TWMS J. App. and Eng. Math. V.14, N.4, 2024, pp. 1374-1389

TESTING LOCAL HYPOTHESES WITH DIFFERENT TYPES OF INCOMPLETE DATA

M. BOUKELOUA[1], §

ABSTRACT. In this work, we consider a general framework of incomplete data which includes many types of censoring and truncation models. Under this framework and assuming that the distribution of interest has a parametric form, we propose local tests for simple and composite hypotheses on the parameter. These tests are based on the $\phi-$divergences, Wald and Rao statistics. We study the asymptotic behaviour of these statistics under the null hypothesis. For the $\phi-$divergences statistics, we study the asymptotic behaviour under the alternative as well. This allows us to approximate the power function of the proposed tests. We also propose local tests of homogeneity which serve to compare the distributions of two samples. Finally, we present the results of an application on real data.

Keywords: Local tests, censored data, truncated data, $\phi-$divergences, tests of homogeneity.

AMS Subject Classification: 62F03, 62F05, 62N03.

1. Introduction

Hypothesis testing constitutes an essential issue in statistics. One of the most popular types of the tests of hypothesis are the parametric tests of the simple null hypothesis $\mathcal{H}_0 : \theta_T = \theta_0$, against the alternative $\mathcal{H}_1 : \theta_T \neq \theta_0$, where $\theta \in \Theta$ is a parameter that describes the distribution of the population, θ_T is the true value of θ and θ_0 is a fixed value in the parameter space Θ. Many tests of this types of hypotheses have been studied in the literature such as Wald, Rao and the likelihood ratio tests. Recently, the theory of $\phi-$divergences between measures, introduced by [1], has been widely applied in statistics. [2, 3] used this theory to study some parametric and semiparametric models. [4] and [5] used it to study semiparametric copula models for complete and censored data, respectively. Furthermore, [6] proposed $\phi-$divergence tests of the hypothesis $\mathcal{H}_0$ against $\mathcal{H}_1$. All the tests we have cited above compare the two distributions characterized by θ_T and

[1] Laboratoire de Génie des Procédés pour le Développement Durable et les Produits de Santé (LGPDDPS)-Ecole Nationale Polytechnique de Constantine - Algeria.
Laboratoire de Biostatistique, Bioinformatique et Méthodologie Mathématique Appliquées aux Sciences de la Santé (BIOSTIM)- Faculté de Médecine - Université Salah Boubnider Constantine 3, Algeria.
e-mail: boukeloua.mohamed@gmail.com; ORCID: https://orcid.org/0000-0002-5522-498X.
§ Manuscript received: March 25, 2023; accepted: September 19, 2023.
TWMS Journal of Applied and Engineering Mathematics, Vol.14, No.4

θ_0 globally, i.e., on the whole support of the variable of interest. However, it happens in certain situations that the two distributions are different but very close on a part of this support. [7] give a real-life example of such a situation. For this kind of situations, the conclusion of the test on the whole support of the distribution may differ if we focus on a specific part of the support. Motivated by this fact, [8] introduced local $\phi-$divergences that allow to quantify the dispersion between two distributions only on a part of their support. Using these local $\phi-$divergences, [7] proposed local test statistics of the hypothesis $\mathcal{H}_0$ against $\mathcal{H}_1$. They gave the asymptotic distribution of these statistics under both the null and the alternative hypotheses.

The tests we have discussed until now are based on complete observations of the variable of interest. However, in the practice, some censoring and/or truncation phenomena may prevent the observation of the variable of interest and provide only a partial information about it. The presence of such phenomena affects considerably the statistical investigation of the data. In the present paper, we consider a general framework of incomplete data which includes some types of censoring and truncation. Under this framework, we propose local $\phi-$divergence tests on the parameter θ. The $\phi-$divergence technique constitutes a useful tool in local hypothesis testing since it facilitates the construction of the local test statistic. In fact, it suffices to multiply the integrand in the $\phi-$divergence by a kernel allowing to focus on a specific part of the integration domain. Moreover, the $\phi-$divergence technique allows to determine the asymptotic distribution of the test statistic under the alternative hypothesis, which is not possible for classical approaches. This helps to give an approximation to the power function of the $\phi-$divergence based test. Concerning our contributions in this paper, we draw on the work of [7] to propose local $\phi-$divergence tests of the simple null hypothesis $\mathcal{H}_0 : \theta_T = \theta_0$, under the general framework of incomplete data. We give the asymptotic distribution of the statistics of these tests under both the null and the alternative hypotheses. This allows us to approximate the power of these tests. We also propose local Wald and Rao type tests and we provide the asymptotic distribution of their statistics under $\mathcal{H}_0$. Then, we study local composite null hypotheses with incomplete data. For these ones, we propose local $\phi-$divergence, Wald, Rao and Lagrange multipliers tests and we provide the asymptotic distribution of their statistics under the null hypothesis. We also study the asymptotic behaviour of the $\phi-$divergence tests statistics under the alternative hypothesis. Furthermore, we consider the problem of comparison of the distributions of two samples of incomplete data. Following [7], we propose local $\phi-$divergence and Wald tests of homogeneity and we provide the asymptotic distribution of their statistics under the null hypothesis. We also give the asymptotic distribution of the $\phi-$divergence tests statistics under the alternative hypothesis. Finally, we apply our proposed tests on a real data set of the time of breast retraction for breast cancer patients. The rest of the paper is organized as follows. In Section 2, we present some types of censored and truncated data on which our study will be based. In Section 3, we give our main results. An application on real data is presented in Section 4 and Section 5 gives some conclusions and perspectives. The proofs are relegated to Appendix A.

2. Some types of incomplete data

We start by presenting some types of censoring and truncation and we give the form of the likelihood function for each type. Let X be a positive real random variable (r.r.v.) of interest. We assume that the distribution of X belongs to a parametric family $\{P_\theta, \, \theta \in \Theta\}$ (Θ being an open set of $\mathbb{R}^d$), dominated by a $\sigma-$finite measure m. Denote by f_θ the

Radon-Nikodym derivative of P_θ with respect to m. We also assume that X may be censored and/or truncated and we denote by (Z, Δ) the couple of the observed variables. In what follows, $(Z_i, \Delta_i)_{1\leq i\leq n}$ represents a sample of independent and identically distributed (i.i.d.) copies of the couple (Z, Δ) and (z_i, δ_i) represents a realization of (Z_i, Δ_i). From now on, for any random variable V, P_V, F_V and S_V denote respectively the probability distribution, the distribution function and the survival function of V; and when P_V is absolutely continuous with respect to m, $f_V = \frac{dP_V}{dm}$ represents its Radon-Nikodym derivative. Moreover, for any function $\psi : \mathbb{R} \to \mathbb{R}$, we denote by $\psi(x^-) = \lim_{t \overset{<}{\to} x} \psi(t)$, when the limit exists and for any vector or matrix A, we denote by $A^\top$ the transpose of A. Here are some types of incomplete data.

Right censored data

In this case, we observe the variables $Z = \min(X, R)$ and $\Delta = 1_{\{X\leq R\}}$, where R is the right censoring variable assumed to be positive and independent of X and $1_{\{.\}}$ denotes the indicator function. The likelihood function of (Z, Δ) is given by

$$L(\theta) = \prod_{i=1}^{n} \left(f_\theta(z_i)S_R(z_i^-)\right)^{\delta_i} \left(S_X(z_i;\theta)f_R(z_i)\right)^{1-\delta_i}.$$

Since we are interested in the parameter θ, we will only consider the functions that depend on θ. So, we study the following pseudo-likelihood function

$$\mathcal{L}(\theta) = \prod_{i=1}^{n} g_\theta(z_i, \delta_i), \text{ where}$$

$$g_\theta(z_i, \delta_i) = f_\theta(z_i)^{\delta_i} S_X(z_i;\theta)^{1-\delta_i}. \tag{1}$$

Doubly censored data

In this case, we observe the variables $Z = \max(\min(X, R), L)$ and

$$\Delta = \begin{cases} 1, & \text{if } L < X \leq R \\ 2, & \text{if } \ X > R \\ 3, & \text{if } \ X \leq L, \end{cases}$$

where R (resp. L) is the right (resp. left) censoring variable with $0 \leq L \leq R$ almost surely $(a.s.)$ and (L, R) is independent from X.

