})\boldsymbol{X})^{-1}\boldsymbol{K})^{-1}\sqrt{N}\boldsymbol{K}^{\mathrm{T}}\left(\widehat{\boldsymbol{\beta}}_{\phi_2}-\boldsymbol{\beta}\right)+\sqrt{N}\mathbf{s}, \end{aligned}$$

and the asymptotic distribution of $\sqrt{N}\left(\widehat{\boldsymbol{\beta}}^{H_0}_{\phi_2}-\boldsymbol{\beta}\right)$ is degenerated under the fixed alternative hypotheses $H_1:\boldsymbol{K}^{\mathrm{T}}\boldsymbol{\beta}=\boldsymbol{m}+\boldsymbol{s}$. ■

The result in the previous theorem is important because it reveals that in order to obtain meaningful asymptotic coverage probabilities of the confidence set $C_{\beta}\left(\widehat{\boldsymbol{\beta}}^*_{\phi}\right)$ we must consider contiguous alternative hypotheses to $H_0$, i.e., we shall consider hypotheses of the type,

$$H_{1,N}:\boldsymbol{\beta}_N=\boldsymbol{\beta}_0+N^{-1/2}\boldsymbol{\Delta},$$

with $\boldsymbol{\beta}_0\in\Theta_0$ and $\boldsymbol{\Delta}\in\mathbb{R}^{k+1}$.

If we consider the function $g(\boldsymbol{\beta}) = \boldsymbol{K}^{\mathrm{T}}\boldsymbol{\beta} - \mathbf{m}$ it is clear that $\Theta_0 = \{\boldsymbol{\beta} \in \Theta : g(\boldsymbol{\beta}) = \mathbf{0}\}$ and the hypothesis $H_{1,N}$ is equivalent to the hypothesis

$$H_{1,N}^{*} : g(\boldsymbol{\beta}_N) = \boldsymbol{N}^{-1/2}\delta (H_{1,N}^{*} : \boldsymbol{K}^{\mathrm{T}}\boldsymbol{\beta}_N = \boldsymbol{m} + \boldsymbol{N}^{-1/2}\delta).$$

A Taylor expansion of $g(\boldsymbol{\beta}_N)$ around $\boldsymbol{\beta}_0 \in \Theta_0$ yields

$$g(\boldsymbol{\beta}_N) = g(\boldsymbol{\beta}_0) + \boldsymbol{K}^{\mathrm{T}}(\boldsymbol{\beta}_N - \boldsymbol{\beta}_0) + o\left(1\right),$$

but $g(\boldsymbol{\beta}_0) = 0$ and $\boldsymbol{\beta}_N - \boldsymbol{\beta}_0 = N^{-1/2}\boldsymbol{\Delta}$, hence

$$g(\boldsymbol{\beta}_N) = N^{-1/2}\boldsymbol{K}^{\mathrm{T}}\boldsymbol{\Delta} + o\left(1\right).$$

Now if we consider $\boldsymbol{\delta} = \boldsymbol{K}^{\mathrm{T}}\boldsymbol{\Delta}$ we have the equivalence in the limit.

On the other hand, we know that

$$N\left\|\boldsymbol{\beta}_N - \widehat{\boldsymbol{\beta}}_{\phi}^{*}\right\|^2_{\mathbf{X}^{\mathrm{T}}\mathbf{W}_N(\widehat{\boldsymbol{\beta}}_{\phi_2})\mathbf{X}} - N\left\|\boldsymbol{\beta}_N - \widehat{\boldsymbol{\beta}}_{\phi}^{*}\right\|^2_{\mathbf{X}^{\mathrm{T}}\mathbf{W}_N(\boldsymbol{\beta}_0)\mathbf{X}} \xrightarrow{P} 0.$$

Therefore in order to study the asymptotic behavior of $C_{\boldsymbol{\beta}}\left(\widehat{\boldsymbol{\beta}}_{\phi}^{*}\right)$ we shall consider that our recentered confidence sets are given by

$$C_{\boldsymbol{\beta}_N}\left(\widehat{\boldsymbol{\beta}}_{\phi}^{*}\right) = \left\{\boldsymbol{\beta}_N : N\left\|\boldsymbol{\beta}_N - \widehat{\boldsymbol{\beta}}_{\phi}^{*}\right\|^2_{\mathbf{X}^{\mathrm{T}}\mathbf{W}_N(\boldsymbol{\beta}_0)\mathbf{X}} \le \chi^2_{k+1,\alpha}\right\}.$$

We need an auxiliary lemma to obtain the asymptotic coverage probabilities of $C_{\boldsymbol{\beta}_N}\left(\widehat{\boldsymbol{\beta}}_{\phi_2}^{H_0}\right)$ and $C_{\boldsymbol{\beta}_N}\left(\widehat{\boldsymbol{\beta}}_{\phi_1,\phi_2}^{\mathrm{Pre}}\right)$.

**Lemma 2.** *We denote by* $\boldsymbol{\Gamma} = \left(\boldsymbol{\Gamma}_1^{\mathrm{T}}, \boldsymbol{\Gamma}_2{}^{\mathrm{T}}\right)^{\mathrm{T}}$, *(*$\boldsymbol{\Gamma}_1$ *is an* $r \times (k+1)$ *matrix and* $\boldsymbol{\Gamma}_2$ *a* $(k+1-r) \times (k+1)$ *matrix), the orthogonal matrix that diagonalizes the idempotent matrix*

$$(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}_N\left(\boldsymbol{\beta}_0\right)\boldsymbol{X})^{-1/2}\boldsymbol{K}^{\mathrm{T}}(\boldsymbol{K}^{\mathrm{T}}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}_N\left(\boldsymbol{\beta}_0\right)\boldsymbol{X})^{-1}\boldsymbol{K})^{-1}\boldsymbol{K}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}_N\left(\boldsymbol{\beta}_0\right)\boldsymbol{X})^{-1/2},$$

$\boldsymbol{\eta}_N^{\mathrm{T}} = (\boldsymbol{\eta}_1^{\mathrm{T}}, \boldsymbol{\eta}_2^{\mathrm{T}})$ *(*$\boldsymbol{\eta}_1$ *is an* $r \times 1$ *random vector and* $\boldsymbol{\eta}_2$ *a* $(k+1-r) \times 1$ *random vector) the random vector defined as*

$$\boldsymbol{\eta}_N = \sqrt{N}\left(\boldsymbol{\Gamma}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}_N\left(\boldsymbol{\beta}_0\right)\boldsymbol{X})^{1/2}\widehat{\boldsymbol{\beta}}_{\phi_2} - \boldsymbol{\Gamma}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}_N\left(\boldsymbol{\beta}_0\right)\boldsymbol{X})^{-1/2}\boldsymbol{K}(\boldsymbol{K}^{\mathrm{T}}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}_N\left(\boldsymbol{\beta}_0\right)\boldsymbol{X})^{-1}\boldsymbol{K})^{-1}\boldsymbol{m}\right) + o_P\left(1\right). \tag{10}$$

*Then, we have:*

(a) $\boldsymbol{\eta}_N - E\left[\boldsymbol{\eta}_N\right] \xrightarrow[N\to\infty]{L} \mathcal{N}(\mathbf{0}, \boldsymbol{I}^{*})$ *where* $\mathbf{I}^{*} = \begin{pmatrix} \mathbf{I}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{k+1-r} \end{pmatrix}$.

