Robust Estimators and Test Statistics for One-Shot Device Testing Under the Exponential Distribution

Narayanaswamy Balakrishnan, Elena Castilla, Nirian Martín, and Leandro Pardo

***Abstract*— This paper develops a new family of estimators, the minimum density power divergence estimators (MDPDEs), for the parameters of the one-shot device model as well as a new family of test statistics, Z-type test statistics based on MDPDEs, for testing the corresponding model parameters. The family of MDPDEs contains as a particular case the maximum likelihood estimator (MLE) considered in Balakrishnan and Ling (2012). Through a simulation study, it is shown that some MDPDEs have a better behavior than the MLE in terms of robustness. At the same time, it can be seen that some Z-type tests based on MDPDEs have a better behavior than the classical Z-test statistic in terms of robustness, as well.**

***Index Terms*— Exponential distribution, minimum density power divergence estimator, one-shot devices, robustness, Z-type tests.**

I. Introduction

THE reliability of a product, system, weapon, or piece of equipment can be defined as the ability of the device to perform as designed for, or, more simply, as the probability that the device does not fail when used. Engineers assess reliability by repeatedly testing the device and observing its failure rate. Certain products, called "one-shot" devices, make this approach challenging. One-shot devices can be used only once and after use the device is either destroyed or must be rebuilt. Consequently, one can only know whether the failure time is either before or after the test time. The outcomes from each of the devices are therefore binary, either right-censored (failure) or left-censored (success). Some examples of one-shot devices are nuclear weapons, space shuttles, automobile air bags, fuel injectors, disposable napkins, heat detectors, and fuses. In survival analysis, these data are called "current status data". For instance, in animal carcinogenicity experiments, one observes whether a tumor occurs at the examination time for each subject.

Due to the advances in manufacturing design and technology, products have now become highly reliable with long lifetimes. This fact would pose a problem in the analysis of data if only few or no failures are observed. For this reason, accelerated life tests are often used by adjusting a controllable factor such as temperature in order to induce more failures in the experiment. Moreover, accelerated life testing would shorten the experimental time and also help to reduce the experimental cost. In this paper, we shall assume that the failure times of devices follow an exponential distribution. In this context, [1] developed the EM algorithm for finding the maximum likelihood estimators (MLEs) of the model parameters. Reference [2] studied a Bayesian approach for one-shot device testing along with an accelerating factor, in which the failure times of devices is assumed to follow once again an exponential distribution. Reference [3] presented two approaches based on the likelihood ratio statistics and the posterior Bayes factor for comparing several exponential accelerated life models. Reference [4] made a comparison of several goodness-of-fit tests for one-shot device testing data.

In Section II, we present a description of the one-shot device model as well as the MLEs of the model parameters. Section III develops the minimum density power divergence estimator (MDPDEs) as a natural extension of the MLE, as well as its asymptotic distribution. In Section IV, Z-type test statistics are introduced for testing some hypotheses about the parameters of the one-shot device model. The influence functions of these Z-type tests are studied in Section V. Some numerical examples are presented in Section VI, with one of them relating to a reliability situation and the other two are from real applications to tumorigenicity experiments. In Section VII, an extensive simulation study is presented in order to empirically illustrate the robustness of the MDPDEs, as well as the Z-type test introduced earlier. A data-driven choice procedure of the optimal tuning parameter given a data set is provided in Section VIII. Finally, some concluding remarks are made in Section IX.

Manuscript received March 16, 2017; revised February 8, 2019; accepted February 22, 2019. Date of publication March 5, 2019; date of current version April 19, 2019. This work was supported in part by the Ministerio de Economia y Competitividad and Ministerio de Educacion, Cultura y Deporte, Spain, under Grant MTM2015-67057-P and Grant FPU16/03104.

N. Balakrishnan is with the Department of Mathematics and Statistics, McMaster University, Hamilton, ON L8S4K1, Canada (e-mail: bala@mcmaster.ca).

E. Castilla and L. Pardo are with the Department of Statistics and O.R., Complutense University of Madrid, 28040 Madrid, Spain (e-mail: elecasti@ucm.es; lpardo@mat.ucm.es).

N. Martín is with the Department of Financial and Actuarial Economics and Statistics, Complutense University of Madrid, 28040 Madrid, Spain (e-mail: nirian@estad.ucm.es).

Communicated by K.-R. Müller, Associate Editor for Machine Learning.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIT.2019.2903244

TABLE I
FAILURES IN THE-SHOT DEVICE TESTING EXPERIMENT OF BALAKRISHNAN AND LING (2012)

	$t_1 = 10$	$t_2 = 20$	$t_3 = 30$
$w_1 = 35$	3	3	7
$w_2 = 45$	1	5	7
$w_3 = 55$	6	7	9

II. MODEL FORMULATION AND MAXIMUM LIKELIHOOD ESTIMATOR

Consider a reliability testing experiment in which at each time, $t_j,\ j = 1, 2, \ldots, J$, K devices are placed under temperatures w_i, $i = 1, \ldots, I$. Thus, IJK devices are tested in total. It is worth noting that a successful detonation occurs if the lifetime is beyond the inspection time, whereas the lifetime will be before the inspection time if the detonation is a failure. For each temperature w_i and at each inspection time t_j, the number of failures, n_{ij}, is then recorded.

In [1], an example is presented, in which 30 devices were tested at temperatures $w_i \in \{35, 45, 55\}$, each with 10 units being detonated at times $t_j \in \{10, 20, 30\}$, respectively. In this example, we have $I = 3$, $J = 3$ and $K = 10$. The number of failures observed is summarized in the 3×3 table given in Table I. In this one-shot device testing experiment, there were in all 48 failures out a total of 90 tested devices.

We shall assume here, as done by [1], that the true lifetimes T_{ijk}, for $i = 1, 2, \ldots, I$, $j = 1, 2, \ldots, J$, $k = 1, \ldots, K$, are independent and identically distributed exponential random variables with probability density function

$$f(t|\lambda) = \lambda \exp(-\lambda t),$$

where $\lambda > 0$ is the unknown failure rate. In practice, we consider inspection times $t_j,\ j = 1, \ldots, J$, rather than $t > 0$, and we relate the parameter λ to an accelerating factor of temperature $w_i > 0$ through a log-linear link function of the form

$$\lambda_{w_i}(\boldsymbol{\alpha}) = \alpha_0 \exp\{\alpha_1 w_i\},$$

where $\alpha_0 > 0$ and $\alpha_1 \in \mathbb{R}$ are unknown parameters. Therefore, the corresponding distribution function is

$$\begin{aligned} F(t_j|\lambda_{w_i}(\boldsymbol{\alpha})) &= 1 - \exp\left\{-\lambda_{w_i}(\boldsymbol{\alpha}) t_j\right\} \\ &= 1 - \exp\left\{-\alpha_0 \exp\{\alpha_1 w_i\} t_j\right\} \end{aligned} \tag{1}$$

and the density function is

$$f(t_j|\lambda_{w_i}(\boldsymbol{\alpha})) = \alpha_0 \exp\{\alpha_1 w_i\} \exp\left\{-\alpha_0 \exp\{\alpha_1 w_i\} t_j\right\}. \tag{2}$$

The data are completely described on K devices, through the contingency table of failures $\boldsymbol{n} = (n_{11}, \ldots, n_{1J}, \ldots, n_{I1}, \ldots, n_{IJ})^T$, collected at the temperatures $\boldsymbol{w} = (w_1, \ldots, w_I)^T$ and the inspection times $\boldsymbol{t} = (t_1, \ldots, t_J)^T$.