The likelihood function of (Z, Δ) is given by

$$L(\theta) = \prod_{i=1}^{n} \left(f_\theta(z_i)\left(S_R(z_i^-) - S_L(z_i^-)\right)\right)^{1\{\delta_i=1\}} \left(S_X(z_i;\theta)f_R(z_i)\right)^{1\{\delta_i=2\}} \left(F_X(z_i;\theta)f_L(z_i)\right)^{1\{\delta_i=3\}}.$$

As in the previous case, we study the following pseudo-likelihood function

$$\mathcal{L}(\theta) = \prod_{i=1}^{n} g_\theta(z_i, \delta_i), \text{ where}$$

$$g_\theta(z_i, \delta_i) = f_\theta(z_i)^{1\{\delta_i=1\}} S_X(z_i;\theta)^{1\{\delta_i=2\}} F_X(z_i;\theta)^{1\{\delta_i=3\}}. \tag{2}$$

Interval censored data, case 1 (current status data)

In this case, we observe the couple (Z, Δ), where Z is a positive random variable independent of X and $\Delta = 1_{\{X\leq Z\}}$. The likelihood function of (Z, Δ) is given by

$$L(\theta) = \prod_{i=1}^{n} F_X(z_i;\theta)^{\delta_i} S_X(z_i;\theta)^{1-\delta_i} f_Z(z_i)$$

and the pseudo-likelihood function is given by

$$\mathcal{L}(\theta) = \prod_{i=1}^{n} g_\theta(z_i, \delta_i), \text{ where}$$

$$g_\theta(z_i, \delta_i) = F_X(z_i; \theta)^{\delta_i} S_X(z_i; \theta)^{1-\delta_i}. \tag{3}$$

Interval censored data, case 2
In this case, we observe the variables $Z = (R, L)$ and

$$\Delta = \begin{cases} 1, & \text{if } L < X \leq R \\ 2, & \text{if } \ X > R \\ 3, & \text{if } \ X \leq L, \end{cases}$$

where R and L are positive variables such that $L < R$ *a.s.* and (R, L) is independent from X. Let $(r_i, l_i, \delta_i)_{1 \leq i \leq n}$ be a realization of the sample $(R_i, L_i, \Delta_i)_{1 \leq i \leq n}$. The likelihood function of (Z, Δ) is given by

$$L(\theta) = \prod_{i=1}^{n} (F_X(r_i; \theta) - F_X(l_i; \theta))^{1_{\{\delta_i = 1\}}} S_X(r_i; \theta)^{1_{\{\delta_i = 2\}}} F_X(l_i; \theta)^{1_{\{\delta_i = 3\}}} f_{R,L}(r_i, l_i)$$

and the pseudo-likelihood function is given by

$$\mathcal{L}(\theta) = \prod_{i=1}^{n} g_\theta(z_i, \delta_i), \text{ where}$$

$$g_\theta(z_i, \delta_i) = g_\theta(r_i, l_i, \delta_i) = (F_X(r_i; \theta) - F_X(l_i; \theta))^{1_{\{\delta_i = 1\}}} S_X(r_i; \theta)^{1_{\{\delta_i = 2\}}} F_X(l_i; \theta)^{1_{\{\delta_i = 3\}}}. \tag{4}$$

The LTRC data model
Let R (resp. L) be a positive variable of censoring (resp. truncation) independent of X. In the left truncated and right censored (LTRC) data model, we observe $Z = (Y, L)$ (where $Y = \min(X, R)$) and $\Delta_1 = 1_{\{X \leq R\}}$ whenever $Y \geq L$ (i.e., when the observation is not left truncated). We also observe the truncation indicator $\Delta_2 = 1_{\{Y \geq L\}}$. Set $\Delta = (\Delta_1, \Delta_2)$ and $(y_i, l_i, \delta_{1i}, \delta_{2i})_{1 \leq i \leq n}$ a realization of the sample $(Y_i, L_i, \Delta_{1i}, \Delta_{2i})_{1 \leq i \leq n}$. The likelihood function is given by

$$L(\theta) = \prod_{i=1}^{n} \left\{ \left(f_\theta(y_i) S_R(y_i^-)\right)^{\delta_{1i}\delta_{2i}} \left[\frac{f_\theta(y_i) S_R(y_i^-)}{S_X(l_i; \theta)}\right]^{\delta_{1i}(1-\delta_{2i})} (S_X(y_i; \theta) f_R(y_i))^{(1-\delta_{1i})\delta_{2i}} \left[\frac{S_X(y_i; \theta) f_R(y_i)}{S_X(l_i; \theta)}\right]^{(1-\delta_{1i})(1-\delta_{2i})} \right\}$$

and the pseudo-likelihood function is given by

$$\mathcal{L}(\theta) = \prod_{i=1}^{n} g_\theta(z_i, \delta_i), \text{ where}$$

$$g_\theta(z_i, \delta_i) = g_\theta(y_i, l_i, \delta_{1i}, \delta_{2i}) = f_\theta(y_i)^{\delta_{1i}\delta_{2i}} \left[\frac{f_\theta(y_i)}{S_X(l_i; \theta)}\right]^{\delta_{1i}(1-\delta_{2i})} S_X(y_i; \theta)^{(1-\delta_{1i})\delta_{2i}} \left[\frac{S_X(y_i; \theta)}{S_X(l_i; \theta)}\right]^{(1-\delta_{1i})(1-\delta_{2i})}. \tag{5}$$

3. Main results

We will propose local tests on the parameter θ, under a general framework of incomplete data which includes all the types of censoring and truncation described in the previous section. Firstly, we begin by defining this general framework.

3.1. **General framework of incomplete data.** In the general framework, we assume that the variable of interest X is not completely observed. So instead of observing X, we observe the variables Z and Δ, where Δ is a discrete variable that indicates which variable is observed (the variable of interest or another latent variable). Under this framework, the pseudo-likelihood function is defined by

$$\mathcal{L}(\theta) = \prod_{i=1}^{n} g_\theta(z_i, \delta_i),$$

where g_θ is the pseudo-density of (Z, Δ). In the particular cases studied in the previous section, g_θ has one of the forms (1)–(5), depending on the considered type of incomplete data.

3.2. **Local $\phi-$divergences.** Our study is based on the local $\phi-$ divergences between the two functions $g_{\theta^{(1)}}$ and $g_{\theta^{(2)}}$ ($\theta^{(1)}$ and $\theta^{(2)}$ being two elements of Θ). Let E and F be the support of the variables Z and Δ respectively. These supports vary according to the considered case. For example, for right censored data, when Z is absolutely continuous, E is a subset of $\mathbb{R}_+$ and when Z is discrete, E is a subset of $\mathbb{N}$. As for F, it is equal to $\{0, 1\}$. The dominating measure m is the Lebesgue measure when Z is absolutely continuous and it is the counting measure when Z is discrete. We also denote by μ the counting measure on F. Following [8] and [7], we define the local $\phi-$ divergence between $g_{\theta^{(1)}}$ and $g_{\theta^{(2)}}$ by

$$D_\phi^\omega\left(g_{\theta^{(1)}}, g_{\theta^{(2)}}\right) = D_\phi^\omega\left(\theta^{(1)}, \theta^{(2)}\right) = \int_F \int_E g_\omega(z,\delta) g_{\theta^{(2)}}(z,\delta) \phi\left(\frac{g_{\theta^{(1)}}(z,\delta)}{g_{\theta^{(2)}}(z,\delta)}\right) dm(z) d\mu(\delta),$$

where ω is a fixed point of Θ and ϕ is a real valued convex function defined on $[0, +\infty[$. We assume that ϕ belongs to the class of convex functions

$$\Phi = \left\{\phi : \phi \text{ is strictly convex at } 1, \phi(1) = \phi'(1) = 0, 0\phi\left(\frac{0}{0}\right) = 0,\right.$$
$$\left. 0\phi\left(\frac{u}{0}\right) = u \lim_{v \to +\infty} \frac{\phi(v)}{v}\right\}.$$

Local $\phi-$divergences are based on the choice of the kernel g_ω for the fixed value $\omega \in \Theta$, which determines the part of the support of X, on which we focus our analysis.