(b) $N\left\|\boldsymbol{\beta}_N - \widehat{\boldsymbol{\beta}}_{\phi_1,\phi_2}^{\mathrm{Pre}}\right\|^2_{\mathbf{X}^{\mathrm{T}}\mathbf{W}_N(\boldsymbol{\beta}_0)\mathbf{X}} = \left\|E\left[\boldsymbol{\eta}_1\right] - \boldsymbol{\eta}_1 I_{\left[\chi^2_{r,\alpha},\infty\right)}\left(\boldsymbol{\eta}_1^{\mathrm{T}}\boldsymbol{\eta}_1 + o_P(1)\right)\right\|^2 + \left\|E\left[\boldsymbol{\eta}_2\right] - \boldsymbol{\eta}_2\right\|^2 + o_P\left(1\right).$

**Proof.** Part (a). Based on the definition of $\boldsymbol{\eta}_N$, given in (10), we have

$$\boldsymbol{\eta}_N - E\left[\boldsymbol{\eta}_N\right] = \boldsymbol{\Gamma}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}_N\left(\boldsymbol{\beta}_0\right)\boldsymbol{X})^{1/2}\sqrt{N}\left(\widehat{\boldsymbol{\beta}}_{\phi_2} - \boldsymbol{\beta}_N\right).$$

Now by (6) we obtain

$$\boldsymbol{\eta}_N - E\left[\boldsymbol{\eta}_N\right] \xrightarrow[N\to\infty]{L} \mathcal{N}(\mathbf{0}, \boldsymbol{I}^{*}).$$

Now we consider part (b). The matrix $\boldsymbol{\Gamma}$ verifies

$$\boldsymbol{\Gamma}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}_N\left(\boldsymbol{\beta}_0\right)\boldsymbol{X})^{-1/2}\boldsymbol{K}^{\mathrm{T}}(\boldsymbol{K}^{\mathrm{T}}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}_N\left(\boldsymbol{\beta}_0\right)\boldsymbol{X})^{-1}\boldsymbol{K})^{-1}\boldsymbol{K}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}_N\left(\boldsymbol{\beta}_0\right)\boldsymbol{X})^{-1/2}\boldsymbol{\Gamma}^{\mathrm{T}} = \begin{pmatrix} \mathbf{I}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}.$$

In [3] it was obtained,

$$T_N^{\phi_1,\phi_2} = \sqrt{N}\left(\boldsymbol{K}^{\mathrm{T}}\widehat{\boldsymbol{\beta}}_{\phi_2} - \boldsymbol{m}\right)^{\mathrm{T}}(\boldsymbol{K}^{\mathrm{T}}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}_N\left(\boldsymbol{\beta}_0\right)\boldsymbol{X})^{-1}\boldsymbol{K})^{-1}\sqrt{N}\left(\boldsymbol{K}^{\mathrm{T}}\widehat{\boldsymbol{\beta}}_{\phi_2} - \boldsymbol{m}\right) + o_P(1).$$

Now we have,

$$\begin{aligned}
T_N^{\phi_1,\phi_2} &= \sqrt{N}\left(\boldsymbol{K}^{\mathrm{T}}\widehat{\boldsymbol{\beta}}_{\phi_2} - \boldsymbol{m}\right)^{\mathrm{T}}(\boldsymbol{K}^{\mathrm{T}}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}_N\left(\boldsymbol{\beta}_0\right)\boldsymbol{X})^{-1}\boldsymbol{K})^{-1}\boldsymbol{K}^{\mathrm{T}}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}_N\left(\boldsymbol{\beta}_0\right)\boldsymbol{X})^{-1/2} \\
&\quad \times (\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}_N(\boldsymbol{\beta}_0)\boldsymbol{X})^{-1/2}\boldsymbol{K}(\boldsymbol{K}^{\mathrm{T}}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}_N\left(\boldsymbol{\beta}_0\right)\boldsymbol{X})^{-1}\boldsymbol{K})^{-1}\sqrt{N}\left(\boldsymbol{K}^{\mathrm{T}}\widehat{\boldsymbol{\beta}}_{\phi_2} - \boldsymbol{m}\right) + o_P(1) \\
&= \sqrt{N}\left(\boldsymbol{\Gamma}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}_N\left(\boldsymbol{\beta}_0\right)\boldsymbol{X})^{1/2}\widehat{\boldsymbol{\beta}}_{\phi_2} - \boldsymbol{\Gamma}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}_N\left(\boldsymbol{\beta}_0\right)\boldsymbol{X})^{-1/2}\boldsymbol{K}(\boldsymbol{K}^{\mathrm{T}}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}_N\left(\boldsymbol{\beta}_0\right)\boldsymbol{X})^{-1}\boldsymbol{K})^{-1}\boldsymbol{m}\right)^{\mathrm{T}} \\
&\quad \times \boldsymbol{\Gamma}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}_N\left(\boldsymbol{\beta}_0\right)\boldsymbol{X})^{-1/2}\boldsymbol{K}(\boldsymbol{K}^{\mathrm{T}}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}_N\left(\boldsymbol{\beta}_0\right)\boldsymbol{X})^{-1}\boldsymbol{K})^{-1}\boldsymbol{K}^{\mathrm{T}}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}_N\left(\boldsymbol{\beta}_0\right)\boldsymbol{X})^{-1/2}\boldsymbol{\Gamma}^{\mathrm{T}} \\
&\quad \times \boldsymbol{\Gamma}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}_N\left(\boldsymbol{\beta}_0\right)\boldsymbol{X})^{-1/2}\boldsymbol{K}(\boldsymbol{K}^{\mathrm{T}}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}_N\left(\boldsymbol{\beta}_0\right)\boldsymbol{X})^{-1}\boldsymbol{K})^{-1}\boldsymbol{K}^{\mathrm{T}}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}_N\left(\boldsymbol{\beta}_0\right)\boldsymbol{X})^{-1/2}\boldsymbol{\Gamma}^{\mathrm{T}}\sqrt{N}\boldsymbol{\Gamma}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}_N\left(\boldsymbol{\beta}_0\right)\boldsymbol{X})^{1/2}\widehat{\boldsymbol{\beta}}_{\phi_2}\boldsymbol{K}^{-1} \\
&\quad - \sqrt{N}\boldsymbol{\Gamma}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}_N\left(\boldsymbol{\beta}_0\right)\boldsymbol{X})^{-1/2}\boldsymbol{K}\boldsymbol{K}^{\mathrm{T}}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}_N\left(\boldsymbol{\beta}_0\right)\boldsymbol{X})^{-1}\boldsymbol{m} + o_P(1).
\end{aligned}$$

Therefore,

$$T_N^{\phi_1,\phi_2} = \boldsymbol{\eta}_N^{\mathrm{T}} \begin{pmatrix} \mathbf{I}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{I}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \boldsymbol{\eta}_N + o_P(1) = \boldsymbol{\eta}_1^{\mathrm{T}} \boldsymbol{\eta}_1 + o_P(1)$$

and the asymptotic distribution of $T_N^{\phi_1,\phi_2}$ is a noncentral chi-square with $r$ degrees of freedom and noncentrality parameter

$$\lambda = E\left[\boldsymbol{\eta}_1^{\mathrm{T}} \boldsymbol{\eta}_1\right] = \boldsymbol{\Delta}^{\mathrm{T}} \boldsymbol{K}^{\mathrm{T}} (\boldsymbol{K}^{\mathrm{T}} (\boldsymbol{X}^{\mathrm{T}} \boldsymbol{W}\,(\boldsymbol{\beta}_0)\,\boldsymbol{X})^{-1} \boldsymbol{K})^{-1} \boldsymbol{K} \boldsymbol{\Delta}. \tag{11}$$

This result follows since $\sqrt{N}(\boldsymbol{\eta}_1 - E[\boldsymbol{\eta}_1])$ converges in law to an $r$-normal random vector with mean vector zero and variance covariance matrix $\mathbf{I}_r$.