We shall consider the theoretical probability vector of successes and failures, $\boldsymbol{p}(\boldsymbol{\alpha})$, defined by

$$\begin{aligned} \boldsymbol{p}(\boldsymbol{\alpha}) = \Big(&\tfrac{F(t_1|\lambda_{w_1}(\boldsymbol{\alpha}))}{IJ}, \tfrac{1-F(t_1|\lambda_{w_1}(\boldsymbol{\alpha}))}{IJ}, \\ &\ldots, \tfrac{F(t_J|\lambda_{w_I}(\boldsymbol{\alpha}))}{IJ}, \tfrac{1-F(t_J|\lambda_{w_I}(\boldsymbol{\alpha}))}{IJ} \Big)^T, \end{aligned}$$

as well as the observed probability vector

$$\widehat{\boldsymbol{p}} = \left(\tfrac{n_{11}}{IJK}, \tfrac{K-n_{11}}{IJK}, \ldots, \tfrac{n_{IJ}}{IJK}, \tfrac{K-n_{IJ}}{IJK} \right)^T,$$

both of dimension $2IJ$. Then, the Kullback-Leibler divergence between the probability vectors $\widehat{\boldsymbol{p}}$ and $\boldsymbol{p}(\boldsymbol{\alpha})$ is given by

$$\begin{aligned} d_{KL}(\widehat{\boldsymbol{p}}, \boldsymbol{p}(\boldsymbol{\alpha})) = \frac{1}{IJ} \sum_{i=1}^{I} \sum_{j=1}^{J} \Bigg(&\frac{n_{ij}}{K} \log \frac{n_{ij}}{K F(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))} \\ &+ \frac{K - n_{ij}}{K} \log \frac{K - n_{ij}}{K\left(1 - F(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))\right)} \Bigg). \end{aligned}$$

It is not difficult to establish the following result.

Theorem 1: The likelihood function

$$\begin{aligned} &\mathcal{L}(\boldsymbol{\alpha} | K, \boldsymbol{n}, \boldsymbol{t}, \boldsymbol{w}) \\ &= \prod_{i=1}^{I} \prod_{j=1}^{J} F(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))^{n_{ij}} \left(1 - F(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))\right)^{K-n_{ij}}, \end{aligned}$$

where $F(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))$ is as given in (1), is related to the Kullback-Leibler divergence between the probability vectors $\widehat{\boldsymbol{p}}$ and $\boldsymbol{p}(\boldsymbol{\alpha})$ through

$$d_{KL}(\widehat{\boldsymbol{p}}, \boldsymbol{p}(\boldsymbol{\alpha})) = \frac{1}{IJK} \left(s - \log \mathcal{L}(\boldsymbol{\alpha} | K, \boldsymbol{n}, \boldsymbol{t}, \boldsymbol{w})\right), \tag{3}$$

with s being a constant not dependent on $\boldsymbol{\alpha}$.

Based on Theorem 1, we have the following definition for the MLEs of α_0 and α_1.

Definition 2: We consider the data given by K, $\boldsymbol{n}$, $\boldsymbol{t}$, $\boldsymbol{w}$ for the one-shot device model. Then, the MLE of $\boldsymbol{\alpha} = (\alpha_0, \alpha_1)^T$, $\widehat{\boldsymbol{\alpha}} = (\widehat{\alpha}_0, \widehat{\alpha}_1)^T$, can be defined as

$$\widehat{\boldsymbol{\alpha}} = \arg\min_{\boldsymbol{\alpha} \in \Theta} d_{KL}(\widehat{\boldsymbol{p}}, \boldsymbol{p}(\boldsymbol{\alpha})), \tag{4}$$

where $\Theta = (\mathbb{R}^+, \mathbb{R})^T$.

III. MINIMUM DENSITY POWER DIVERGENCE ESTIMATOR

Based on expression (4), we can think of defining an estimator by minimizing any distance or divergence measure between the probability vectors $\widehat{\boldsymbol{p}}$ and $\boldsymbol{p}(\boldsymbol{\alpha})$. There are many different divergence measures (or distances) known in the literature, see, for instance, [6] and [7], and the natural question is whether all of them are valid to define estimators with good properties. Initially, the answer is yes, but we must think in terms of efficiency as well as robustness of the defined estimators. From an asymptotic point of view, it is well-known that the MLE is a BAN (Best Asymptotically Normal) estimator, but at the same time we know that the MLE has a very poor behavior, in general, with regard to robustness. It is well-known that a gain in robustness leads to a loss in efficiency. Therefore, the distances (divergence measures) that we must use are those which result in estimators having good properties in terms of robustness with low loss of efficiency. The density power divergence measure introduced by [8] has the required properties and has been studied for many different problems until now. For more details, see [9], [10] and the references therein.

Based on [11], the MDPDE of $\boldsymbol{\alpha}$ is first introduced, and later in Result 4 it is shown that this estimator can be considered as a natural extension of (4).

Definition 3: Let y_{ijk}, $i = 1, 2, \ldots, I$, $j = 1, 2, \ldots, J$, $k = 1, \ldots, K$, be a sequence of independent Bernoulli random variables, $y_{ijk} \overset{ind}{\sim} Ber(\pi_{ij}(\boldsymbol{\alpha}))$, such that $\pi_{ij}(\boldsymbol{\alpha}) = F\left(t_j|\lambda_{w_i}(\boldsymbol{\alpha})\right)$ and $n_{ij} = \sum_{k=1}^{K} y_{ijk}$. The MDPDE of $\boldsymbol{\alpha}$, with tuning parameter $\beta \geq 0$, is given by

$$\widehat{\boldsymbol{\alpha}}_\beta = \arg\min_{\boldsymbol{\alpha}\in\Theta} \frac{1}{IJK} \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} V_{ij}\left(y_{ijk}, \beta\right), \tag{5}$$

where

$$\begin{aligned} V_{ij}\left(y_{ijk}, \beta\right) &= \pi_{ij}^{\beta+1}(\boldsymbol{\alpha}) + (1-\pi_{ij}(\boldsymbol{\alpha}))^{\beta+1} \\ &\quad - \frac{1+\beta}{\beta}\left(\pi_{ij}^{y_{ijk}}(\boldsymbol{\alpha})(1-\pi_{ij}(\boldsymbol{\alpha}))^{1-y_{ijk}}\right)^{\beta}. \end{aligned}$$

For more details about the interpretation of Definition 3, see [11, Formula 2.3], in which $\pi_{ij}^{y_{ijk}}(\boldsymbol{\alpha})(1-\pi_{ij}(\boldsymbol{\alpha}))^{1-y_{ijk}}$ plays the role of the density in our context. Notice that the expression to be minimized in (5) can be simplified as

$$\begin{aligned} &\frac{1}{IJK}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}\left\{\pi_{ij}^{\beta+1}(\boldsymbol{\alpha}) + (1-\pi_{ij}(\boldsymbol{\alpha}))^{\beta+1}\right. \\ &\left.-\frac{1+\beta}{\beta}\left(\pi_{ij}^{y_{ijk}}(\boldsymbol{\alpha})(1-\pi_{ij}(\boldsymbol{\alpha}))^{1-y_{ijk}}\right)^{\beta}\right\} \\ &= \frac{1}{IJ}\sum_{i=1}^{I}\sum_{j=1}^{J}\left\{\pi_{ij}^{\beta+1}(\boldsymbol{\alpha}) + (1-\pi_{ij}(\boldsymbol{\alpha}))^{\beta+1}\right. \\ &\left.-\frac{1+\beta}{\beta}\frac{n_{ij}}{K}\pi_{ij}^{\beta}(\boldsymbol{\alpha}) - \frac{1+\beta}{\beta}\frac{K-n_{ij}}{K}(1-\pi_{ij}(\boldsymbol{\alpha}))^{\beta}\right\}. \end{aligned} \tag{6}$$