The local $\phi-$ divergence $D_\phi^\omega\left(\theta^{(1)}, \theta^{(2)}\right)$ satisfies $D_\phi^\omega\left(\theta^{(1)}, \theta^{(2)}\right) \geq 0$ with equality if and only if $g_{\theta^{(1)}} = g_{\theta^{(2)}}$. Moreover, we assume that the parametric family $\{g_\theta, \theta \in \Theta\}$ is identifiable, i.e., $g_{\theta^{(1)}} = g_{\theta^{(2)}}$ implies that $\theta^{(1)} = \theta^{(2)}$ for all $\theta^{(1)}$ and $\theta^{(2)} \in \Theta$.

In the sequel, we will be interested in some tests on the parameter θ with local simple and composite null hypotheses. We will also study some local tests of homogeneity.

3.3. **Local tests with simple null hypothesis.** In this paragraph, we will study local tests with simple null hypothesis. In particular, we will construct local $\phi-$divergence, Wald and Rao test statistics and we will determine their asymptotic distributions under the null hypothesis, using standard assumptions in the parametric setting. For the local $\phi-$divergence statistics, we will also determine the asymptotic distribution under the alternative hypothesis, which allows us to give an approximation to the power function.

Let θ_T be the true value of θ and θ_0 be a fixed point in Θ. As in [7], we consider the local null hypothesis defined by

$$\mathcal{H}_0 : g_{\theta_T}(z,\delta) = g_{\theta_0}(z,\delta) \text{ for a given } g_\omega, \omega \in \Theta,$$

which we briefly write

$$\mathcal{H}_0^w : \theta_T = \theta_0.$$

To test the hypothesis $\mathcal{H}_0^\omega$ against the alternative $\mathcal{H}_1^\omega : \theta_T \neq \theta_0$, we make use of the following local $\phi-$divergence test statistic

$$T_{\phi,n}^\omega(\widehat{\theta}_n, \theta_0) = \frac{2nD_\phi^\omega\left(\widehat{\theta}_n, \theta_0\right)}{\phi''(1)},$$

where $\widehat{\theta}_n$ is the maximum pseudo-likelihood estimate (MPLE) of θ_T.
We will also study Wald and Rao tests of the hypothesis $\mathcal{H}_0^\omega$ against $\mathcal{H}_1^\omega$. For that, let us define the Fisher information matrix

$$I(\theta) = \left(\int_F \int_E f_{Z,\Delta}(z,\delta;\theta) \frac{\partial \log g_\theta(z,\delta)}{\partial \theta_i} \frac{\partial \log g_\theta(z,\delta)}{\partial \theta_j} dm(z) d\mu(\delta)\right)_{1 \leq i,j \leq d}$$

and the local information matrix

$$I^\omega(\theta) = \left(\int_F \int_E g_\omega(z,\delta) g_\theta(z,\delta) \frac{\partial \log g_\theta(z,\delta)}{\partial \theta_i} \frac{\partial \log g_\theta(z,\delta)}{\partial \theta_j} dm(z) d\mu(\delta)\right)_{1 \leq i,j \leq d}.$$

The local Wald and Rao test statistics are defined respectively by

$$W_n^\omega = n\left(\widehat{\theta}_n - \theta_0\right)^\top I^\omega(\widehat{\theta}_n)\left(\widehat{\theta}_n - \theta_0\right)$$

and

$$R_n^\omega = \frac{1}{n} U_n^\top(\theta_0) I^\omega(\theta_0)^{-1} U_n(\theta_0),$$

where

$$U_n(\theta_0) = \left(\sum_{i=1}^n \frac{\partial \log g_\theta(Z_i,\Delta_i)}{\partial \theta_1}, \ldots, \sum_{i=1}^n \frac{\partial \log g_\theta(Z_i,\Delta_i)}{\partial \theta_d}\right)_{\theta=\theta_0}^\top.$$

We will give the asymptotic distribution of the test statistics $T_{\phi,n}^\omega(\widehat{\theta}_n, \theta_0)$, W_n^ω and R_n^ω under the following assumptions.

H1: The third partial derivatives of $g_\theta(z,\delta)$ with respect to θ exist for all $\theta \in \Theta$.

H2: The first, second and third partial derivatives of $g_\theta(z,\delta)$ with respect to θ are absolutely bounded from functions $\alpha(z,\delta)$, $\beta(z,\delta)$ and $\gamma(z,\delta)$ respectively and $\int_F \int_E \alpha(z,\delta) dm(z) d\mu(\delta) < \infty$, $\int_F \int_E \beta(z,\delta) dm(z) d\mu(\delta) < \infty$ and $\int_F \int_E \gamma(z,\delta) f_{Z,\Delta}(z,\delta;\theta) dm(z) d\mu(\delta) < \infty$.

H3: For each $\theta \in \Theta$, the matrices $I(\theta)$ and $I^\omega(\theta)$ exist, they are positive definite and their elements are continuous functions of θ.

H4: The function $\phi \in \Phi$ is twice continuously differentiable with $\phi''(1) > 0$.

H5: For each $\theta_0 \in \Theta$ there exists an open neighborhood $N(\theta_0)$ such that for all $\theta \in N(\theta_0)$ and $1 \leq i,j \leq d$ we have

$$\begin{aligned} &\frac{\partial}{\partial \theta_i} \int_F \int_E g_\omega(z,\delta) g_{\theta_0}(z,\delta) \phi\left(\frac{g_\theta(z,\delta)}{g_{\theta_0}(z,\delta)}\right) dm(z) d\mu(\delta) \\ &= \int_F \int_E \frac{\partial}{\partial \theta_i}\left(g_\omega(z,\delta) g_{\theta_0}(z,\delta) \phi\left(\frac{g_\theta(z,\delta)}{g_{\theta_0}(z,\delta)}\right)\right) dm(z) d\mu(\delta), \end{aligned}$$

$$\frac{\partial^2}{\partial\theta_i\partial\theta_j}\int_F\int_E g_\omega(z,\delta)g_{\theta_0}(z,\delta)\phi\left(\frac{g_\theta(z,\delta)}{g_{\theta_0}(z,\delta)}\right)dm(z)d\mu(\delta)$$
$$=\int_F\int_E \frac{\partial^2}{\partial\theta_i\partial\theta_j}\left(g_\omega(z,\delta)g_{\theta_0}(z,\delta)\phi\left(\frac{g_\theta(z,\delta)}{g_{\theta_0}(z,\delta)}\right)\right)dm(z)d\mu(\delta)$$

and these expressions are continuous on $N(\theta_0)$.

In all the sequel, $(V_n)_{n\in\mathbb{N}}$ represents a sequence of independent and identically distributed standard normal random variables.

Theorem 3.1. *Under $\mathcal{H}_0^\omega$ and the assumptions* **H1 − H5**, *the statistics $T^\omega_{\phi,n}(\widehat{\theta}_n,\theta_0)$ and W_n^ω converge in distribution to $\sum_{i=1}^r a_iV_i^2$, where $r = rank\left(I(\theta_0)^{-1}I^\omega(\theta_0)I(\theta_0)^{-1}\right)$ and $a_1,\dots,a_r$ are the non zero eigenvalues of the matrix $I^\omega(\theta_0)I(\theta_0)^{-1}$.*
Moreover, the statistic R_n^ω converges in distribution to $\sum_{i=1}^s b_iV_i^2$, where $s = rank\left(I(\theta_0)I^\omega(\theta_0)^{-1}I(\theta_0)\right)$ and $b_1,\dots,b_s$ are the non zero eigenvalues of the matrix $I^\omega(\theta_0)^{-1}I(\theta_0)$.