Using the definition of $\widehat{\boldsymbol{\beta}}^{\mathrm{Pre}}_{\phi_1,\phi_2}$ given in (8), we obtain

$$\begin{aligned}
N \left\| \boldsymbol{\beta}_N - \widehat{\boldsymbol{\beta}}^{\mathrm{Pre}}_{\phi_1,\phi_2} \right\|^2_{\mathbf{X}^{\mathrm{T}}\mathbf{W}_N(\boldsymbol{\beta}_0)\mathbf{X}}
&= \sqrt{N} \left( (\boldsymbol{X}^{\mathrm{T}} \boldsymbol{W}_N\,(\boldsymbol{\beta}_0)\,\boldsymbol{X})^{1/2} \left( \boldsymbol{\beta}_N - \widehat{\boldsymbol{\beta}}^{H_0}_{\phi_2} \right) - (\boldsymbol{X}^{\mathrm{T}} \boldsymbol{W}_N\,(\boldsymbol{\beta}_0)\,\boldsymbol{X})^{1/2} \left( \widehat{\boldsymbol{\beta}}_{\phi_2} - \widehat{\boldsymbol{\beta}}^{H_0}_{\phi_2} \right) I_{[\chi^2_{r,\alpha},\infty)} \left( \boldsymbol{\eta}_1^{\mathrm{T}} \boldsymbol{\eta}_1 + o_P(1) \right) \right)^{\mathrm{T}} \\
&\quad \times \sqrt{N} \left( (\boldsymbol{X}^{\mathrm{T}} \boldsymbol{W}_N\,(\boldsymbol{\beta}_0)\,\boldsymbol{X})^{1/2} \left( \boldsymbol{\beta}_N - \widehat{\boldsymbol{\beta}}^{H_0}_{\phi_2} \right) - (\boldsymbol{X}^{\mathrm{T}} \boldsymbol{W}_N\,(\boldsymbol{\beta}_0)\,\boldsymbol{X})^{1/2} \left( \widehat{\boldsymbol{\beta}}_{\phi_2} - \widehat{\boldsymbol{\beta}}^{H_0}_{\phi_2} \right) I_{[\chi^2_{r,\alpha},\infty)} \left( \boldsymbol{\eta}_1^{\mathrm{T}} \boldsymbol{\eta}_1 + o_P(1) \right) \right) \\
&= \sqrt{N} \left( \boldsymbol{\Gamma} (\boldsymbol{X}^{\mathrm{T}} \boldsymbol{W}_N\,(\boldsymbol{\beta}_0)\,\boldsymbol{X})^{1/2} \left( \boldsymbol{\beta}_N - \widehat{\boldsymbol{\beta}}^{H_0}_{\phi_2} \right) - \boldsymbol{\Gamma} (\boldsymbol{X}^{\mathrm{T}} \boldsymbol{W}_N\,(\boldsymbol{\beta}_0)\,\boldsymbol{X})^{1/2} \left( \widehat{\boldsymbol{\beta}}_{\phi_2} - \widehat{\boldsymbol{\beta}}^{H_0}_{\phi_2} \right) I_{[\chi^2_{r,\alpha},\infty)} \left( \boldsymbol{\eta}_1^{\mathrm{T}} \boldsymbol{\eta}_1 + o_P(1) \right) \right)^{\mathrm{T}} \\
&\quad \times \sqrt{N} \left( \boldsymbol{\Gamma} (\boldsymbol{X}^{\mathrm{T}} \boldsymbol{W}_N\,(\boldsymbol{\beta}_0)\,\boldsymbol{X})^{1/2} \left( \boldsymbol{\beta}_N - \widehat{\boldsymbol{\beta}}^{H_0}_{\phi_2} \right) - \boldsymbol{\Gamma} (\boldsymbol{X}^{\mathrm{T}} \boldsymbol{W}_N\,(\boldsymbol{\beta}_0)\,\boldsymbol{X})^{1/2} \left( \widehat{\boldsymbol{\beta}}_{\phi_2} - \widehat{\boldsymbol{\beta}}^{H_0}_{\phi_2} \right) I_{[\chi^2_{r,\alpha},\infty)} \left( \boldsymbol{\eta}_1^{\mathrm{T}} \boldsymbol{\eta}_1 + o_P(1) \right) \right) \\
&= \boldsymbol{\xi}_N^{\mathrm{T}} \boldsymbol{\xi}_N
\end{aligned}$$

where

$$\boldsymbol{\xi}_N = \sqrt{N} \boldsymbol{\Gamma} (\boldsymbol{X}^{\mathrm{T}} \boldsymbol{W}_N\,(\boldsymbol{\beta}_0)\,\boldsymbol{X})^{1/2} \left( \boldsymbol{\beta}_N - \widehat{\boldsymbol{\beta}}^{H_0}_{\phi_2} \right) - \boldsymbol{\Gamma} (\boldsymbol{X}^{\mathrm{T}} \boldsymbol{W}_N\,(\boldsymbol{\beta}_0)\,\boldsymbol{X})^{1/2} \left( \widehat{\boldsymbol{\beta}}_{\phi_2} - \widehat{\boldsymbol{\beta}}^{H_0}_{\phi_2} \right) I_{[\chi^2_{r,\alpha},\infty)} \left( \boldsymbol{\eta}_1^{\mathrm{T}} \boldsymbol{\eta}_1 + o_P(1) \right). \tag{12}$$

Denoting

$$A = \sqrt{N} \boldsymbol{\Gamma} (\boldsymbol{X}^{\mathrm{T}} \boldsymbol{W}_N\,(\boldsymbol{\beta}_0)\,\boldsymbol{X})^{1/2} \left( \widehat{\boldsymbol{\beta}}_{\phi_2} - \widehat{\boldsymbol{\beta}}^{H_0}_{\phi_2} \right),$$

we can write

$$\begin{aligned}
A &= \sqrt{N} \boldsymbol{\Gamma} (\boldsymbol{X}^{\mathrm{T}} \boldsymbol{W}_N\,(\boldsymbol{\beta}_0)\,\boldsymbol{X})^{1/2} (\boldsymbol{X}^{\mathrm{T}} \boldsymbol{W}_N\,(\boldsymbol{\beta}_0)\,\boldsymbol{X})^{-1} \boldsymbol{K} (\boldsymbol{K}^{\mathrm{T}} (\boldsymbol{X}^{\mathrm{T}} \boldsymbol{W}_N\,(\boldsymbol{\beta}_0)\,\boldsymbol{X})^{-1} \boldsymbol{K})^{-1} (\boldsymbol{K}^{\mathrm{T}} \widehat{\boldsymbol{\beta}}_{\phi_2} - \boldsymbol{m}) + o_P(1) \\
&= \sqrt{N} \left( \boldsymbol{\Gamma} (\boldsymbol{X}^{\mathrm{T}} \boldsymbol{W}_N\,(\boldsymbol{\beta}_0)\,\boldsymbol{X})^{-1/2} \boldsymbol{K} (\boldsymbol{K}^{\mathrm{T}} (\boldsymbol{X}^{\mathrm{T}} \boldsymbol{W}_N\,(\boldsymbol{\beta}_0)\,\boldsymbol{X})^{-1} \boldsymbol{K})^{-1} \boldsymbol{K}^{\mathrm{T}} (\boldsymbol{X}^{\mathrm{T}} \boldsymbol{W}_N\,(\boldsymbol{\beta}_0)\,\boldsymbol{X})^{-1/2} \boldsymbol{\Gamma}^{\mathrm{T}} \right) \\
&\quad \times \left( \boldsymbol{\Gamma} (\boldsymbol{X}^{\mathrm{T}} \boldsymbol{W}_N\,(\boldsymbol{\beta}_0)\,\boldsymbol{X})^{1/2} \widehat{\boldsymbol{\beta}}_{\phi_2} - \boldsymbol{\Gamma} (\boldsymbol{X}^{\mathrm{T}} \boldsymbol{W}_N\,(\boldsymbol{\beta}_0)\,\boldsymbol{X})^{-1/2} \boldsymbol{K} (\boldsymbol{K}^{\mathrm{T}} (\boldsymbol{X}^{\mathrm{T}} \boldsymbol{W}_N\,(\boldsymbol{\beta}_0)\,\boldsymbol{X})^{-1} \boldsymbol{K})^{-1} \boldsymbol{m} \right) + o_P(1) \\
&= \begin{pmatrix} \mathbf{I}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \boldsymbol{\eta}_1 \\ \boldsymbol{\eta}_2 \end{pmatrix} + o_P(1) = \begin{pmatrix} \boldsymbol{\eta}_1 \\ \mathbf{0} \end{pmatrix} + o_P(1)
\end{aligned}$$

and denoting $B = \sqrt{N} \boldsymbol{\Gamma} (\boldsymbol{X}^{\mathrm{T}} \boldsymbol{W}_N\,(\boldsymbol{\beta}_0)\,\boldsymbol{X})^{1/2} \left( \boldsymbol{\beta}_N - \widehat{\boldsymbol{\beta}}^{H_0}_{\phi_2} \right)$,