The following result provides an alternative expression for $\widehat{\boldsymbol{\alpha}}_\beta$, given in Definition 3, which is closer to (4) in its expression, since only a divergence measure between two probabilities is involved. Given two probability vectors $\boldsymbol{p} = (p_1, \ldots, p_M)^T$ and $\boldsymbol{q} = (q_1, \ldots, q_M)^T$, the power density divergence measure between $\boldsymbol{p}$ and $\boldsymbol{q}$, with tuning parameter $\beta > 0$, is given by

$$d_\beta(\boldsymbol{p}, \boldsymbol{q}) = \sum_{j=1}^{M}\left\{q_j^{\beta+1} - (1+\tfrac{1}{\beta})q_j^{\beta}p_j + \tfrac{1}{\beta}p_j^{1+\beta}\right\},$$

and for $\beta = 0$,

$$d_0(\boldsymbol{p}, \boldsymbol{q}) = \lim_{\beta\to 0^+} d_\beta(\boldsymbol{p}, \boldsymbol{q}) = d_{KL}(\boldsymbol{p}, \boldsymbol{q}).$$

Therefore, the density power divergence measure between the probability vectors $\widehat{\boldsymbol{p}}$ and $\boldsymbol{p}(\boldsymbol{\alpha})$, with tuning parameter $\beta > 0$, has the expression

$$\begin{aligned} d_\beta(\widehat{\boldsymbol{p}}, \boldsymbol{p}(\boldsymbol{\alpha})) &= \frac{1}{(IJ)^{\beta+1}}\sum_{i=1}^{I}\sum_{j=1}^{J}\left\{\pi_{ij}^{1+\beta}(\boldsymbol{\alpha}) - \tfrac{\beta+1}{\beta}\pi_{ij}^{\beta}(\boldsymbol{\alpha})\frac{n_{ij}}{K}\right. \\ &\quad + \tfrac{1}{\beta}\left(\frac{n_{ij}}{K}\right)^{1+\beta} + \left(1-\pi_{ij}(\boldsymbol{\alpha})\right)^{1+\beta} \\ &\quad - \tfrac{\beta+1}{\beta}\left(1-\pi_{ij}(\boldsymbol{\alpha})\right)^{\beta}\tfrac{K-n_{ij}}{K} \\ &\quad \left. + \tfrac{1}{\beta}\left(\frac{K-n_{ij}}{K}\right)^{1+\beta}\right\}, \end{aligned} \tag{7}$$

and for $\beta = 0$

$$d_{\beta=0}(\widehat{\boldsymbol{p}}, \boldsymbol{p}(\boldsymbol{\alpha})) = \lim_{\beta\to 0^+} d_\beta(\widehat{\boldsymbol{p}}, \boldsymbol{p}(\boldsymbol{\alpha})) = d_{KL}(\widehat{\boldsymbol{p}}, \boldsymbol{p}(\boldsymbol{\alpha})).$$

Theorem 4: The MDPDE of $\boldsymbol{\alpha}$, with tuning parameter $\beta \geq 0$, given in Definition 3, can be alternatively defined as

$$\widehat{\boldsymbol{\alpha}}_\beta = \arg\min_{\boldsymbol{\alpha}\in\Theta} d_\beta(\widehat{\boldsymbol{p}}, \boldsymbol{p}(\boldsymbol{\alpha})), \tag{8}$$

where $d_\beta(\widehat{\boldsymbol{p}}, \boldsymbol{p}(\boldsymbol{\alpha}))$ is as in (7).

In the following result, the estimating equations needed to get the MDPDEs are presented.

Theorem 5: The MDPDE of $\boldsymbol{\alpha}$ with tuning parameter $\beta \geq 0$, $\widehat{\boldsymbol{\alpha}}_\beta$, can be obtained as the solution of the following equations:

$$\begin{aligned} &\sum_{i=1}^{I}\sum_{j=1}^{J}\left(K\ F(t_j|\lambda_{w_i}(\boldsymbol{\alpha})) - n_{ij}\right) f(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))t_j \\ &\qquad \left[F^{\beta-1}(t_j|\lambda_{w_i}(\boldsymbol{\alpha})) + \left(1 - F(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))\right)^{\beta-1}\right] = 0, \end{aligned} \tag{9}$$

$$\begin{aligned} &\sum_{i=1}^{I}\sum_{j=1}^{J}\left(K\ F(t_j|\lambda_{w_i}(\boldsymbol{\alpha})) - n_{ij}\right) f(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))t_j w_i \\ &\qquad \left[F^{\beta-1}(t_j|\lambda_{w_i}(\boldsymbol{\alpha})) + \left(1 - F(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))\right)^{\beta-1}\right] = 0. \end{aligned} \tag{10}$$

In the following results, the asymptotic distribution of the MDPDE of $\boldsymbol{\alpha}$, $\widehat{\boldsymbol{\alpha}}_\beta$, for the one-shot device model is presented.

Theorem 6: The asymptotic distribution of the MDPDE $\widehat{\boldsymbol{\alpha}}_\beta$ is given by

$$\sqrt{K}\left(\widehat{\boldsymbol{\alpha}}_\beta - \boldsymbol{\alpha}_0\right) \underset{K\to\infty}{\overset{\mathcal{L}}{\longrightarrow}} \mathcal{N}\left(\boldsymbol{0}, \bar{\boldsymbol{J}}_\beta^{-1}(\boldsymbol{\alpha}_0)\bar{\boldsymbol{K}}_\beta(\boldsymbol{\alpha}_0)\bar{\boldsymbol{J}}_\beta^{-1}(\boldsymbol{\alpha}_0)\right),$$

where

$$\begin{aligned} \bar{\boldsymbol{J}}_\beta(\boldsymbol{\alpha}) &= \sum_{i=1}^{I}\begin{pmatrix} \frac{1}{\alpha_0^2} & \frac{w_i}{\alpha_0} \\ \frac{w_i}{\alpha_0} & w_i^2 \end{pmatrix}\sum_{j=1}^{J} t_j^2 f^2(t_j|\lambda_{w_i}(\boldsymbol{\alpha})) \\ &\quad \times\left[F^{\beta-1}(t_j|\lambda_{w_i}(\boldsymbol{\alpha})) + (1-F(t_j|\lambda_{w_i}(\boldsymbol{\alpha})))^{\beta-1}\right], \end{aligned} \tag{11}$$

$$\begin{aligned} \bar{\boldsymbol{K}}_\beta(\boldsymbol{\alpha}) &= \sum_{i=1}^{I}\begin{pmatrix} \frac{1}{\alpha_0^2} & \frac{w_i}{\alpha_0} \\ \frac{w_i}{\alpha_0} & w_i^2 \end{pmatrix}\sum_{j=1}^{J} t_j^2 f^2(t_j|\lambda_{w_i}(\boldsymbol{\alpha})) \\ &\quad \times\left\{\left[F^{2\beta-1}(t_j|\lambda_{w_i}(\boldsymbol{\alpha})) + (1-F(t_j|\lambda_{w_i}(\boldsymbol{\alpha})))^{2\beta-1}\right]\right. \\ &\quad \left. - \left[F^{\beta}(t_j|\lambda_{w_i}(\boldsymbol{\alpha})) - (1-F(t_j|\lambda_{w_i}(\boldsymbol{\alpha})))^{\beta}\right]^2\right\}, \end{aligned} \tag{12}$$

and $F(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))$ and $f(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))$ are as given in (1) and (2), respectively.