From this theorem, the critical region of the local $\phi-$divergence test at level $\alpha\in(0,1)$ is $CR=\{T^\omega_{\phi,n}(\widehat{\theta}_n,\theta_0)>q_{1-\alpha}\}$, where $q_{1-\alpha}$ is the $(1-\alpha)-$quantile of the limiting distribution of $T^\omega_{\phi,n}(\widehat{\theta}_n,\theta_0)$. In the practice, the quantile $q_{1-\alpha}$ can be approximated by a Monte Carlo approach as described in [7]. The critical regions of the local Wald and Rao tests can be defined in the same way.

The next theorem gives the asymptotic distribution of $T^\omega_{\phi,n}(\widehat{\theta}_n,\theta_0)$ under the alternative hypothesis $\mathcal{H}_1^\omega$.

Theorem 3.2. *Under $\mathcal{H}_1^\omega$ and the assumptions* **H1-H5**, *we have*

$$\sqrt{n}\left(D^\omega_\phi(\widehat{\theta}_n,\theta_0)-D^\omega_\phi(\theta_T,\theta_0)\right)\xrightarrow{\mathcal{D}}\mathcal{N}(0,\sigma^2),$$

where $\sigma^2=T^\top I(\theta_T)^{-1}T$ with $T=(t_1,\dots,t_d)^\top$, $t_i=\left.\frac{\partial D^\omega_\phi(\theta,\theta_0)}{\partial\theta_i}\right|_{\theta=\theta_T}$, $1\le i\le d$.

Thanks to this theorem, we can approximate the power function $\theta_T\in\Theta\mapsto\pi(\theta_T)=P_{\theta_T}(CR)$. Indeed, we have

$$\pi(\theta_T)\approx 1-F_{\mathcal{N}}\left(\frac{\sqrt{n}}{\sigma}\left(\frac{q_{1-\alpha}}{2n}\phi''(1)-D^\omega_\phi(\theta_T,\theta_0)\right)\right),$$

where $F_{\mathcal{N}}$ is the cumulative distribution function of the standard normal distribution. From this approximation, we can compute the sample size that ensures a specified power π. Let n_0 be the positive root of the equation

$$\pi=1-F_{\mathcal{N}}\left(\frac{\sqrt{n}}{\sigma}\left(\frac{q_{1-\alpha}}{2n}\phi''(1)-D^\omega_\phi(\theta_T,\theta_0)\right)\right)$$

which can be written into

$$n_0=\frac{a+b-\sqrt{a(a+2b)}}{2D^\omega_\phi(\theta_T,\theta_0)^2},$$

where $a=\sigma^2\left[F_{\mathcal{N}}^{-1}(1-\pi)\right]^2$ and $b=q_{1-\alpha}\phi''(1)D^\omega_\phi(\theta_T,\theta_0)$. The required sample size is then $\lfloor n_0\rfloor+1$ ($\lfloor x\rfloor$ denotes the integer part of x). In practice, we can replace θ_T by $\widehat{\theta}_n$ in $D^\omega_\phi(\theta_T,\theta_0)$ and σ and $q_{1-\alpha}$ can be estimated by the Monte Carlo approach described in [7].

3.4. **Local tests with composite null hypothesis.** In this paragraph, we will study local tests with composite null hypothesis. In particular, we will construct local $\phi-$divergence, Wald, Rao and Lagrange multipliers test statistics and we will determine their asymptotic distributions under the null hypothesis. We will also determine the asymptotic behaviour of the local $\phi-$divergence statistics under the alternative hypothesis, which allows us to conclude that the local $\phi-$divergence test is consistent.
Consider the local composite null hypothesis

$$\mathcal{H}_0^\omega : h(\theta_T) = 0 \text{ against the alternative } \mathcal{H}_1^\omega : h(\theta_T) \neq 0,$$

where h is a function defined from Θ to $\mathbb{R}^p$ $(p < d)$.
The hypothesis $\mathcal{H}_0^\omega$ can be transformed to a simple one by considering a function $\widetilde{h} : B \subseteq \mathbb{R}^{d-p} \longrightarrow \Theta$ so that $\mathcal{H}_0^\omega$ and $\mathcal{H}_1^\omega$ are equivalent to the hypotheses

$$\mathcal{H}_0^\omega : \theta_T = \widetilde{h}(\beta) \text{ and } \mathcal{H}_1^\omega : \theta_T \neq \widetilde{h}(\beta),$$

for some $\beta \in B$.
Let $\widetilde{\theta}_n$ be the MPLE of θ_T satisfying the constraint $h(\widetilde{\theta}_n) = 0$. Under the assumptions

H6: For all $\theta \in \Theta$ such that $h(\theta) = 0$, the matrix $H(\theta) = \nabla_\theta h(\theta)$ exists, it has full rank and its elements are continuous functions of θ

and

H7: For all $\beta \in B$, the matrix $\widetilde{H}(\beta) = \nabla_\beta \widetilde{h}(\beta)$ exists, it has full rank and its elements are continuous functions of β,

we have in view of lemma 3.1 of [7]

$$\sqrt{n}\left(\widetilde{\theta}_n - \theta_T\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \left(I_d - I^{-1}(\theta_T) B(\theta_T)\right) I^{-1}(\theta_T) \left(I_d - B(\theta_T) I^{-1}(\theta_T)\right)\right),$$

where $B(\theta) = H(\theta)\left[H^\top(\theta) I^{-1}(\theta) H(\theta)\right]^{-1} H^\top(\theta)$.
We also set $A(\theta) = B(\theta) I^{-1}(\theta) I^\omega(\theta) I^{-1}(\theta) B(\theta)$.
We are interested in the following test statistics of the hypothesis $\mathcal{H}_0^\omega$ against $\mathcal{H}_1^\omega$.

- The local $\phi-$divergence statistic

$$T_{\phi,n}^\omega\left(\widehat{\theta}_n, \widetilde{\theta}_n\right) = \frac{2n D_\phi^\omega\left(\widehat{\theta}_n, \widetilde{\theta}_n\right)}{\phi''(1)}.$$

- The Wald statistic

$$W_n^{w,c} = n h(\widehat{\theta}_n)^\top \left(H(\widehat{\theta}_n) I^\omega(\widehat{\theta}_n)^{-1} H(\widehat{\theta}_n)\right)^{-1} h(\widehat{\theta}_n).$$

- The Rao statistic

$$R_n^{\omega,c} = \frac{1}{n} U_n(\widetilde{\theta}_n)^\top \left[I^\omega(\widetilde{\theta}_n)\right]^{-1} U_n(\widetilde{\theta}_n).$$

- The Lagrange multipliers statistic:
  Consider the constrained optimization problem

$$\begin{cases} \max_{\theta \in \Theta} \mathcal{L}(\theta) \\ h(\theta) = 0. \end{cases}$$

  The Lagrangian of this problem is

$$\mathcal{L}(\theta) + h(\theta)^\top \lambda,$$

where λ is the Lagrange multiplier.
Let $(\widetilde{\theta}_n, \widetilde{\lambda}_n)$ be the solution of this problem, the Lagrange multipliers test statistic is defined by

$$M_n^{\omega,c} = \frac{1}{n}\widetilde{\lambda}_n^\top \Gamma^\omega(\widetilde{\theta}_n)\widetilde{\lambda}_n,$$

where $\Gamma^\omega(\theta) = H^\top(\theta)\left[I^\omega(\theta)\right]^{-1} H(\theta)$.
We also set $\Gamma(\theta) = H^\top(\theta)\left[I(\theta)\right]^{-1} H(\theta)$.

Now, we will give the asymptotic distributions of these statistics under $\mathcal{H}_0^\omega$.