$$\begin{aligned}
B &= \sqrt{N} \boldsymbol{\Gamma} (\boldsymbol{X}^{\mathrm{T}} \boldsymbol{W}_N\,(\boldsymbol{\beta}_0)\,\boldsymbol{X})^{1/2} \left( \boldsymbol{\beta}_N - \widehat{\boldsymbol{\beta}}_{\phi_2} + (\boldsymbol{X}^{\mathrm{T}} \boldsymbol{W}_N\,(\boldsymbol{\beta}_0)\,\boldsymbol{X})^{-1} \boldsymbol{K} (\boldsymbol{K}^{\mathrm{T}} (\boldsymbol{X}^{\mathrm{T}} \boldsymbol{W}_N\,(\boldsymbol{\beta}_0)\,\boldsymbol{X})^{-1} \boldsymbol{K})^{-1} \left( \boldsymbol{K}^{\mathrm{T}} \widehat{\boldsymbol{\beta}}_{\phi_2} - \boldsymbol{m} \right) \right) + o_P(1) \\
&= \sqrt{N} \boldsymbol{\Gamma} (\boldsymbol{X}^{\mathrm{T}} \boldsymbol{W}_N\,(\boldsymbol{\beta}_0)\,\boldsymbol{X})^{1/2} \left( \boldsymbol{\beta}_N - \widehat{\boldsymbol{\beta}}_{\phi_2} \right) + \sqrt{N} \boldsymbol{\Gamma} (\boldsymbol{X}^{\mathrm{T}} \boldsymbol{W}_N\,(\boldsymbol{\beta}_0)\,\boldsymbol{X})^{-1/2} \boldsymbol{K} (\boldsymbol{K}^{\mathrm{T}} (\boldsymbol{X}^{\mathrm{T}} \boldsymbol{W}_N\,(\boldsymbol{\beta}_0)\,\boldsymbol{X})^{-1} \boldsymbol{K})^{-1} \left( \boldsymbol{K}^{\mathrm{T}} \widehat{\boldsymbol{\beta}}_{\phi_2} - \boldsymbol{m} \right) + o_P(1) \\
&= \sqrt{N} \boldsymbol{\Gamma} (\boldsymbol{X}^{\mathrm{T}} \boldsymbol{W}_N\,(\boldsymbol{\beta}_0)\,\boldsymbol{X})^{1/2} \left( \boldsymbol{\beta}_N - \widehat{\boldsymbol{\beta}}_{\phi_2} \right) + \begin{pmatrix} \boldsymbol{\eta}_1 \\ \mathbf{0} \end{pmatrix} + o_P(1).
\end{aligned}$$

But

$$\boldsymbol{\eta} - E[\boldsymbol{\eta}] = \left( \boldsymbol{\eta}_1^{\mathrm{T}}, \boldsymbol{\eta}_2^{\mathrm{T}} \right)^{\mathrm{T}} - \left( E[\boldsymbol{\eta}_1]^{\mathrm{T}}, E[\boldsymbol{\eta}_2]^{\mathrm{T}} \right)^{\mathrm{T}} = \boldsymbol{\Gamma} (\boldsymbol{X}^{\mathrm{T}} \boldsymbol{W}_N\,(\boldsymbol{\beta}_0)\,\boldsymbol{X})^{1/2} \sqrt{N} \left( \widehat{\boldsymbol{\beta}}_{\phi_2} - \boldsymbol{\beta}_N \right)$$

and

$$\sqrt{N} \boldsymbol{\Gamma} (\boldsymbol{X}^{\mathrm{T}} \boldsymbol{W}_N\,(\boldsymbol{\beta}_0)\,\boldsymbol{X})^{1/2} \left( \boldsymbol{\beta}_N - \widehat{\boldsymbol{\beta}}^{H_0}_{\phi_2} \right) = \left( E[\boldsymbol{\eta}_1]^{\mathrm{T}}, (E[\boldsymbol{\eta}_2] - \boldsymbol{\eta}_2)^{\mathrm{T}} \right)^{\mathrm{T}}.$$

Therefore, the random vector $\boldsymbol{\xi}_N$ defined in (12) can be written as

$$\boldsymbol{\xi}_N = \begin{pmatrix} E[\boldsymbol{\eta}_1] \\ E[\boldsymbol{\eta}_2] - \boldsymbol{\eta}_2 \end{pmatrix} - \begin{pmatrix} \boldsymbol{\eta}_1 \\ \mathbf{0} \end{pmatrix} I_{[\chi^2_{r,\alpha},\infty)} \left( \boldsymbol{\eta}_1^{\mathrm{T}} \boldsymbol{\eta}_1 + o_P(1) \right) + o_P(1)$$

and

$$N \left\| \boldsymbol{\beta}_N - \widehat{\boldsymbol{\beta}}^{\mathrm{Pre}}_{\phi_1,\phi_2} \right\|^2_{\mathbf{X}^{\mathrm{T}}\mathbf{W}_N(\boldsymbol{\beta}_0)\mathbf{X}} = \left\| E[\boldsymbol{\eta}_1] - \boldsymbol{\eta}_1 I_{[\chi^2_{r,\alpha},\infty)} \left( \boldsymbol{\eta}_1^{\mathrm{T}} \boldsymbol{\eta}_1 + o_P(1) \right) \right\|^2 + \left\| E[\boldsymbol{\eta}_2] - \boldsymbol{\eta}_2 \right\|^2 + o_P(1). \quad \blacksquare$$

In the following theorem we are going to obtain the coverage probabilities of the sets $C_{\beta_N}(\widehat{\boldsymbol{\beta}}_{\phi}^{*})$ with $\widehat{\boldsymbol{\beta}}_{\phi}^{*}$ equal to $\widehat{\boldsymbol{\beta}}_{\phi_2}^{H_0}$ or $\widehat{\boldsymbol{\beta}}_{\phi_1,\phi_2}^{\mathrm{Pre}}$.

**Theorem 3.** *We have, under $H_{1,N}$:*

(a) $\lim_{N\to\infty}\Pr\left(C_{\beta_N}(\widehat{\boldsymbol{\beta}}_{\phi_1,\phi_2}^{\mathrm{Pre}})\right) = G_r(\chi^2_{r,\alpha};\lambda)G_{k+1-r}(\chi^2_{k+1,\alpha}-\lambda;0) + \lim_{N\to\infty}\int_0^{\chi^2_{k+1,\alpha}}\Pr(\|E[\boldsymbol{\eta}_1]-\boldsymbol{\eta}_1\|^2 \le \chi^2_{k+1,\alpha}-t;\|\boldsymbol{\eta}_1\|^2 > \chi^2_{r,\alpha})\mathrm{d}G_{k+1-r}(t;0)$.

(b) $\lim_{N\to\infty}\Pr\left(C_{\beta_N}(\widehat{\boldsymbol{\beta}}_{\phi_2}^{H_0})\right) = G_{k+1-r}(\chi^2_{k+1,\alpha}-\lambda;0)$.

*By $G_a(b;\mu)$ we are denoting the distribution function of a noncentral chi-square random variable with noncentrality parameter $\mu$ and "a" degrees of freedom evaluated at "b".*

**Proof.** (a) We denote $l = \lim_{N\to\infty}\Pr\left(C_{\beta_N}(\widehat{\boldsymbol{\beta}}_{\phi_1,\phi_2}^{\mathrm{Pre}})\right)$. We have,

$$\begin{aligned}
l &= \lim_{N\to\infty}\Pr\left(\left\|E[\boldsymbol{\eta}_1]-\boldsymbol{\eta}_1 I_{[\chi^2_{r,\alpha},\infty)}\left(\boldsymbol{\eta}_1^{\mathrm{T}}\boldsymbol{\eta}_1+o_P(1)\right)\right\|^2 + \|E[\boldsymbol{\eta}_2]-\boldsymbol{\eta}_2\|^2 \le \chi^2_{k+1,\alpha}\right)\\
&= \lim_{N\to\infty}\Pr\left(\|E[\boldsymbol{\eta}_1]\|^2 + \|E[\boldsymbol{\eta}_2]-\boldsymbol{\eta}_2\|^2 < \chi^2_{k+1,\alpha};\|\boldsymbol{\eta}_1\|^2 \le \chi^2_{r,\alpha}\right)\\
&\quad + \lim_{N\to\infty}\Pr\left(\|E[\boldsymbol{\eta}_1]-\boldsymbol{\eta}_1\|^2 + \|E[\boldsymbol{\eta}_2]-\boldsymbol{\eta}_2\|^2 < \chi^2_{k+1,\alpha};\|\boldsymbol{\eta}_1\|^2 > \chi^2_{r,\alpha}\right)\\
&= G_r(\chi^2_{r,\alpha};\lambda)G_{k+1-r}(\chi^2_{k+1,\alpha}-\lambda;0) + \lim_{N\to\infty}\int_0^{\chi^2_{k+1,\alpha}}\Pr\left(\|E[\boldsymbol{\eta}_1]-\boldsymbol{\eta}_1\|^2 \le \chi^2_{k+1,\alpha}-t;\|\boldsymbol{\eta}_1\|^2 > \chi^2_{r,\alpha}\right)\mathrm{d}G_{k+1-r}(t;0).
\end{aligned}$$