Since $\widehat{\boldsymbol{\alpha}}_{\beta=0}$ is the MLE of $\boldsymbol{\alpha}$, obtained by maximizing $\log\mathcal{L}(\boldsymbol{\alpha}\,|K, \boldsymbol{n}, \boldsymbol{t}, \boldsymbol{w})$, or equivalently by minimizing

$$\begin{aligned} d_{\beta=0}(\widehat{\boldsymbol{p}}, \boldsymbol{p}(\boldsymbol{\alpha})) &= \lim_{\beta\to 0^-} d_\beta(\widehat{\boldsymbol{p}}, \boldsymbol{p}(\boldsymbol{\alpha})) = d_{KL}(\widehat{\boldsymbol{p}}, \boldsymbol{p}(\boldsymbol{\alpha})) \\ &= \frac{1}{IJK}\left(s - \log\mathcal{L}(\boldsymbol{\alpha}\,|K, \boldsymbol{n}, \boldsymbol{t}, \boldsymbol{w})\right), \end{aligned}$$

the following result relates the asymptotic distribution of $\widehat{\boldsymbol{\alpha}}_{\beta=0}$ given previously in terms of $\bar{\boldsymbol{J}}_{\beta=0}(\boldsymbol{\alpha}_0)$ and $\bar{\boldsymbol{K}}_{\beta=0}(\boldsymbol{\alpha}_0)$, with respect to the Fisher information matrix, well-known in the classical asymptotic theory of the MLEs.

Theorem 7: The asymptotic distribution of the MLE of $\boldsymbol{\alpha}$, $\widehat{\boldsymbol{\alpha}}_{\beta=0}$, is

$$\sqrt{K}\left(\widehat{\boldsymbol{\alpha}}_{\beta=0}-\boldsymbol{\alpha}_0\right) \underset{K\to\infty}{\xrightarrow{\mathcal{L}}} \mathcal{N}\left(\mathbf{0}, \tfrac{1}{IJ}\boldsymbol{I}_F^{-1}\left(\boldsymbol{\alpha}_0\right)\right),$$

where

$$\boldsymbol{I}_F(\boldsymbol{\alpha}) = \frac{1}{IJ}\sum_{i=1}^{I}\begin{pmatrix} \frac{1}{\alpha_0^2} & \frac{w_i}{\alpha_0} \\ \frac{w_i}{\alpha_0} & w_i^2 \end{pmatrix}\sum_{j=1}^{J} t_j^2 \frac{f^2(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))}{F(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))(1-F(t_j|\lambda_{w_i}(\boldsymbol{\alpha})))}$$

is the Fisher Information matrix for the one-shot device model. In addition, relating the theory of MDPDEs with the Fisher Information matrix, we have

$$\boldsymbol{J}_{\beta=0}(\boldsymbol{\alpha}) = \boldsymbol{K}_{\beta=0}(\boldsymbol{\alpha}) = \boldsymbol{I}_F(\boldsymbol{\alpha}).$$

Remark 8: A referee suggested that the mínimum DPD estimator could be adapted to the context of empirical likelihood. In order to fit the model to the framework of empirical likelihood, we need to consider a sample of independent and identically distributed observations, while this model considers in principle non-homogeneous observations since every test condition has a pre-specified inspection time t_j, $j = 1, \ldots, J$, and temperature w_i $i = 1, \ldots, I$. To overcome this difficulty, we could consider that K individuals are randomly assigned to be at test condition (C_1, C_2) such that $\Pr(C_1 = i, C_2 = j) = \frac{1}{IJ}$ and $(w(C_1)|C_1 = i) = w_i$, $(t(C_2)|C_2 = j) = t_j$. In the new setting, we could consider

$$\begin{aligned} &Y_k(C_1, C_2) \overset{ind}{\sim} \mathrm{Ber}(\Pi(C_1, C_2|\boldsymbol{\alpha})),\ k = 1, \ldots, K, \\ &\Pi(C_1, C_2|\boldsymbol{\alpha}) = F(t(C_2)|\lambda_{w(C_1)}(\boldsymbol{\alpha})), \end{aligned}$$

with $Y_k(C_1, C_2)$ and $\Pi(C_1, C_2|\boldsymbol{\alpha})$ being the random generators of the indices of $(Y_k(C_1, C_2)|C_1 = i, C_2 = j) = Y_{ijk}$ and $(\Pi(C_1, C_2|\boldsymbol{\alpha})|C_1 = i, C_2 = j) = \pi_{ij}(\boldsymbol{\alpha})$, respectively, in such a way that now the random variables are also identically distributed. In this setting now, for $\beta > 0$,

$$\widehat{\boldsymbol{\alpha}}_\beta = \arg\min_{\boldsymbol{\alpha}\in\Theta_{IJK}} E\left[d_\beta^*\left(\widehat{p}(C_1, C_2), \Pi(C_1, C_2|\boldsymbol{\alpha})\right)\right], \quad (13)$$

where $\widehat{p}(C_1, C_2) = \frac{1}{K}\sum_{k=1}^{K} Y_k(C_1, C_2)$ is the random generator of the indices of $\widehat{p}_{ij} = (\widehat{p}(C_1, C_2)|C_1 = i, C_2 = j) = n_{ij}/K$ with n_{ij} being a realization of $\sum_{k=1}^{K} Y_{ijk}$ and

$$\begin{aligned} &E\left[d_\beta^*\left(\widehat{p}(C_1, C_2), \Pi(C_1, C_2|\boldsymbol{\alpha})\right)\right] \\ &\quad = E\left[E\left[d_\beta^*\left(\widehat{p}(C_1, C_2), \Pi(C_1, C_2|\boldsymbol{\alpha})\right)|C_1, C_2\right]\right]. \end{aligned}$$

Notice that the expression of the objective function in (13) is equivalent to (5), but the parameter space Θ_{IJK} should contain the sample moment constraints. Using a notation similar to the one in [12], we could consider

$$\widehat{\boldsymbol{a}}_\beta = \arg\min_{\boldsymbol{\alpha}\in\Theta_{IJK}} \frac{1}{IJ}\sum_{i=1}^{I}\sum_{j=1}^{J} \overline{V}_K\left(\pi_{ij}(\boldsymbol{\alpha})\right)$$

with

$$\begin{aligned} \overline{V}_K\left(\pi_{ij}(\boldsymbol{\alpha})\right) &= \frac{1}{K}\sum_{k=1}^{K} V_k\left(\pi_{ij}(\boldsymbol{\alpha})\right), \\ V_k\left(\pi_{ij}(\boldsymbol{\alpha})\right) &= \pi_{ij}^{\beta+1}(\boldsymbol{\alpha}) + \left(1-\pi_{ij}(\boldsymbol{\alpha})\right)^{\beta+1} \\ &\quad - \tfrac{\beta+1}{\beta}\left(Y_{ijk}\pi_{ij}^{\beta}(\boldsymbol{\alpha}) + (1-Y_{ijk})\left(1-\pi_{ij}(\boldsymbol{\alpha})\right)^{\beta}\right), \\ \frac{1}{K}\sum_{k=1}^{K} V_k\left(\pi_{ij}(\boldsymbol{\alpha})\right) &= \pi_{ij}^{\beta+1}(\boldsymbol{\alpha}) + \left(1-\pi_{ij}(\boldsymbol{\alpha})\right)^{\beta+1} \\ &\quad - \tfrac{\beta+1}{\beta}\left(\widehat{p}_{ij}\pi_{ij}^{\beta}(\boldsymbol{\alpha}) + (1-\widehat{p}_{ij})\left(1-\pi_{ij}(\boldsymbol{\alpha})\right)^{\beta}\right). \end{aligned}$$

Other results in relation to empirical likelihood and distances or divergences can be found in [13] and [14], and the references therein.