Theorem 3.3. *Under $\mathcal{H}_0^\omega$ and assumptions* **H1-H3**, **H6** *and* **H7**, *we have*

i) If **H4** *and* **H5** *are satisfied, then*

$$T_{\phi,n}^\omega\left(\widehat{\theta}_n, \widetilde{\theta}_n\right) \xrightarrow{\mathcal{D}} \sum_{i=1}^{r_1} a_i V_i^2,$$

where $r_1 = rank(I(\theta_T)^{-1}A(\theta_T)I(\theta_T)^{-1})$ and $a_1, \ldots, a_{r_1}$ are the non zero eigenvalues of $A(\theta)I(\theta)^{-1}$.

ii)

$$W_n^{\omega,c} \xrightarrow{\mathcal{D}} \sum_{i=1}^{r_2} b_i V_i^2,$$

where

$$\begin{aligned} r_2 = rank(&H(\theta_T)^\top I(\theta_T)^{-1}H(\theta_T)(H(\theta_T)^\top I^\omega(\theta_T)^{-1}H(\theta_T))^{-1} \\ &H(\theta_T)^\top I(\theta_T)^{-1}H(\theta_T)) \end{aligned}$$

and $b_1, \ldots, b_{r_2}$ are the non zero eigenvalues of $(H(\theta_T)^\top I^\omega(\theta_T)^{-1}H(\theta_T))^{-1}H(\theta_T)^\top I(\theta_T)^{-1}H(\theta_T)$.

iii)

$$R_n^{\omega,c} \xrightarrow{\mathcal{D}} \sum_{i=1}^{r_3} c_i V_i^2,$$

where $r_3 = rank(B(\theta_T)I(\theta_T)^{-1}B(\theta_T)I^\omega(\theta_T)^{-1}B(\theta_T)I(\theta_T)^{-1}B(\theta_T))$ and $c_1, \ldots, c_{r_3}$ are the non zero eigenvalues of $I^\omega(\theta_T)^{-1}B(\theta_T)I(\theta_T)^{-1}B(\theta_T)$.

iv)

$$M_n^{\omega,c} \xrightarrow{\mathcal{D}} \sum_{i=1}^{r_4} d_i V_i^2,$$

where $r_4 = rank(\Gamma^\omega(\theta_T))$ and $d_1, \ldots, d_{r_4}$ are the non zero eigenvalues of $\Gamma^{-1}(\theta_T)\Gamma^\omega(\theta_T)$.

The following theorem deals with the asymptotic behaviour of the local $\phi-$divergence test statistic under $\mathcal{H}_1^\omega$.

Theorem 3.4. *Assume that there exists a unique $\theta^* \in \Theta$ that maximizes $E(\log g_\theta(Z, \Delta))$ under the constraint $h(\theta) = 0$, then under $\mathcal{H}_1^\omega$ and the assumptions* **H1-H7**, *the test statistic $T_{\phi,n}^\omega(\widehat{\theta}_n, \widetilde{\theta}_n)$ tends in probability to infinity.*

From this theorem, we deduce that the power of the local $\phi-$divergence test tends to 1 as n tends to infinity, i.e., it is a consistent test.

3.5. **Local tests of homogeneity.** In this paragraph, we will study local tests of homogeneity. In particular, we will construct local $\phi-$divergence and Wald test statistics and we will determine their asymptotic distributions under the null hypothesis. As in the previous paragraphs, we will also determine, for the local $\phi-$divergence statistics, the asymptotic distribution under the alternative hypothesis, which allows us to give an approximation to the power function.

Let $(Z_i, \Delta_i)_{1\leq i\leq n}$ be an observed sample associated to the variable of interest X, with probability density function $f_{\theta^{(1)}}$ and let $(\widetilde{Z}_i, \widetilde{\Delta}_i)_{1\leq i\leq m}$ be an observed sample associated to the variable of interest $\widetilde{X}$, with probability density function $f_{\theta^{(2)}}$. We assume that the two samples are independent and that the sample sizes n and m are asymptotically linked by the following relation

$$\lim_{\substack{n\to\infty \\ m\to\infty}} \frac{m}{m+n} = \rho \in (0,1).$$

We want to test the local null hypothesis $\mathcal{H}_0^\omega : \theta^{(1)} = \theta^{(2)}$ against the alternative $\mathcal{H}_1^\omega : \theta^{(1)} \neq \theta^{(2)}$. For that, we use the local $\phi-$divergence and Wald test statistics. The first one is defined by

$$T^\omega_{\phi,n,m}(\widehat{\theta}^{(1)}_n, \widehat{\theta}^{(2)}_m) = \frac{2nm D^\omega_\phi(\widehat{\theta}^{(1)}_n, \widehat{\theta}^{(2)}_m)}{(m+n)\phi''(1)},$$

where $\widehat{\theta}^{(1)}_n$ and $\widehat{\theta}^{(2)}_m$ are the MPLE's of $\theta^{(1)}$ and $\theta^{(2)}$ on the basis of the samples $(Z_i, \Delta_i)_{1\leq i\leq n}$ and $(\widetilde{Z}_i, \widetilde{\Delta}_i)_{1\leq i\leq m}$, respectively.

Moreover, the Wald statistic is given by

$$W^\omega_{n,m} = nm \left(\widehat{\theta}^{(1)}_n - \widehat{\theta}^{(2)}_m\right)^\top \left[m I^\omega(\widehat{\theta}^{(1)}_n)^{-1} + n I^\omega(\widehat{\theta}^{(2)}_m)^{-1}\right]^{-1} \left(\widehat{\theta}^{(1)}_n - \widehat{\theta}^{(2)}_m\right).$$

The next theorem gives the asymptotic distributions of these statistics under $\mathcal{H}_0^\omega$.

Theorem 3.5. *Assume that the assumptions* **H1-H5** *hold, so under $\mathcal{H}_0^\omega$, the statistics $T^\omega_{\phi,n,m}(\widehat{\theta}^{(1)}_n, \widehat{\theta}^{(2)}_m)$ and $W^\omega_{n,m}$ converge in distribution, as n and m tend to infinity, to $\sum_{i=1}^r a_i V_i^2$, where $r = rank(I(\theta^{(1)})^{-1} I^\omega(\theta^{(1)}) I(\theta^{(1)})^{-1})$ and $a_1, \ldots, a_r$ are the non zero eigenvalues of the matrix $I^\omega(\theta^{(1)}) I(\theta^{(1)})^{-1}$.*

In order to get an approximation of the power of the local $\phi-$divergence test, we will give the asymptotic distribution of $T^\omega_{\phi,n,m}(\widehat{\theta}^{(1)}_n, \widehat{\theta}^{(2)}_m)$ under the alternative hypothesis $\mathcal{H}_1^\omega$.

Theorem 3.6. *Assume that the assumptions* **H1-H5** *hold and that the function ϕ also satisfies the following assumption.*

H8: *For all $1 \leq i \leq d$, we have*

$$\begin{aligned}
&\frac{\partial}{\partial\theta^{(1)}_i} \int_F \int_E g_\omega(z,\delta) g_{\theta^{(2)}}(z,\delta) \phi\left(\frac{g_{\theta^{(1)}}(z,\delta)}{g_{\theta^{(2)}}(z,\delta)}\right) dm(z) d\mu(\delta) \\
&= \int_F \int_E g_\omega(z,\delta) \frac{\partial g_{\theta^{(1)}}}{\partial\theta^{(1)}_i}(z,\delta) \phi'\left(\frac{g_{\theta^{(1)}}(z,\delta)}{g_{\theta^{(2)}}(z,\delta)}\right) dm(z) d\mu(\delta)
\end{aligned}$$

and

$$\frac{\partial}{\partial\theta_i^{(2)}}\int_F\int_E g_\omega(z,\delta)g_{\theta^{(2)}}(z,\delta)\phi\left(\frac{g_{\theta^{(1)}}(z,\delta)}{g_{\theta^{(2)}}(z,\delta)}\right)dm(z)d\mu(\delta)$$
$$=\int_F\int_E g_\omega(z,\delta)\left[\frac{\partial g_{\theta^{(2)}}}{\partial\theta_i^{(2)}}(z,\delta)\phi\left(\frac{g_{\theta^{(1)}}(z,\delta)}{g_{\theta^{(2)}}(z,\delta)}\right)\right.$$
$$\left.-\frac{\partial g_{\theta^{(2)}}}{\partial\theta_i^{(2)}}(z,\delta)\frac{g_{\theta^{(1)}}(z,\delta)}{g_{\theta^{(2)}}(z,\delta)}\phi'\left(\frac{g_{\theta^{(1)}}(z,\delta)}{g_{\theta^{(2)}}(z,\delta)}\right)\right]dm(z)d\mu(\delta).$$