(b) It is well known that $\widehat{\boldsymbol{\beta}}_{\phi_1,\phi_2}^{\mathrm{Pre}} = \widehat{\boldsymbol{\beta}}_{\phi_2}^{H_0}$ if $T_N^{\phi_1,\phi_2} < \chi^2_{r,\alpha}$, therefore based on the previous Lemma we have

$$N\left\|\boldsymbol{\beta}_N-\widehat{\boldsymbol{\beta}}_{\phi_2}^{H_0}\right\|^2_{\mathbf{X}^{\mathrm{T}}\mathbf{W}_N(\beta_0)\mathbf{X}} = \|E[\boldsymbol{\eta}_1]\|^2 + \|E[\boldsymbol{\eta}_2]-\boldsymbol{\eta}_2\|^2.$$

Therefore

$$\begin{aligned}
\lim_{N\to\infty}\Pr\left(C_{\beta_N}(\widehat{\boldsymbol{\beta}}_{\phi_2})\right) &= \lim_{N\to\infty}\Pr\left(\|E[\boldsymbol{\eta}_1]\|^2 + \|E[\boldsymbol{\eta}_2-\boldsymbol{\eta}_2]\|^2 \le \chi^2_{k+1,\alpha}\right)\\
&= \lim_{N\to\infty}\Pr\left(\|E[\boldsymbol{\eta}_2]-\boldsymbol{\eta}_2\|^2 \le \chi^2_{k+1,\alpha}-\lambda\right)\\
&= G_{k+1-r}(\chi^2_{k+1,\alpha}-\lambda;0). \quad \blacksquare
\end{aligned}$$

**Remark 4.** We know that under $H_{1,N}$

$$\lim_{N\to\infty}\Pr\left(N\left\|\boldsymbol{\beta}_N-\widehat{\boldsymbol{\beta}}_{\phi_2}\right\|^2_{\mathbf{X}^{\mathrm{T}}\mathbf{W}_N(\beta_0)\mathbf{X}} \le \chi^2_{k+1,\alpha}\right) = 1-\alpha$$

and this probability does not depend on $\lambda$, i.e.,

$$\lim_{N\to\infty}\Pr\left(C_{\beta_N}(\widehat{\boldsymbol{\beta}}_{\phi_2})\right) = 1-\alpha.$$

If we consider

$$C_{\beta_N}(\widehat{\boldsymbol{\beta}}_{\phi_2}^{H_0}) = \left\{\boldsymbol{\beta}_N : N\left\|\boldsymbol{\beta}_N-\widehat{\boldsymbol{\beta}}_{\phi_2}^{H_0}\right\|^2_{\mathbf{X}^{\mathrm{T}}\mathbf{W}_N(\beta_0)\mathbf{X}} \le \chi^2_{k+1,\alpha}\right\},$$

we have by (b) in Theorem 3 that

$$\lim_{N\to\infty}\Pr\left(C_{\beta_N}(\widehat{\boldsymbol{\beta}}_{\phi_2}^{H_0})\right) = G_{k+1-r}(\chi^2_{k+1,\alpha}-\lambda;0).$$

We can observe that $G_{k+1-r}(\chi^2_{k+1,\alpha}-\lambda;0)$ is a decreasing function on $\lambda$. At $\lambda = 0$, it attains the maximum value $G_{k+1-r}(\chi^2_{k+1,\alpha};0)$ and it tends to zero as $\lambda \to \chi^2_{k+1,\alpha}$. The coverage probabilities of $C_{\beta}(\widehat{\boldsymbol{\beta}}_{\phi_2})$ and $C_{\beta}(\widehat{\boldsymbol{\beta}}_{\phi_2}^{H_0})$ are equal if $\lambda = \chi^2_{k+1,\alpha} - G^{-1}_{k+1-r}(1-\alpha;0)$.

The asymptotic coverage probability of $C_{\beta_N}(\widehat{\boldsymbol{\beta}}_{\phi_1,\phi_2}^{\mathrm{Pre}})$ depends on the noncentrality parameter $\lambda$ in the following way:

**Theorem 5.** *The following results hold:*

(i) *If $0 \le \lambda < \chi^2_{k+1,\alpha}$, then*

$$\lim_{N\to\infty}\Pr\left(C_{\beta_N}(\widehat{\boldsymbol{\beta}}_{\phi_1,\phi_2}^{\mathrm{Pre}})\right) \ge 1-\alpha.$$

(ii) *If* $\chi^2_{k+1,\alpha} \leq \lambda \leq \left((\chi^2_{k+1,\alpha})^{1/2} + (\chi^2_{r,\alpha})^{1/2}\right)^2$, *then*

$$\lim_{N\to\infty} \Pr\left(C_{\boldsymbol{\beta}_N}(\widehat{\boldsymbol{\beta}}^{\mathrm{Pre}}_{\phi_1,\phi_2})\right) \leq 1-\alpha.$$

(iii) *If* $\lambda > \left((\chi^2_{k+1,\alpha})^{1/2} + (\chi^2_{r,\alpha})^{1/2}\right)^2$, *then*

$$\lim_{N\to\infty} \Pr\left(C_{\boldsymbol{\beta}_N}(\widehat{\boldsymbol{\beta}}^{\mathrm{Pre}}_{\phi_1,\phi_2})\right) = 1-\alpha.$$

**Proof.** (i) We assume $\lambda < \chi^2_{k+1,\alpha}$. We denote

$$l = \lim_{N\to\infty} \Pr\left(C_{\boldsymbol{\beta}_N}(\widehat{\boldsymbol{\beta}}^{\mathrm{Pre}}_{\phi_1,\phi_2})\right), \tag{13}$$

we have

$$\begin{aligned}
l &= \lim_{N\to\infty} \Pr\left\{\left\|E[\boldsymbol{\eta}_2]-\boldsymbol{\eta}_2\right\|^2 + \lambda \leq \chi^2_{k+1,\alpha};\ \left\|\boldsymbol{\eta}_1\right\|^2 \leq \chi^2_{r,\alpha}\right\} \\
&\quad + \lim_{N\to\infty} \int_0^{\chi^2_{k+1,\alpha}} \Pr\left(\left\|E[\boldsymbol{\eta}_1]-\boldsymbol{\eta}_1\right\|^2 \leq \chi^2_{k+1,\alpha}-t;\ \left\|\boldsymbol{\eta}_1\right\|^2 > \chi^2_{r,\alpha}\right) \mathrm{d}G_{k+1-r}(t;0) \\
&= \lim_{N\to\infty} \int_0^{\chi^2_{k+1,\alpha}} \Pr\left\{\left\|E[\boldsymbol{\eta}_1]\right\|^2 \leq \chi^2_{k+1,\alpha}-t;\ \left\|\boldsymbol{\eta}_1\right\|^2 \leq \chi^2_{r,\alpha}\right\} \mathrm{d}G_{k+1-r}(t;0) \\
&\quad + \lim_{N\to\infty} \int_0^{\chi^2_{k+1,\alpha}} \Pr\left(\left\|E[\boldsymbol{\eta}_1]-\boldsymbol{\eta}_1\right\|^2 \leq \chi^2_{k+1,\alpha}-t;\ \left\|\boldsymbol{\eta}_1\right\|^2 > \chi^2_{r,\alpha}\right) \mathrm{d}G_{k+1-r}(t;0) \\
&\geq \lim_{N\to\infty} \int_0^{\chi^2_{k+1,\alpha}} \Pr\left(\left\|E[\boldsymbol{\eta}_1]-\boldsymbol{\eta}_1\right\|^2 \leq \chi^2_{k+1,\alpha}-t;\ \left\|\boldsymbol{\eta}_1\right\|^2 \leq \chi^2_{r,\alpha}\right) \mathrm{d}G_{k+1-r}(t;0) \\
&\quad + \lim_{N\to\infty} \int_0^{\chi^2_{k+1,\alpha}} \Pr\left(\left\|E[\boldsymbol{\eta}_1]-\boldsymbol{\eta}_1\right\|^2 \leq \chi^2_{k+1,\alpha}-t;\ \left\|\boldsymbol{\eta}_1\right\|^2 > \chi^2_{r,\alpha}\right) \mathrm{d}G_{k+1-r}(t;0) \\
&= \lim_{N\to\infty} \Pr\left\{\left\|E[\boldsymbol{\eta}_1]-\boldsymbol{\eta}_1\right\|^2 + \left\|E[\boldsymbol{\eta}_2]-\boldsymbol{\eta}_2\right\|^2 \leq \chi^2_{k+1,\alpha}\right\} \\
&= \lim_{N\to\infty} \Pr\left(\boldsymbol{\eta}^{\mathrm{T}}\boldsymbol{\eta} \leq \chi^2_{k+1,\alpha}\right) = \Pr\left(\chi^2_{k+1} \leq \chi^2_{k+1,\alpha}\right) = 1-\alpha.
\end{aligned}$$