IV. Robust Z-Type Tests

For testing the null hypothesis of a linear combination of $\boldsymbol{\alpha} = (\alpha_0, \alpha_1)^T$, H_0: $m_0\alpha_0 + m_1\alpha_1 = d$, or equivalently

$$H_0\colon \boldsymbol{m}^T\boldsymbol{\alpha} = d, \quad (14)$$

where $\boldsymbol{m}^T = (m_0, m_1)$, it is important to know the asymptotic distribution of the MDPDE of $\boldsymbol{\alpha}$. In particular, in case we wish to test if the different temperatures do not affect the model of the one-shot devices, we need to consider $\boldsymbol{m}^T = (m_0, m_1) = (0, 1)$ and $d = 0$. In the following definition, we present Z-type test statistics based on $\widehat{\boldsymbol{\alpha}}_\beta$. Since Z-type test statistics are a particular case of the Wald-type test, we can say that this type of robust test statistics have been considered previously by [9] and [10].

Definition 9: Let $\widehat{\boldsymbol{\alpha}}_\beta = (\widehat{\alpha}_{0,\beta}, \widehat{\alpha}_{1,\beta})^T$ be the MDPDE of $\boldsymbol{\alpha} = (\alpha_0, \alpha_1)^T$. Then, the family of Z-type test statistics for testing (14) is given by

$$Z_K(\widehat{\boldsymbol{\alpha}}_\beta) = \sqrt{\frac{K}{\boldsymbol{m}^T\bar{\boldsymbol{J}}_\beta^{-1}(\widehat{\boldsymbol{\alpha}}_\beta)\bar{\boldsymbol{K}}_\beta(\widehat{\boldsymbol{\alpha}}_\beta)\bar{\boldsymbol{J}}_\beta^{-1}(\widehat{\boldsymbol{\alpha}}_\beta)\boldsymbol{m}}}(\boldsymbol{m}^T\widehat{\boldsymbol{\alpha}}_\beta - d). \quad (15)$$

In the following theorem, the asymptotic distribution of $Z_K(\widehat{\boldsymbol{\alpha}}_\beta)$ is presented.

Theorem 10: The asymptotic distribution of Z-type test statistics, $Z_K(\widehat{\boldsymbol{\alpha}}_\beta)$, defined in (15), is standard normal.

Based on Theorem 10, the null hypothesis in (14) will be rejected, with significance level α, if $\left|Z_K(\widehat{\boldsymbol{\alpha}}_\beta)\right| > z_{\frac{\alpha}{2}}$, where $z_{\frac{\alpha}{2}}$ is a right hand side quantile of order $\frac{\alpha}{2}$ of a normal distribution. Now, we shall present a result providing an approximation, to the power function, for the test statistics defined in (15).

Theorem 11: Let $\boldsymbol{\alpha}^* \in \Theta$ be the true value of the parameter $\boldsymbol{\alpha}$ so that

$$\widehat{\boldsymbol{\alpha}}_\beta \underset{K\to\infty}{\xrightarrow{\mathcal{P}}} \boldsymbol{\alpha}^* \in \Theta$$

and $\boldsymbol{m}^T\boldsymbol{\alpha}^* \neq d$. Then, the approximate power function of the test statistic in (15) at $\boldsymbol{\alpha}^*$ is as given below, where $\Phi(\cdot)$ is the standard normal distribution function,

$$\begin{aligned} &\pi\left(\boldsymbol{\alpha}^*\right) \simeq 2\times \quad (16) \\ &\left(1-\Phi\left(z_{\frac{\alpha}{2}} - \sqrt{\frac{K}{\boldsymbol{m}^T\bar{\boldsymbol{J}}_\beta^{-1}(\boldsymbol{\alpha}^*)\bar{\boldsymbol{K}}_\beta(\boldsymbol{\alpha}^*)\bar{\boldsymbol{J}}_\beta^{-1}(\boldsymbol{\alpha}^*)\boldsymbol{m}}}(\boldsymbol{m}^T\boldsymbol{\alpha}^* - d)\right)\right). \end{aligned}$$

Remark 12: Based on the above results, it is possible to provide an explicit expression for the number of devices as

$$K = \left[\frac{\boldsymbol{m}^T \bar{\boldsymbol{J}}_\beta^{-1}(\boldsymbol{\alpha}^*)\bar{\boldsymbol{K}}_\beta(\boldsymbol{\alpha}^*)\bar{\boldsymbol{J}}_\beta^{-1}(\boldsymbol{\alpha}^*)\boldsymbol{m}}{\boldsymbol{m}^T\boldsymbol{\alpha}^* - d} \left(z_{\frac{\alpha}{2}} - \Phi^{-1}(1 - \tfrac{\pi^*}{2}) \right)^2 \right] + 1,$$

placed under temperatures w_i, $i = 1, \ldots, I$, at each time, t_j, $j = 1, 2, \ldots, J$, necessary to get a fixed power π^* for a specific significance level α. Here, $[\cdot]$ denotes the largest integer less than or equal to $\cdot$.

V. Robustness of the Z-Type Tests

An important concept in robustness theory is the influence function. For any estimator defined in terms of an statistical functional $\boldsymbol{U}(F)$ from the true distribution F, its influence function is defined as

$$\mathcal{IF}(x, \boldsymbol{U}, F) = \lim_{\varepsilon \downarrow 0} \frac{\boldsymbol{U}(F_\varepsilon) - \boldsymbol{U}(F)}{\varepsilon} = \left.\frac{\partial \boldsymbol{U}(F_\varepsilon)}{\partial \varepsilon}\right|_{\varepsilon = 0^+}, \qquad (17)$$

where $F_\varepsilon = (1 - \varepsilon)F + \varepsilon\Delta_x$, with ε being the contamination proportion and Δ_x being the degenerate distribution at the contamination point x. Thus, the (first-order) influence function (IF), as a function of x, measures the standardized asymptotic bias (in its first-order approximation) caused by the infinitesimal contamination at the point x. The maximum of this IF over x indicates the extent of bias due to contamination and so smaller its value, the more robust the estimator is.

Let us now consider the one-shot device model under the exponential distribution presented here. Let $f_{ij,\boldsymbol{\alpha}}$ represent the probability mass function for each individual from the i-th temperature and j-th inspection time, $i = 1, \ldots, I$, $j = 1, \ldots, J$, associated with the model. As in each combination ij, K individuals are tested, denoting by $Y_{ij}^{(k)}$ ($k = 1, \ldots, K$) the random variable which has Bernoulli distribution with probability of success $\pi_{ij}(\boldsymbol{\alpha})$, we have

$$f_{ij,\boldsymbol{\alpha}}(y) = y\pi_{ij}(\boldsymbol{\alpha}) + (1 - y)(1 - \pi_{ij}(\boldsymbol{\alpha})), \text{ for } y \in \{0, 1\}.$$

Let g_{ij} represent the true probability mass function for each individual from the ij-th group of K individuals $i = 1, \ldots, I$, $j = 1, \ldots, J$. In vectorial notation, we let

$$\boldsymbol{g} = (g_{11} \otimes \boldsymbol{1}_K^T, \ldots, g_{IJ} \otimes \boldsymbol{1}_K^T)^T,$$
$$\boldsymbol{f}_{\boldsymbol{\alpha}} = (f_{11,\boldsymbol{\alpha}} \otimes \boldsymbol{1}_K^T, \ldots, f_{IJ,\boldsymbol{\alpha}} \otimes \boldsymbol{1}_K^T)^T.$$