So, under $\mathcal{H}_1^\omega$, we have

$$\sqrt{\frac{nm}{m+n}}\left(D_\phi^\omega(\widehat{\theta}_n^{(1)},\widehat{\theta}_m^{(2)})-D_\phi^\omega(\theta^{(1)},\theta^{(2)})\right)\underset{\substack{n\to\infty\\ m\to\infty}}{\xrightarrow{\mathcal{D}}}\mathcal{N}\left(0,\sigma_\phi^2(\theta^{(1)},\theta^{(2)})\right),$$

where

$$\sigma_\phi^2(\theta^{(1)},\theta^{(2)})=\rho T_1^\top I(\theta^{(1)})^{-1}T_1+(1-\rho)T_2^\top I(\theta^{(2)})^{-1}T_2$$

with

$$T_1=(t_{11},\dots,t_{1d})^\top,\ t_{1i}=\int_F\int_E g_\omega(z,\delta)\frac{\partial g_{\theta^{(1)}}}{\partial\theta_i^{(1)}}(z,\delta)\phi'\left(\frac{g_{\theta^{(1)}}(z,\delta)}{g_{\theta^{(2)}}(z,\delta)}\right)dm(z)d\mu(\delta)$$

and $T_2=(t_{21},\dots,t_{2d})^\top$,

$$t_{2i}=\int_F\int_E g_\omega(z,\delta)\left[\frac{\partial g_{\theta^{(2)}}}{\partial\theta_i^{(2)}}(z,\delta)\phi\left(\frac{g_{\theta^{(1)}}(z,\delta)}{g_{\theta^{(2)}}(z,\delta)}\right)\right.$$
$$\left.-\frac{\partial g_{\theta^{(2)}}}{\partial\theta_i^{(2)}}(z,\delta)\frac{g_{\theta^{(1)}}(z,\delta)}{g_{\theta^{(2)}}(z,\delta)}\phi'\left(\frac{g_{\theta^{(1)}}(z,\delta)}{g_{\theta^{(2)}}(z,\delta)}\right)\right]dm(z)d\mu(\delta).$$

The critical region of the local $\phi-$divergence test at level $\alpha\in(0,1)$ is given by $CR=\left\{T_{\phi,n,m}^\omega(\widehat{\theta}_n^{(1)},\widehat{\theta}_m^{(2)})>q_{1-\alpha}\right\}$, where $q_{1-\alpha}$ is the $(1-\alpha)-$quantile of the asymptotic distribution of $T_{\phi,n,m}^\omega(\widehat{\theta}_n^{(1)},\widehat{\theta}_m^{(2)})$ under $\mathcal{H}_0^\omega$. Proceeding as in the one sample case, we can approximate the power function as follows

$$\pi\approx 1-F_{\mathcal{N}}\left[\frac{1}{\sigma_\phi(\theta^{(1)},\theta^{(2)})}\sqrt{\frac{nm}{m+n}}\left(\frac{m+n}{nm}\frac{\phi''(1)}{2}q_{1-\alpha}-D_\phi^\omega(\theta^{(1)},\theta^{(2)})\right)\right].$$

where $F_{\mathcal{N}}$ is the cumulative distribution function of the standard normal distribution.

4. Real data application

[10, 11] reported a study on the cosmetic results of breast cancer patients, treated either by radiotherapy only or by radiotherapy and chemotherapy. During the period from 1976 to 1980, the patients were followed in order to record the time to the cosmetic retraction of the breast. At the begginig, the patients were observed every 4 to 6 months, but, as their recovery progressed, the checking times became more distant. Therefore, the time of breast retraction is case two interval censored. This dataset has also been used by [12] and [13, 14] who suggested the log-normal distribution to fit the data. For our part, we consider the sample of patients treated by radiotherapy and chemotherapy (composed of 48 patients) and we compare its distribution with the log-normal distribution with parameter $\theta_0=(m_0,\sigma_0^2)=(3,0.7^2)$. The graphs of the density of this distribution and that of the kernel estimated density of the data are given in the left panel of Figure 1. We use the gamma kernel to calculate the estimated density of the data. Overall, the two

graphs are distant, but they are very close in a certain zone at the right. To highlight this zone, we add in the right panel of Figure 1, the graph of the truncated normal kernel with parameter $\omega = (\mu_\omega, \sigma_\omega^2) = (25, 0.3^2)$. This latter is given by

$$k_\omega(x) = \frac{1}{\sigma_\omega \sqrt{2\pi} F_{\mathcal{N}}\left(\frac{\mu_\omega}{\sigma_\omega}\right)} \exp\left\{-\frac{(x-\mu_\omega)^2}{2\sigma_\omega^2}\right\} 1_{\{x>0\}}.$$

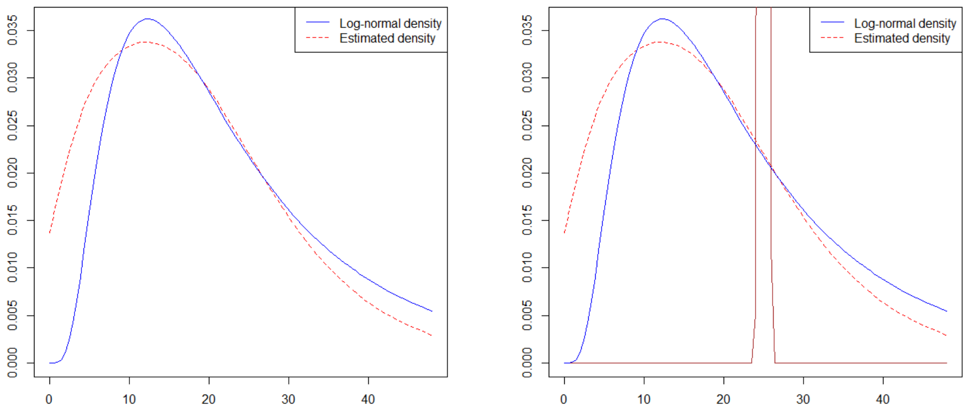

FIGURE 1. Graphs of the log-normal density with parameter $\theta_0 = (3, 0.7^2)$ and the kernel estimated density of the data.

To confirm our observation, we use the global and local $\phi-$divergence, Wald and Rao tests on this set of data, at the significance level $\alpha = 0.05$. The divergences we use are the Kullback-Leibler (KL), modified Kullback-Leibler (KL_m) and the $\lambda-$power divergences introduced by [15] (for different values of λ). They correspond respectively to the functions: $\phi_{KL}(x) = x\log(x) - x + 1$, $\phi_{KL_m}(x) = -\log(x) + x - 1$ and $\phi_\lambda(x) = \dfrac{x^{\lambda+1} - x - \lambda(x-1)}{\lambda(\lambda+1)}$ (for $\lambda \neq 0$ and $\lambda \neq -1$). For $\lambda = -0.5$ (resp. $\lambda = 1$), we obtain the Hellinger (resp. the χ^2) divergence. In the case of global tests, the critical value q is the $(1-\alpha)-$quantile of the χ_2^2 distribution and in the case of local tests, it is calculated from Theorem 3.1. The kernel we use for local tests is $g_\omega(r, l) = k_\omega(r)k_\omega(l)$. Our obtained results are given in Tables 1 and 2.

The test	The test statistic	q	Decision
KL divergence	10748.1		Reject $\mathcal{H}_0$
Modified KL divergence	30891.67		Reject $\mathcal{H}_0$
Hellinger divergence ($\lambda = -0.5$)	14660.05		Reject $\mathcal{H}_0$
Power divergence with $\lambda = 0.5$	9189.61	5.991465	Reject $\mathcal{H}_0$
χ^2 divergence ($\lambda = 1$)	8586.346		Reject $\mathcal{H}_0$
Wald	126.7467		Reject $\mathcal{H}_0$
Rao	2.850749		Accept $\mathcal{H}_0$

TABLE 1. The obtained results for the global tests.