(ii) We assume $\chi^2_{k+1,\alpha} \leq \lambda \leq \left((\chi^2_{k+1,\alpha})^{1/2} + (\chi^2_{r,\alpha})^{1/2}\right)^2$. On the other hand we have established before that

$$l = G_r(\chi^2_{r,\alpha};\lambda)G_{k+1-r}(\chi^2_{k+1,\alpha}-\lambda;0) + \lim_{N\to\infty} \int_0^{\chi^2_{k+1,\alpha}} \Pr\left(\left\|E[\boldsymbol{\eta}_1]-\boldsymbol{\eta}_1\right\|^2 \leq \chi^2_{k+1,\alpha}-t;\ \left\|\boldsymbol{\eta}_1\right\|^2 > \chi^2_{r,\alpha}\right) \mathrm{d}G_{k+1-r}(t;0),$$

where $l$ was defined in (13). But if $\lambda \geq \chi^2_{k+1,\alpha}$ then $\lambda \geq \chi^2_{k+1,\alpha} \geq \chi^2_{r,\alpha}$, hence $\chi^2_{r,\alpha}-\lambda \leq 0$ and $G_{k+1-r}(\chi^2_{k+1,\alpha}-\lambda;0) = 0$. Therefore

$$\begin{aligned}
l &= \lim_{N\to\infty} \int_0^{\chi^2_{k+1,\alpha}} \Pr\left(\left\|E[\boldsymbol{\eta}_1]-\boldsymbol{\eta}_1\right\|^2 \leq \chi^2_{k+1,\alpha}-t;\ \left\|\boldsymbol{\eta}_1\right\|^2 > \chi^2_{r,\alpha}\right) \mathrm{d}G_{k+1-r}(t;0) \\
&\leq \lim_{N\to\infty} \int_0^{\chi^2_{k+1,\alpha}} \Pr\left(\left\|E[\boldsymbol{\eta}_1]-\boldsymbol{\eta}_1\right\|^2 \leq \chi^2_{k+1,\alpha}-t\right) \mathrm{d}G_{k+1-r}(t;0) \\
&= \lim_{N\to\infty} \Pr\left(\left\|E[\boldsymbol{\eta}_1]-\boldsymbol{\eta}_1\right\|^2 + \left\|E[\boldsymbol{\eta}_2]-\boldsymbol{\eta}_2\right\|^2 \leq \chi^2_{k+1,\alpha}\right) \\
&= \Pr(\chi^2_{k+1} \leq \chi^2_{k+1,\alpha}) = 1-\alpha.
\end{aligned}$$

(iii) If $\lambda > \left((\chi^2_{k+1,\alpha})^{1/2} + (\chi^2_{r,\alpha})^{1/2}\right)^2$, then $G_{k+1-r}(\chi^2_{k+1,\alpha}-\lambda;0) = 0$, hence

$$l = \lim_{N\to\infty} \int_0^{\chi^2_{k+1,\alpha}} \Pr\left(\left\|E[\boldsymbol{\eta}_1]-\boldsymbol{\eta}_1\right\|^2 \leq \chi^2_{k+1,\alpha}-t;\ \left\|\boldsymbol{\eta}_1\right\|^2 > \chi^2_{r,\alpha}\right) \mathrm{d}G_{k+1-r}(t;0).$$

On the other hand if $\left((\chi^2_{k+1,\alpha})^{1/2} + (\chi^2_{r,\alpha})^{1/2}\right)^2 < \lambda$ then $(\chi^2_{k+1,\alpha})^{1/2} + (\chi^2_{r,\alpha})^{1/2} < \lambda^{1/2}$ and further

$$\left\|E[\boldsymbol{\eta}_1]-\boldsymbol{\eta}_1\right\|^2 \leq \chi^2_{k+1,\alpha}-t \Longrightarrow \left\|E[\boldsymbol{\eta}_1]-\boldsymbol{\eta}_1\right\| \leq \sqrt{\chi^2_{k+1,\alpha}-t},$$

then

$$\left\|E[\boldsymbol{\eta}_1]\right\| - \left\|\boldsymbol{\eta}_1\right\| = \lambda^{1/2} - \left\|\boldsymbol{\eta}_1\right\| \leq \left\|\boldsymbol{\eta}_1 - E[\boldsymbol{\eta}_1]\right\| \Longrightarrow \lambda^{1/2} \leq \left\|\boldsymbol{\eta}_1\right\| + \sqrt{\chi^2_{k+1,\alpha}-t}.$$

Since

$$\sqrt{\chi^2_{k+1,\alpha}-t} + (\chi^2_{r,\alpha})^{1/2} \leq (\chi^2_{k+1,\alpha})^{1/2} + (\chi^2_{r,\alpha})^{1/2} < \lambda^{1/2} \leq \left\|\boldsymbol{\eta}_1\right\| + \sqrt{\chi^2_{k+1,\alpha}-t}$$

hence $\|\boldsymbol{\eta}_1\| > (\chi^2_{r,\alpha})^{1/2} \Longrightarrow \|\boldsymbol{\eta}_1\|^2 > \chi^2_{r,\alpha}$.

Therefore

$$\begin{aligned} l &= \lim_{N\to\infty} \int_0^{\chi^2_{k+1,\alpha}} \Pr\left(\|E[\boldsymbol{\eta}_1] - \boldsymbol{\eta}_1\|^2 \le \chi^2_{k+1,\alpha} - t;\ \|\boldsymbol{\eta}_1\|^2 > \chi^2_{r,\alpha}\right) \mathrm{d}G_{k+1-r}(t;0) \\ &= \lim_{N\to\infty} \int_0^{\chi^2_{k+1,\alpha}} \Pr\left(\|E[\boldsymbol{\eta}_1] - \boldsymbol{\eta}_1\|^2 \le \chi^2_{k+1,\alpha} - t\right) \mathrm{d}G_{k+1-r}(t;0) \\ &= \Pr(\chi^2_{k+1} \le \chi^2_{k+1,\alpha}) = 1-\alpha. \quad \blacksquare \end{aligned}$$

## 4. Simulation results

We study the coverage probability (CP) of the confidence sets based on preliminary minimum ($\phi_1$,$\phi_2$)-divergence test estimators, $\widehat{\boldsymbol{\beta}}^{\mathrm{Pre}}_{\phi_1,\phi_2}$, under the null hypothesis as well as under contiguous alternative hypotheses using Monte Carlo experiments. Our idea is to check the advantage of using the minimum $\phi$-divergence estimators instead of the MLE as well as $\phi$-divergence test statistics instead of the classical likelihood-ratio test or Pearson test statistic. In our study we shall consider the power divergence measures introduced and studied in [2], the expression of the function associated with this family of divergence measures is

$$\phi_\lambda(x) = \begin{cases} \dfrac{x^{\lambda+1} - x - \lambda(x-1)}{\lambda(\lambda+1)}, & \lambda \ne 0, -1 \\ x\log x - x + 1, & \lambda = 0 \\ \log x + x - 1, & \lambda = -1. \end{cases}$$