The DPD for the sample, for $\beta > 0$, is defined as

$$d_\beta(\boldsymbol{g}, \boldsymbol{f}_{\boldsymbol{\alpha}}) = \frac{1}{IJ} \sum_{i=1}^I \sum_{j=1}^J d_\beta(g_{ij}, f_{ij,\boldsymbol{\alpha}}), \qquad (18)$$

where

$$\begin{aligned} d_\beta(g_{ij}, f_{ij,\boldsymbol{\alpha}}) &\propto \frac{1}{K} \sum_{k=1}^K H_{\beta,\boldsymbol{\alpha}}(Y_{ij}^{(k)}) \\ &= \pi_{ij}^{\beta+1}(\boldsymbol{\alpha}) + (1 - \pi_{ij}(\boldsymbol{\alpha}))^{\beta+1} \\ &\quad - \frac{\beta + 1}{\beta} \left[\frac{N_{ij}}{K}\pi_{ij}^\beta(\boldsymbol{\alpha}) + \frac{1 - N_{ij}}{K}(1 - \pi_{ij}(\boldsymbol{\alpha}))^\beta \right], \end{aligned}$$

$$\begin{aligned} H_{\beta,\boldsymbol{\alpha}}(Y_{ij}^{(k)}) &= \pi_{ij}^{\beta+1}(\boldsymbol{\alpha}) + (1 - \pi_{ij}(\boldsymbol{\alpha}))^{\beta+1} \\ &\quad - \frac{\beta + 1}{\beta} f_{ij,\boldsymbol{\alpha}}^\beta(Y_{ij}^{(k)}), \\ f_{ij,\boldsymbol{\alpha}}^\beta(Y_{ij}^{(k)}) &= Y_{ij}^{(k)}\pi_{ij}^\beta(\boldsymbol{\alpha}) + (1 - Y_{ij}^{(k)})(1 - \pi_{ij}(\boldsymbol{\alpha}))^\beta, \end{aligned}$$

and $N_{ij} = \sum_{k=1}^K Y_{ij}^{(k)}$.

Then, the statistical functional $\boldsymbol{U}_\beta(\boldsymbol{G})$ corresponding to the proposed MDPDE, $\widehat{\boldsymbol{\alpha}}_\beta$, of $\boldsymbol{\alpha}$ is defined as the minimizer of

$$H(\boldsymbol{\alpha}) = \sum_{i=1}^I \sum_{j=1}^J d_\beta(g_{ij}, f_{ij,\boldsymbol{\alpha}})$$

whenever it exists. When the assumption of the model holds with true parameter $\boldsymbol{\alpha}_0$, we have $g_{ij} = f_{ij,\boldsymbol{\alpha}_0}$ and $H(\boldsymbol{\alpha}) = IJd_\beta(\boldsymbol{p}(\boldsymbol{\alpha}_0), \boldsymbol{p}(\boldsymbol{\alpha}))$; this is minimized at $\boldsymbol{\alpha} = \boldsymbol{\alpha}_0$ implying the Fisher consistency of the MDPDE functional $\boldsymbol{U}_\beta(\boldsymbol{G})$ in our model.

The influence function of $\widehat{\boldsymbol{\alpha}}_\beta = \boldsymbol{U}_\beta(\boldsymbol{G})$, with respect to the k_0-th observation of the ij_0-th group of observations, is given by

$$\begin{aligned} \mathcal{IF}\left(ij_0, k_0, x, \boldsymbol{U}_\beta, \boldsymbol{F}_{\boldsymbol{\alpha}}\right) &= \left.\frac{\partial \boldsymbol{U}_\beta\left(\boldsymbol{F}_{\boldsymbol{\alpha},x,\varepsilon}\right)}{\partial \varepsilon}\right|_{\varepsilon=0^+} \\ &= \boldsymbol{J}_\beta^{-1}(\boldsymbol{\alpha})\left(F(t_j|\lambda_{w_i}(\boldsymbol{\alpha})) - x\right) f(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))t_i\boldsymbol{d}_i \\ &\quad \times \left[F^{\beta-1}(t_j|\lambda_{w_i}(\boldsymbol{\alpha})) + (1 - F(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))^{\beta-1} \right], \qquad (19) \end{aligned}$$

where $\boldsymbol{d}_i = (\frac{1}{\alpha_0}, w_i)^T$,

$$\boldsymbol{F}_{\boldsymbol{\alpha},x,\varepsilon} = (1 - \varepsilon)\boldsymbol{F}_{\boldsymbol{\alpha}} + \varepsilon\Delta_x\boldsymbol{e}_{ij_0,k_0}, \quad \Delta_x(z) = I(x \le z),$$
$$\boldsymbol{F}_{\boldsymbol{\alpha}} = (F(t_1|\lambda_{w_1}(\boldsymbol{\alpha})) \otimes \boldsymbol{1}_K^T, \ldots, F(t_J|\lambda_{w_I}(\boldsymbol{\alpha})) \otimes \boldsymbol{1}_K^T)^T,$$

$\boldsymbol{e}_{ij_0,jk_0}$ is a vector with 1 in the position (ij_0, k_0) in lexicographical order.

The influence function of $\widehat{\boldsymbol{\alpha}}_\beta = \boldsymbol{U}_\beta(\widehat{\boldsymbol{G}})$ with respect to all the observations is given by

$$\begin{aligned} \mathcal{IF}\left(\underline{x}, \boldsymbol{U}_\beta, \boldsymbol{F}_{\boldsymbol{\alpha}}\right) &= \left.\frac{\partial \boldsymbol{U}_\beta\left(\boldsymbol{F}_{\boldsymbol{\alpha},\underline{x},\varepsilon}\right)}{\partial \varepsilon}\right|_{\varepsilon=0^+} \\ &= \boldsymbol{J}_\beta^{-1}(\boldsymbol{\alpha}) \sum_{i=1}^I \sum_{j=1}^J (KF(t_j|\lambda_{w_i}(\boldsymbol{\alpha})) - x_{ij}) f(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))t_i\boldsymbol{d}_i \\ &\quad \times \left[F^{\beta-1}(t_j|\lambda_{w_i}(\boldsymbol{\alpha})) + (1 - F(t_j|\lambda_{w_i}(\boldsymbol{\alpha})))^{\beta-1} \right], \qquad (20) \end{aligned}$$

where

$$\boldsymbol{F}_{\boldsymbol{\alpha},\underline{x},\varepsilon} = (1 - \varepsilon)\boldsymbol{F}_{\boldsymbol{\alpha}} + \varepsilon \sum_{i=1}^I \sum_{j=1}^J \sum_{k=1}^K \Delta_{x_{ijk}}\boldsymbol{e}_{ij,k},$$

$$\Delta_{x_{ijk}}(y) = I(x_{ijk} \le y), \quad x_{ij} = \sum_{k=1}^{K_i} x_{ijk}.$$

The functional associated with the Z-type tests, $Z_K(\widehat{\boldsymbol{\alpha}}_\beta)$ evaluated at $\boldsymbol{U}_\beta(\boldsymbol{F}_{\boldsymbol{\alpha}})$ is given by

$$Z_\beta(\boldsymbol{F}_{\boldsymbol{\alpha}}) = \sqrt{\frac{K}{\boldsymbol{m}^T\boldsymbol{J}_\beta^{-1}(\boldsymbol{U}_\beta(\boldsymbol{F}_{\boldsymbol{\alpha}}))\boldsymbol{K}_\beta(\boldsymbol{U}_\beta(\boldsymbol{F}_{\boldsymbol{\alpha}}))\boldsymbol{J}_\beta^{-1}(\boldsymbol{U}_\beta(\boldsymbol{F}_{\boldsymbol{\alpha}}))\boldsymbol{m}}} \times (\boldsymbol{m}^T\boldsymbol{U}_\beta(\boldsymbol{F}_{\boldsymbol{\alpha}}) - \boldsymbol{d})$$

The influence function with respect to the k_0-th observation of the ij_0-th group of observations, of the functional associated with the Z -type test statistics for testing the composite null hypothesis in (14), is then given by

$$\mathcal{IF}\left(i_0, j_0, x, Z_\beta, \boldsymbol{F_\alpha}\right) = \left.\frac{\partial Z_\beta\left(\boldsymbol{F}_{\boldsymbol{\alpha},x,\varepsilon}\right)}{\partial\varepsilon}\right|_{\varepsilon=0^+}.$$