The test	The test statistic	q	Decision
KL divergence	0.116991	3.105955	Accept $\mathcal{H}_0$
Modified KL divergence	0.152990		Accept $\mathcal{H}_0$
Hellinger divergence ($\lambda = -0.5$)	0.131920		Accept $\mathcal{H}_0$
Power divergence with $\lambda = 0.5$	0.106491		Accept $\mathcal{H}_0$
χ^2 divergence ($\lambda = 1$)	0.099273		Accept $\mathcal{H}_0$
Wald	1.234309		Accept $\mathcal{H}_0$
Rao	1434.473	2890.549	Accept $\mathcal{H}_0$

Table 2. The obtained results for the local tests.

Except the Rao test, all tests reject the global null hypothesis. Moreover, all tests accept the local null hypothesis, which confirms our observation of Figure 1.

5. Conclusions

Under a general framework of incomplete data, we have introduced local tests in parametric models for simple and composite null hypotheses. We have also introduced local tests of homogeneity. These tests are based on $\phi-$divergences, Wald and Rao statistics. In the future, it would be interesting to look at local model selection (see [9]) for incomplete data. It would also be interesting to study local nonparametric procedures to test the goodness-of-fit, the homogeneity and the independence.

References

[1] Csiszár, E., (1963), Eine informationstheoretische ungleichung und ihre anwendung auf den beweis der ergodizität von markoffschen ketten, Magyar Tud. Akad. Mat. Kutató Int. Közl., 8, pp. 85-108.
[2] Broniatowski, M. and Keziou, A., (2009), Parametric estimation and tests through divergences and the duality technique, Journal of Multivariate Analysis, 100(1), pp. 16-36.
[3] Broniatowski, M. and Keziou, A., (2012), Divergences and duality for estimation and test under moment condition models, Journal of Statistical Planning and Inference, 142(9), pp. 2554-2573.
[4] Bouzebda, S. and Keziou, A., (2010), New estimates and tests of independence in semiparametric copula models, Kybernetika, 46(1), pp. 178-201.
[5] Boukeloua, M., (2021), Study of semiparametric copula models via divergences with bivariate censored data, Communications in Statistics-Theory and Methods, 50(23), pp. 5429-5452.
[6] Salicrú, M., Morales, D., Menéndez, M. L. and Pardo, L., (1994), On the applications of divergence type measures in testing statistical hypotheses, Journal of Multivariate Analysis, 51(2), pp. 372-391.
[7] Avlogiaris, G., Micheas, A. and Zografos, K., (2016), On testing local hypotheses via local divergence, Statistical Methodology, 31, pp. 20-42.
[8] Avlogiaris, G., Micheas, A. and Zografos, K., (2016), On local divergences between two probability measures, Metrika, 79(3), pp. 303-333.
[9] Avlogiaris, G., Micheas, A. C. and Zografos, K., (2019), A criterion for local model selection, Sankhyā A: The Indian Journal of Statistics, 81(2), pp. 406-444.
[10] Beadle, G. F., Silver, B., Botnick, L., Hellman, S. and Harris, J. R., (1984), Cosmetic results following primary radiation therapy for early breast cancer, Cancer, 54, pp. 2911-2918.
[11] Beadle, G. F., Come, S., Henderson, I. C., Silver, B., Hellman, S. and Harris, J. R., (1984), The effect of adjuvant chemotherapy on the cosmic results after primary radiation treatment for early stage breast cancer, International Journal of Radiation Oncology, Biology and Physics, 10, pp. 2131-2137.
[12] Finkelstein, D. M. and Wolfe, R. A., (1985), A semiparametric model for regression analysis of interval-censored failure time data, Biometrics, 41, pp. 933-945.
[13] Lindsey, J. C. and Ryan, L. M., (1998), Tutorial in biostatistics: Methods for interval-censored data, Statistics in Medicine, 17(2), pp. 219-238.
[14] Lindsey, J. C. and Ryan, L. M., (1999), Erratum to "Tutorial in biostatistics: Methods for interval-censored data"[Statistics in Medicine 17 (1998) 219-238], Statistics in Medicine, 18(7), pp. 890.

[15] Cressie, N. and Read, T. R., (1984), Multinomial goodness-of-fit tests, Journal of the Royal Statistical Society: Series B (Methodological), 46(3), pp. 440-464.
[16] Dik, J. J. and de Gunst, M. C. M., (1985), The distribution of general quadratic forms in normal variables, Statistica Neerlandica, 39(1), pp. 14-26.
[17] Pardo, L., (2006), Statistical inference based on divergence mesures, Chapman & Hall/CRC, Madrid.
[18] Sen, P. K. and Singer J. M., (1993), Large sample methods in statistics: an introduction with applications, Chapman & Hall/CRC, New York.

Appendix A. Proofs

Proof of Theorem 3.1. - The result for $T^{\omega}_{\phi,n}(\widehat{\theta}_n, \theta_0)$ can be proved following the same steps of the proof of Theorem 2.1 of [7]. The difference lies in the formulas of $D^{\omega}_{\phi}(\theta, \theta_0)$, $I(\theta)$ and $I^{\omega}(\theta)$, where the functions f_θ and f_ω in [7] are respectively replaced by g_θ and g_ω. In particular, using a Taylor expansion and the fact that $\nabla_\theta D^{\omega}_{\phi}(\theta, \theta_0)|_{\theta=\theta_0} = 0$ and $\nabla^{\top}_{\theta} \nabla_\theta D^{\omega}_{\phi}(\theta, \theta_0)|_{\theta=\theta_0} = \phi"(1) I^{\omega}(\theta_0)$, we get

$$D^{\omega}_{\phi}\left(\widehat{\theta}_n, \theta_0\right) = \frac{1}{2}\left(\widehat{\theta}_n - \theta_0\right)^{\top} \phi"(1) I^{\omega}(\theta_0)\left(\widehat{\theta}_n - \theta_0\right) + o_p\left(n^{-1}\right).$$

So that

$$T^{\omega}_{\phi,n}(\widehat{\theta}_n, \theta_0) = n\left(\widehat{\theta}_n - \theta_0\right)^{\top} I^{\omega}(\theta_0)\left(\widehat{\theta}_n - \theta_0\right) + o_p(1)$$

and the claimed result follows from the fact that

$$\sqrt{n}\left(\widehat{\theta}_n - \theta_0\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, I(\theta_0)^{-1}\right), \tag{6}$$

thanks to Corollary 2.1 of [16].
- The results for W^{ω}_n and R^{ω}_n follow by the same steps of the proof of Theorem 2.2 of [7]. Here too, the difference lies in the formulas of $I^{\omega}(\theta)$ and $U_n(\theta)$, where the functions f_θ and f_ω in [7] are respectively replaced by g_θ and g_ω. In particular, the convergence of W_n follows from (6) and the fact that $I^{\omega}(\widehat{\theta}_n) \xrightarrow{P} I^{\omega}(\theta_0)$, thanks to Corollary 2.1 of [16]. Moreover, the convergence of R^{ω}_n follows, once again, from this corollary and the fact that

$$\frac{1}{\sqrt{n}} U_n(\theta_0) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, I(\theta_0)\right).$$

□

Proof of Theorem 3.2. The proof is similar to that of Theorem 9.2 of [17]. □

Proof of Theorem 3.3. - The proof of i) is similar to that of Theorem 3.1 of [7].
- The proof of ii) is similar to that of Theorem 3.2 of [7].
- Proof of iii):
Thanks to equations (5.6.20) page 242 and (5.6.2) page 237 of [18], we have

$$\begin{aligned}\frac{1}{\sqrt{n}} U_n(\widetilde{\theta}_n) &= \frac{1}{\sqrt{n}} U_n(\theta_T) - I(\theta_T)\sqrt{n}\left(\widetilde{\theta}_n - \theta_T\right) + o_P(1) \\ &= I(\theta_T)\sqrt{n}\left(\widehat{\theta}_n - \theta_T\right) - I(\theta_T)\sqrt{n}\left(\widetilde{\theta}_n - \theta_T\right) + o_P(1).\end{aligned}$$