This family will be used for testing and estimating. That it is to say, we consider for our study the family of preliminary test estimators

$$\widehat{\boldsymbol{\beta}}^{\mathrm{Pre}}_{\lambda_1,\lambda_2} \equiv \widehat{\boldsymbol{\beta}}^{\mathrm{Pre}}_{\phi_{\lambda_1},\phi_{\lambda_2}} = \widehat{\boldsymbol{\beta}}^{H_0}_{\phi_{\lambda_2}} I_{(0,\chi^2_{r,\alpha})}(T_N^{\phi_{\lambda_1},\phi_{\lambda_2}}) + \widehat{\boldsymbol{\beta}}_{\phi_{\lambda_2}} I_{[\chi^2_{r,\alpha},\infty)}(T_N^{\phi_{\lambda_1},\phi_{\lambda_2}}),$$

for some choices of the parameters $\lambda_1$ and $\lambda_2$. More concretely we shall use $\lambda_1 = -1/2, 0, 2/3, 1$ and 2 and $\lambda_2 = 0, 2/3$ and 1. It is interesting to note that for $\lambda_2 = 0$, $\widehat{\boldsymbol{\beta}}_{\phi_0}$ and $\widehat{\boldsymbol{\beta}}^{H_0}_{\phi_0}$ are the unrestricted and restricted MLE of $\boldsymbol{\beta}$ respectively. Note that $T_N^{\phi_0,\phi_0} = LR + o_P(1)$, where $LR$ is the likelihood-ratio test.

The logistic regression model considered in the simulation study consists of a dichotomous dependent variable and three normally distributed with zero mean and unit variance explanatory variables. We generated 10 000 samples of different sample sizes $\boldsymbol{n} = (n_1, \ldots, n_n)^{\mathrm{T}} \in \mathcal{N} = \{n^1, n^2, n^3, n^4, n^5\}$ with $n_i^1 = 15, n_i^2 = 30, n_i^3 = 80, i = 1, \ldots, 8, n^4 = (25, 25, 25, 25, 10, 10, 10, 10)$ and $n^5 = (40, 40, 15, 15, 5, 5, 25, 25)$. The regression coefficients $\boldsymbol{\beta}^{\mathrm{T}} = (\beta_0, \beta_1, \beta_2, \beta_3)$ were generated from a uniform over $(0, 2)$ .

We analyze the CP under the null hypothesis $\boldsymbol{\beta} \in \Theta_0$ as well as the contiguous alternative hypotheses

$$H_{1,N} : \boldsymbol{\beta}_N = \boldsymbol{\beta} + N^{-1/2}\boldsymbol{\Delta},$$

with $\boldsymbol{\beta} \in \Theta_0$ and different values of $\boldsymbol{\Delta}$, $\boldsymbol{\Delta}_1 = (0, 0, 0, 30)$ , $\boldsymbol{\Delta}_2 = (0, 0, 0, 20)$ , $\boldsymbol{\Delta}_3 = (0, 0, 0, -20)$ and $\boldsymbol{\Delta}_4 = (0, 0, 0, -30)$. We present the results obtained in Tables 1–5.

**Table 1**
CP of the estimates for $\boldsymbol{\Delta} = \mathbf{0}$

| $\lambda_1$ | $\lambda_2$ | $n^1$ | $n^2$ | $n^3$ | $n^4$ | $n^5$ |
|---|---|---|---|---|---|---|
| 0 | −1/2 | 0.9734 | 0.9670 | 0.9606 | 0.9757 | 0.9720 |
| | 0 | 0.9658 | 0.9670 | 0.9655 | 0.9692 | 0.9651 |
| | 2/3 | 0.9484 | 0.9551 | 0.9578 | 0.9437 | 0.9421 |
| | 1 | 0.9344 | 0.9440 | 0.9512 | 0.9259 | 0.9299 |
| | 2 | 0.9022 | 0.9015 | 0.9238 | 0.8790 | 0.8874 |
| 2/3 | −1/2 | 0.9738 | 0.9673 | 0.9607 | 0.9760 | 0.9714 |
| | 0 | 0.9647 | 0.9661 | 0.9658 | 0.9684 | 0.9628 |
| | 2/3 | 0.9461 | 0.9538 | 0.9571 | 0.9406 | 0.9396 |
| | 1 | 0.9312 | 0.9426 | 0.9506 | 0.9226 | 0.9260 |
| | 2 | 0.8990 | 0.8980 | 0.9227 | 0.8754 | 0.8817 |
| 1 | −1/2 | 0.9739 | 0.9672 | 0.9612 | 0.9759 | 0.9711 |
| | 0 | 0.9647 | 0.9663 | 0.9659 | 0.9677 | 0.9627 |
| | 2/3 | 0.9446 | 0.9530 | 0.9568 | 0.9391 | 0.9382 |
| | 1 | 0.9304 | 0.9417 | 0.9502 | 0.9219 | 0.9249 |
| | 2 | 0.8975 | 0.8972 | 0.9218 | 0.8738 | 0.8802 |

**Table 2**
CP of the estimates for $\boldsymbol{\Delta} = \boldsymbol{\Delta}_1$

| $\lambda_1$ | $\lambda_2$ | $n^1$ | $n^2$ | $n^3$ | $n^4$ | $n^5$ |
|---|---|---|---|---|---|---|
| 0 | −1/2 | 0.9572 | 0.9378 | 0.9000 | 0.9637 | 0.9558 |
| | 0 | 0.9066 | 0.8953 | 0.8618 | 0.9144 | 0.9147 |
| | 2/3 | 0.8231 | 0.8142 | 0.8027 | 0.8451 | 0.8560 |
| | 1 | 0.7891 | 0.7786 | 0.7748 | 0.8151 | 0.8326 |
| | 2 | 0.7145 | 0.7016 | 0.6973 | 0.7542 | 0.7787 |
| 2/3 | −1/2 | 0.9552 | 0.9322 | 0.8893 | 0.9626 | 0.9545 |
| | 0 | 0.9023 | 0.8883 | 0.8495 | 0.9109 | 0.9129 |
| | 2/3 | 0.8178 | 0.8027 | 0.7838 | 0.8413 | 0.8525 |
| | 1 | 0.7822 | 0.7669 | 0.7546 | 0.8119 | 0.8286 |
| | 2 | 0.7078 | 0.6886 | 0.6758 | 0.7509 | 0.7751 |
| 1 | −1/2 | 0.9542 | 0.9290 | 0.8845 | 0.9619 | 0.9541 |
| | 0 | 0.9012 | 0.8846 | 0.8430 | 0.9101 | 0.9123 |
| | 2/3 | 0.8162 | 0.7989 | 0.7751 | 0.8398 | 0.8515 |
| | 1 | 0.7804 | 0.7635 | 0.7435 | 0.8105 | 0.8277 |
| | 2 | 0.7059 | 0.6831 | 0.6653 | 0.7499 | 0.7725 |