But,

$$\frac{\partial Z_\beta\left(\boldsymbol{F}_{\boldsymbol{\alpha},x,\varepsilon}\right)}{\partial\varepsilon} = \frac{\partial \boldsymbol{l}\left(\boldsymbol{U}_\beta\left(\boldsymbol{F}_{\boldsymbol{\alpha},x,\varepsilon}\right)\right)}{\partial\varepsilon}(\boldsymbol{m}^T\boldsymbol{U}_\beta(\boldsymbol{F}_{\boldsymbol{\alpha},x,\varepsilon}) - \boldsymbol{d}) + l\left(\boldsymbol{U}_\beta\left(\boldsymbol{F}_{\boldsymbol{\alpha},x,\varepsilon}\right)\right)\boldsymbol{m}^T\frac{\partial \boldsymbol{U}_\beta\left(\boldsymbol{F}_{\boldsymbol{\alpha},x,\varepsilon}\right)}{\partial\varepsilon},$$

with

$$l\left(\boldsymbol{U}_\beta\left(\boldsymbol{F}_{\boldsymbol{\alpha},x,\varepsilon}\right)\right) = \sqrt{\frac{K}{\boldsymbol{m}^T\boldsymbol{J}_\beta^{-1}(\boldsymbol{U}_\beta(\boldsymbol{F}_{\boldsymbol{\alpha},x,\varepsilon}))\boldsymbol{K}_\beta(\boldsymbol{U}_\beta(\boldsymbol{F}_{\boldsymbol{\alpha},x,\varepsilon}))\boldsymbol{J}_\beta^{-1}(\boldsymbol{U}_\beta(\boldsymbol{F}_{\boldsymbol{\alpha},x,\varepsilon}))\boldsymbol{m}}}.$$

Now, we have

$$\mathcal{IF}\left(i_0, j_0, x, Z_\beta, \boldsymbol{F_\alpha}\right) = l(\boldsymbol{\alpha}_0)\boldsymbol{m}^T\left.\frac{\partial \boldsymbol{U}_\beta\left(\boldsymbol{F}_{\boldsymbol{\alpha},x,\varepsilon}\right)}{\partial\varepsilon}\right|_{\varepsilon=0^+}, \quad (21)$$

where $\left.\frac{\partial \boldsymbol{U}_\beta(\boldsymbol{F}_{\boldsymbol{\alpha},x,\varepsilon})}{\partial\varepsilon}\right|_{\varepsilon=0^+}$ is the influence function of the estimator given in (19).

Similarly, in all the indices we have

$$\begin{aligned}\mathcal{IF}\left(\underline{x}, Z_\beta, \boldsymbol{F_\alpha}\right) &= \left.\frac{\partial Z_\beta\left(\boldsymbol{F}_{\boldsymbol{\alpha},\underline{x},\varepsilon}\right)}{\partial\varepsilon}\right|_{\varepsilon=0^+} \quad (22)\\ &= l(\boldsymbol{\alpha})\boldsymbol{m}^T\left.\frac{\partial \boldsymbol{U}_\beta\left(\boldsymbol{F}_{\boldsymbol{\alpha},\underline{x},\varepsilon}\right)}{\partial\varepsilon}\right|_{\varepsilon=0^+}\\ &= l(\boldsymbol{\alpha})\boldsymbol{m}^T\boldsymbol{J}_\beta^{-1}(\boldsymbol{\alpha})\\ &\times\sum_{i=1}^{I}\sum_{j=1}^{J}(KF(t_j|\lambda_{w_i}(\boldsymbol{\alpha})) - x_{ij})f(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))t_i\boldsymbol{d}_i\\ &\times\left[F^{\beta-1}(t_j|\lambda_{w_i}(\boldsymbol{\alpha})) + (1-F(t_j|\lambda_{w_i}(\boldsymbol{\alpha})))^{\beta-1}\right].\end{aligned}$$

Let $h_0(w,t;\boldsymbol{\alpha},\beta) = h(w,t;\boldsymbol{\alpha},\beta)w$ and $h_1(w,t;\boldsymbol{\alpha},\beta) = h(w,t;\boldsymbol{\alpha},\beta)\frac{1}{\alpha_0}$, where

$$\begin{aligned}h(w,t;\boldsymbol{\alpha},\beta) = &\left[\exp\{\alpha_1 w - \alpha_0\exp\{\alpha_1 w\}t\}\right.\\ &(1-\exp\{-\alpha_0\exp\{\alpha_1 w\}t\})^{\beta-1}\\ &\left.+\exp\{\alpha_1 w - \beta\alpha_0\exp\{\alpha_1 w\}t\}\right]\alpha_0 t,\end{aligned}$$

be factors of the influence function of $\widehat{\alpha}_{\beta,0}$ and $\widehat{\alpha}_{\beta,1}$ respectively, given in vectorial form in either (19) or (20). Based on this, might be commented on conditions for boundedness of the influence influence functions presented in this paper, either for estimation (19-20) or testing (21-22), that they are bounded on x or $\underline{x}$, however if $\beta = 0$ the norm of the bidimensional influence functions can be very large on $(w,t) = (w_I, t_J)$ when $w_1 < w_2 < \ldots < w_I, t_1 < t_2 < \ldots < t_J$, in comparison with $\beta > 0$, since

$$\begin{aligned}\lim_{\substack{w\to+\infty\\(\alpha_1<0)}} h_0(w,t;\boldsymbol{\alpha},\beta) &= \lim_{\substack{w\to+\infty\\(\alpha_1>0)}} h_1(w,t;\boldsymbol{\alpha},\beta)\\ &= \lim_{t\to+\infty} h_0(w,t;\boldsymbol{\alpha},\beta)\\ &= \lim_{t\to+\infty} h_1(w,t;\boldsymbol{\alpha},\beta)\\ &\begin{cases} = \infty, \text{ if } \beta = 0\\ < \infty, \text{ if } \beta > 0\end{cases}.\end{aligned}$$

In this regard, the lower-right corner cell of the contingency table, $(i_0, j_0) = (I, J)$, could be a strong leverage cell and this fact justifies our robust statistical approach for appropriate inference procedure in the one-shot device model with the exponential distribution. In accelerated processes, inspection time (t) tends not to be large, and the influence of large values is naturally associated to temperature (w). In addition, it is interesting to note that

$$F(t|\lambda_w(\boldsymbol{\alpha})) \underset{w\to+\infty}{\longrightarrow} \begin{cases}1, \text{ if } \alpha_1 > 0\\ 0, \text{ if } \alpha_1 < 0\end{cases},$$

which matches with the degenerated distributions for cells (with either K or 0 observations w.p. 1), and similarly

$$\lambda_w(\boldsymbol{\alpha}) \underset{w\to+\infty}{\longrightarrow} \begin{cases}+\infty, \text{ if } \alpha_1 > 0\\ 0, \quad\ \text{ if } \alpha_1 < 0\end{cases}$$

approaches both boundaries of $\lambda_w(\boldsymbol{\alpha}) \in (0,+\infty)$. Even though the distribution function approaches the upper-limit in the upper-bound of $t \in (0,+\infty)$, $F(t|\lambda_w(\boldsymbol{\alpha})) \underset{t\to+\infty}{\longrightarrow} 1$, the influence functions are unbounded only for $\beta = 0$ (MLE and the MLE based classical Z-type test) and bounded for $\beta > 0$ (robust Z -type tests). When the distribution function approaches the lower-limit in the lower-bound of $t \in (0,+\infty)$, $F(t|\lambda_w(\boldsymbol{\alpha})) \underset{t\to 0^+}{\longrightarrow} 0$, the influence functions are all bounded.

VI. Real Data Examples

In this section, we present some numerical examples to illustrate the inferential results developed in the preceding sections. The first one is an application to the reliability example considered in Section II, and the other two are real applications to tumorigenicity experiments considered earlier by other authors.