So, lemma 3.1 of [7] allows to write

$$\begin{aligned}\frac{1}{\sqrt{n}} U_n(\widetilde{\theta}_n) &= I(\theta_T)\sqrt{n}\left(\widehat{\theta}_n - \theta_T\right) - I(\theta_T)(I_d - I(\theta_T)^{-1}B(\theta_T))\sqrt{n}\left(\widehat{\theta}_n - \theta_T\right) + o_P(1) \\ &= B(\theta_T)\sqrt{n}\left(\widehat{\theta}_n - \theta_T\right) + o_P(1).\end{aligned}$$

Therefore

$$\frac{1}{\sqrt{n}}U_n(\widetilde{\theta}_n) \xrightarrow{\mathcal{D}} \mathcal{N}(0, B(\theta_T)I(\theta_T)^{-1}B(\theta_T)), \text{ as } n \to \infty$$

and since $I^\omega(\widetilde{\theta}_n)^{-1} \xrightarrow{P} I^\omega(\theta_T)^{-1}$, we deduce by Corollary 2.1 of [16] that

$$R_n^{\omega,c} \xrightarrow{\mathcal{D}} \sum_{i=1}^{r_3} c_i V_i^2,$$

where $r_3 = rank(B(\theta_T)I(\theta_T)^{-1}B(\theta_T)I^\omega(\theta_T)^{-1}B(\theta_T)I(\theta_T)^{-1}B(\theta_T))$ and $c_1, \ldots, c_{r_3}$ are the non zero eigenvalues of $I^\omega(\theta_T)^{-1}B(\theta_T)I(\theta_T)^{-1}B(\theta_T)$.

- Proof of iv):

Equations (5.6.23) page 243 and (5.6.2) page 237 of [18] allow to write

$$\begin{aligned}
\frac{1}{\sqrt{n}}\widetilde{\lambda}_n &= -(H(\theta_T)^\top I(\theta_T)^{-1}H(\theta_T))^{-1}H(\theta_T)^\top I(\theta_T)^{-1}\left(\frac{1}{\sqrt{n}}U_n(\theta_T)\right) + o_P(1) \\
&= -(H(\theta_T)^\top I(\theta_T)^{-1}H(\theta_T))^{-1}H(\theta_T)^\top \sqrt{n}\left(\widehat{\theta}_n - \theta_T\right) + o_P(1) \\
&= -\Gamma(\theta_T)^{-1}H(\theta_T)^\top \sqrt{n}\left(\widehat{\theta}_n - \theta_T\right) + o_P(1). \qquad (7)
\end{aligned}$$

Otherwise, the Taylor Young formula permits to write

$$\begin{aligned}
\sqrt{n}h(\widehat{\theta}_n) &= \sqrt{n}h(\theta_T) + \sqrt{n}H(\theta_T)^\top\left(\widehat{\theta}_n - \theta_T\right) + o_P(\sqrt{n}(\widehat{\theta}_n - \theta_T)) \\
&= \sqrt{n}H(\theta_T)^\top\left(\widehat{\theta}_n - \theta_T\right) + o_P(1),
\end{aligned}$$

so $\sqrt{n}H(\theta_T)^\top\left(\widehat{\theta}_n - \theta_T\right) = \sqrt{n}h(\widehat{\theta}_n) + o_P(1)$. Combining this with (7), we get

$$\frac{1}{\sqrt{n}}\widetilde{\lambda}_n = -\sqrt{n}\Gamma(\theta_T)^{-1}h(\widehat{\theta}_n) + o_P(1)$$

and

$$\begin{aligned}
M_n^{\omega,c} &= \frac{1}{n}\widetilde{\lambda}_n^\top \Gamma^\omega(\widetilde{\theta}_n)\widetilde{\lambda}_n \\
&= nh(\widehat{\theta}_n)^\top \Gamma(\theta_T)^{-1}\Gamma^\omega(\widetilde{\theta}_n)\Gamma(\theta_T)^{-1}h(\widehat{\theta}_n) + o_P(1) \\
&= nh(\widehat{\theta}_n)^\top \Gamma(\theta_T)^{-1}\Gamma^\omega(\theta_T)\Gamma(\theta_T)^{-1}h(\widehat{\theta}_n) + o_P(1)
\end{aligned}$$

by the continuity of $\Gamma^\omega(\theta)$ in θ.

Moreover, proceeding as in the proof of Theorem 5.4.1 of [18], we get

$$\sqrt{n}\left(h(\widehat{\theta}_n) - h(\theta_T)\right) = \sqrt{n}h(\widehat{\theta}_n) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \Gamma(\theta_T)\right).$$

So, the claimed result follows from Corollary 2.1 of [16].

□

Proof of Theorem 3.4. Proceeding as in [7] (proof of Theorem 3.1), we get

$$T_{\phi,n}^\omega(\widehat{\theta}_n, \widetilde{\theta}_n) = n\left[\left(\widehat{\theta}_n - \widetilde{\theta}_n\right)^\top I^\omega(\widetilde{\theta}_n)\left(\widehat{\theta}_n - \widetilde{\theta}_n\right) + o_p(\|\widehat{\theta}_n - \widetilde{\theta}_n\|^2)\right].$$

So, the claimed result follows from the fact that $\widehat{\theta}_n \xrightarrow{P} \theta_T$ and $\widetilde{\theta}_n \xrightarrow{P} \theta^* \neq \theta_T$. □

Proof of Theorem 3.5. - To obtain the asymptotic distribution of $T_{\phi,n,m}^\omega$, one can proceed as in Theorem 4.1 of [7].

- For $W^{\omega}_{n,m}$, we have under $\mathcal{H}^{\omega}_0$

$$W^{\omega}_{n,m} = mn\left(\widehat{\theta}^{(1)}_n - \widehat{\theta}^{(2)}_m\right)^{\top}\left[m(I^{\omega}(\theta^{(1)}) + o_P(1))^{-1} + n(I^{\omega}(\theta^{(1)}) + o_P(1))^{-1}\right]^{-1}$$
$$\left(\widehat{\theta}^{(1)}_n - \widehat{\theta}^{(2)}_m\right)$$
$$= \frac{mn}{m+n}\left(\widehat{\theta}^{(1)}_n - \widehat{\theta}^{(2)}_m\right)^{\top}(I^{\omega}(\theta^{(1)}) + o_P(1))\left(\widehat{\theta}^{(1)}_n - \widehat{\theta}^{(2)}_m\right).$$

In view of [17] page 443, we have

$$\sqrt{\frac{mn}{m+n}}\left(\widehat{\theta}^{(1)}_n - \widehat{\theta}^{(2)}_m\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, I(\theta^{(1)})^{-1}\right), \tag{8}$$

which implies that

$$\sqrt{\frac{mn}{m+n}}\left(\widehat{\theta}^{(1)}_n - \widehat{\theta}^{(2)}_m\right) = O_P(1).$$

So

$$W^{\omega}_{n,m} = \frac{mn}{m+n}\left(\widehat{\theta}^{(1)}_n - \widehat{\theta}^{(2)}_m\right)^{\top} I^{\omega}(\theta^{(1)})\left(\widehat{\theta}^{(1)}_n - \widehat{\theta}^{(2)}_m\right) + o_P(1)$$

and the claimed result follows from (8) thanks to Corollary 2.1 of [16].

□

Proof of Theorem 3.6. The proof follows by the same arguments used in [17], pages 441-442. □

Mohamed Boukeloua is currently an associate professor of mathematics in the National Polytechnic Institute of Constantine, Algeria. He is a member of Laboratory of Process Engineering for Sustainable Development and Health Products in the same institute. He is also a member of Laboratory of Biostatistics, Bioinformatics and Mathematical Methodology Applied on Health Sciences. He holds a PhD in Mathematical Statistics. His research interests are Parametric and Non parametric Inference, Censored data, $\phi-$divergences and their applications, Copula models, Kernel estimation, Non parametric regression, Dependent data, Bayesian inference.