**Table 3**
CP of the estimates for $\boldsymbol{\Delta} = \boldsymbol{\Delta}_2$

| $\lambda_1$ | $\lambda_2$ | $n^1$ | $n^2$ | $n^3$ | $n^4$ | $n^5$ |
|---|---|---|---|---|---|---|
| 0 | −1/2 | 0.9587 | 0.9481 | 0.9232 | 0.9664 | 0.9612 |
| | 0 | 0.9336 | 0.9309 | 0.9169 | 0.9429 | 0.9461 |
| | 2/3 | 0.8844 | 0.8933 | 0.8917 | 0.9014 | 0.9140 |
| | 1 | 0.8603 | 0.8741 | 0.8736 | 0.8770 | 0.8974 |
| | 2 | 0.7968 | 0.8220 | 0.8328 | 0.8178 | 0.8564 |
| 2/3 | −1/2 | 0.9581 | 0.9475 | 0.9200 | 0.9665 | 0.9610 |
| | 0 | 0.9330 | 0.9301 | 0.9136 | 0.9427 | 0.9460 |
| | 2/3 | 0.8841 | 0.8916 | 0.8888 | 0.9008 | 0.9139 |
| | 1 | 0.8594 | 0.8720 | 0.8708 | 0.8767 | 0.8971 |
| | 2 | 0.7961 | 0.8191 | 0.8296 | 0.8170 | 0.8558 |
| 1 | −1/2 | 0.9581 | 0.9466 | 0.9192 | 0.9665 | 0.9612 |
| | 0 | 0.9331 | 0.9296 | 0.9116 | 0.9425 | 0.9461 |
| | 2/3 | 0.8840 | 0.8911 | 0.8878 | 0.9006 | 0.9140 |
| | 1 | 0.8595 | 0.8714 | 0.8699 | 0.8766 | 0.8970 |
| | 2 | 0.7961 | 0.8182 | 0.8279 | 0.8161 | 0.8554 |

**Table 4**
CP of the estimates for $\boldsymbol{\Delta} = \boldsymbol{\Delta}_3$

| $\lambda_1$ | $\lambda_2$ | $n^1$ | $n^2$ | $n^3$ | $n^4$ | $n^5$ |
|---|---|---|---|---|---|---|
| 0 | −1/2 | 0.7939 | 0.7717 | 0.8021 | 0.7704 | 0.7928 |
| | 0 | 0.8547 | 0.8229 | 0.8350 | 0.8230 | 0.8238 |
| | 2/3 | 0.8945 | 0.8644 | 0.8574 | 0.8748 | 0.8696 |
| | 1 | 0.9048 | 0.8742 | 0.8647 | 0.8859 | 0.8803 |
| | 2 | 0.9185 | 0.8890 | 0.8728 | 0.8991 | 0.8899 |
| 2/3 | −1/2 | 0.8307 | 0.8057 | 0.8270 | 0.8189 | 0.8622 |
| | 0 | 0.8808 | 0.8487 | 0.8545 | 0.8646 | 0.8869 |
| | 2/3 | 0.9122 | 0.8824 | 0.8749 | 0.9000 | 0.9140 |
| | 1 | 0.9200 | 0.8919 | 0.8788 | 0.9080 | 0.9185 |
| | 2 | 0.9307 | 0.8999 | 0.8841 | 0.9144 | 0.9207 |
| 1 | −1/2 | 0.8466 | 0.8194 | 0.8366 | 0.8384 | 0.8835 |
| | 0 | 0.8934 | 0.8615 | 0.8646 | 0.8814 | 0.9073 |
| | 2/3 | 0.9197 | 0.8922 | 0.8824 | 0.9121 | 0.9265 |
| | 1 | 0.9267 | 0.9003 | 0.8844 | 0.9188 | 0.9281 |
| | 2 | 0.9354 | 0.9064 | 0.8891 | 0.9214 | 0.9298 |

From Tables 2 and 3 that correspond with $\boldsymbol{\Delta}_1$, $\boldsymbol{\Delta}_2$ it is clear that $\widehat{\boldsymbol{\beta}}^{\mathrm{Pre}}_{0,-1/2}$ is preferred to the rest. For $\boldsymbol{\Delta} = \mathbf{0}$, this estimator is the first or second best. However, for $\boldsymbol{\Delta}_3$, $\boldsymbol{\Delta}_4$ it can be seen from Tables 4 and 5 that $\widehat{\boldsymbol{\beta}}^{\mathrm{Pre}}_{1,2}$ is preferred to the rest. Therefore, $\widehat{\boldsymbol{\beta}}^{\mathrm{Pre}}_{2/3,2/3}$ can be considered as a good compromise for all the cases. Note that if we want to use the LRT ($\lambda_1 = 0$) statistic for the preliminary estimator, the largest CP corresponds to $\lambda_2 = -1/2$ for $\boldsymbol{\Delta} = \mathbf{0}$, $\boldsymbol{\Delta}_1$, $\boldsymbol{\Delta}_2$ and $\lambda_2 = 2$ for $\boldsymbol{\Delta}_3$ and $\boldsymbol{\Delta}_4$. So, $\widehat{\boldsymbol{\beta}}^{\mathrm{Pre}}_{0,2/3}$ is a good compromise between these two. On the other hand, we can fix the MLE ($\lambda_2 = 0$) for obtaining the preliminary

**Table 5**
CP of the estimates for $\boldsymbol{\Delta} = \boldsymbol{\Delta}_4$

| $\lambda_1$ | $\lambda_2$ | $n^1$ | $n^2$ | $n^3$ | $n^4$ | $n^5$ |
|---|---|---|---|---|---|---|
| 0 | $-1/2$ | 0.9613 | 0.9473 | 0.9353 | 0.9552 | 0.9456 |
| | 0 | 0.9637 | 0.9581 | 0.9472 | 0.9549 | 0.9605 |
| | 2/3 | 0.9670 | 0.9606 | 0.9486 | 0.9608 | 0.9660 |
| | 1 | 0.9688 | 0.9605 | 0.9469 | 0.9616 | 0.9669 |
| | 2 | 0.9708 | 0.9585 | 0.9402 | 0.9625 | 0.9665 |
| 2/3 | $-1/2$ | 0.9618 | 0.9483 | 0.9385 | 0.9575 | 0.9482 |
| | 0 | 0.9642 | 0.9588 | 0.9488 | 0.9579 | 0.9616 |
| | 2/3 | 0.9674 | 0.9610 | 0.9493 | 0.9621 | 0.9664 |
| | 1 | 0.9692 | 0.9609 | 0.9479 | 0.9631 | 0.9671 |
| | 2 | 0.9711 | 0.9589 | 0.9407 | 0.9635 | 0.9669 |
| 1 | $-1/2$ | 0.9622 | 0.9490 | 0.9395 | 0.9583 | 0.9489 |
| | 0 | 0.9647 | 0.9593 | 0.9491 | 0.9588 | 0.9624 |
| | 2/3 | 0.9678 | 0.9614 | 0.9497 | 0.9629 | 0.9665 |
| | 1 | 0.9696 | 0.9612 | 0.9483 | 0.9636 | 0.9674 |
| | 2 | 0.9713 | 0.9590 | 0.9411 | 0.9638 | 0.9671 |

estimator and to look for the best statistic. In this case, for $\boldsymbol{\Delta} = \mathbf{0}, \boldsymbol{\Delta}_1, \boldsymbol{\Delta}_2$ LRT is the best but for $\boldsymbol{\Delta}_3, \boldsymbol{\Delta}_4$ the minimum chi-square statistic is the best, so a good compromise for all $\boldsymbol{\Delta}$ seems to be the statistic corresponding with $\lambda_1 = 2/3$.

## Acknowledgements

The research in this paper was supported in part by Grants MTM2006-06872 and UCM2007-910707.

## References

[1] P.C. Chiou, A.K.Md.E. Saleh, Preliminary test confidence sets for the mean of a multivariate normal distribution, Journal of Propagations in Probability and Statistics 2 (2002) 177–189.
[2] N. Cressie, T.R.C. Read, Multinomial goodness-of-fit tests, Journal of the Royal Statistical Society, Series B 46 (1984) 440–464.
[3] M.L. Menéndez, J.A. Pardo, L. Pardo, Phi-divergence test statistics for testing linear hypotheses in logistic regression models, Communication in Statistics (Theory and Methods) 37 (4) (2008) 494–507.
[4] M.L. Menéndez, L. Pardo, M.C. Pardo, Preliminary phi-divergence test estimators for linear restrictions in a logistic regression model, Statistical Papers (2008), doi:10.1007/s00362-007-0078-z.
[5] J.A. Pardo, L. Pardo, M.C. Pardo, Minimum $\phi$-divergence estimators in Logistic Regression Models, Statistical Papers 47 (2005) 92–108.
[6] L. Pardo, Statistical Inference Based on Divergence Measures, Chapman & Hall/CRC, New York, 2006.
[7] A.K.Md.E. Saleh, Theory of Preliminary Test and Stein-Type Estimation with Applications, John Wiley, 2006.