A. *Example 1 (Reliability Experiment)*

Based on the example introduced in Section II, in this section, the MDPDEs of the parameters of the one-shot device model are considered. As tuning parameter, $\beta \in \{0,\ 0.1,\ 0.2,\ 0.3,\ 0.4,\ 0.5,\ 0.6,\ 0.7,\ 0.8,\ 0.9,\ 1,\ 2,\ 3,\ 4\}$ are taken. In Table II, apart from the MDPDEs of $\boldsymbol{\alpha}$, the MDPDEs of the reliability function

$$R(t|\lambda_{w_0}(\boldsymbol{\alpha})) = 1 - F(t|\lambda_{w_0}(\boldsymbol{\alpha})) = e^{-\lambda_{w_0}t} = \exp(-\alpha_0 e^{\alpha_1 w_0}t)$$

are also presented at mission times (time points at which we are interested in the reliability of the unit) $t \in \{10, 20, 30\}$, namely $R(10|\lambda_{w_0}(\widehat{\boldsymbol{\alpha}}_\beta))$, $R(20|\lambda_{w_0}(\widehat{\boldsymbol{\alpha}}_\beta))$, $R(30|\lambda_{w_0}(\widehat{\boldsymbol{\alpha}}_\beta))$,

TABLE II

MDPDEs of the Parameters, the Reliability Function at Times $t \in \{10, 20, 30\}$, and Mean Lifetime at Normal Temperature of $25°C$ in One-Shot Device Testing Experiment Considered by Balakrishnan and Ling (2012)

β	$\widehat{\alpha}_{0,\beta}$	$\widehat{\alpha}_{1,\beta}$	$R(10\|\lambda_{25}(\widehat{\boldsymbol{\alpha}}_\beta))$	$R(20\|\lambda_{25}(\widehat{\boldsymbol{\alpha}}_\beta))$	$R(30\|\lambda_{25}(\widehat{\boldsymbol{\alpha}}_\beta))$	$E[T(\lambda_{25}(\widehat{\boldsymbol{\alpha}}_\beta))]$
0	0.00487	0.04732	0.85300	0.72761	0.62065	62.89490
0.1	0.00489	0.04722	0.85288	0.72741	0.62039	62.83953
0.2	0.00490	0.04714	0.85277	0.72722	0.62016	62.79031
0.3	0.00491	0.04706	0.85268	0.72706	0.61995	62.74654
0.4	0.00492	0.04700	0.85260	0.72693	0.61978	62.70965
0.5	0.00493	0.04695	0.85253	0.72681	0.61963	62.67944
0.6	0.00494	0.04690	0.85247	0.72671	0.61950	62.65188
0.7	0.00495	0.04687	0.85246	0.72669	0.61947	62.64457
0.8	0.00495	0.04683	0.85236	0.72651	0.61925	62.59732
0.9	0.00496	0.04681	0.85233	0.72646	0.61918	62.58398
1	0.00496	0.04681	0.85239	0.72656	0.61931	62.61131
2	0.00496	0.04679	0.85231	0.72644	0.61915	62.57739
3	0.00494	0.04687	0.85255	0.72684	0.61966	62.68584
4	0.00491	0.04700	0.85292	0.72748	0.62048	62.85869

as well as the MDPDEs of the mean of the lifetime $T(\lambda_{w_0}(\boldsymbol{\alpha}))$, namely,

$$E[T(\lambda_{w_0}(\boldsymbol{\alpha}))] = \frac{1}{\lambda_{w_0}(\boldsymbol{\alpha})} = \frac{1}{\alpha_0 e^{\alpha_1 w_0}},$$

under the normal operating temperature $w_0 = 25$.

Table II shows that the mean lifetime obtained by the MLE ($\beta = 0$) is greater than that obtained from the alternative MPDPDEs. However, results for all considered choices of β seem to be quite similar. The choice of β will be discussed in detail in Section VIII.

B. Example 2 (ED01 Data)

In 1974, the National Center for Toxicological Research made an experiment on 24000 female mice randomized to a control group or one of seven dose levels of a known carcinogen, called 2-Acetylaminofluorene (2-AAF). Table 1 in [15] shows the results obtained when the highest dose level (150 parts per million) was administered. The original study considered four different outcomes: Number of animals dying tumor free (DNT) and with tumor (DWT), and sacrified without tumor (SNT) and with tumor (SWT), summarized over three time intervals at 12, 18 and 33 months. A total of 3355 mice were involved in the experiment.

Reference [16] made an analysis combining SNT and SWT as the sacrificed group ($r = 0$); and denoting the cause of DNT as natural death ($r = 1$), and the cause of DWT as death due to cancer ($r = 2$). This modified data are presented in Table III, while MDPDEs of the model parameters and the corresponding estimates of mean lifetimes are presented in Table IV. Here $w = 0$ refers to control group and $w = 1$ is the test group, while $E(T_1)$ and $E(T_2)$ are the estimated mean lifetimes for sacrifice or nature death ($r = 0, 1$) and death due to cancer ($r = 2$), respectively.

From Table IV, some MDPDEs of α_{11} are seen to be negative. As pointed out in [5], this can be due to the fact that the true value of it may be quite close to zero. In fact, for the values of the tuning parameter $\beta \in \{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7\}$, the estimators of α_{11} are very close to zero,

TABLE III

Number of Mice Sacrified ($r = 0$) and Died Without Tumor ($r = 1$) and With Tumor ($r = 2$) From the ED01 Data

		$r=0$	$r=1$	$r=2$
$IT_1 = 12$	$w=0$	115	22	8
	$w=1$	110	49	16
$IT_2 = 18$	$w=0$	780	42	8
	$w=1$	540	54	26
$IT_3 = 33$	$w=0$	675	200	85
	$w=1$	510	64	51

meaning that the drug will not increase the hazard rate of the natural death outcome. Furthermore, if we look at the estimates of mean lifetimes, these last estimators show a reduction when the carcinogenic drug is administered, but the other ones, $\beta \in \{0, 0.8, 0.9, 1\}$, do not show this behavior (see Figure 1). Thus, in this case, we observe that the MDPDEs with tuning parameter $\beta \in \{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7\}$ give a more meaningful result in the context of the laboratory experiment than, in particular, the MLE ($\beta = 0$). The simulation study presented in this paper will prove how, in a general case, MDPDEs with these tuning parameters will also present a better behavior in terms of robustness. However, this conclusion can be seen as a post-hoc choice based on expert knowledge about the experiment (2-AAF is a carcinogenic). In Section VIII, a data-driven choice of the optimal β for this specific data will be carried out and will further corroborate our suggestion.

C. Example 3 (Benzidine Dihydrochloride Data)

The benzidine dihydrochloride experiment was also conducted at the National Center for Toxicological Research to examine the incidence in mice of liver tumors induced by the drug, and studied by [15] and [16]. The inspection times used on the mice were 9.37, 14.07 and 18.7 months. In Table V, the numbers of mice sacrified ($r = 0$), died without

TABLE IV
MDPDEs OF THE PARAMETERS AND THE MEAN LIFETIMES OF THE ED01 EXPERIMENT

β	$\widehat{\alpha}_{10}$	$\widehat{\alpha}_{11}$	$E_{w=0}(T_1)$	$E_{w=1}(T_1)$	$\widehat{\alpha}_{20}$	$\widehat{\alpha}_{21}$	$E_{w=0}(T_2)$	$E_{w=1}(T_2)$	$E_{w=0}(T)$	$E_{w=1}(T)$
0	0.00594	−0.12980	168.333	191.665	0.00216	0.27620	463.425	351.582	123.447	124.000
0.1	0.00702	0.09355	142.352	129.639	0.00250	0.32870	399.794	287.795	104.988	89.392
0.2	0.00698	0.06495	143.302	134.290	0.00250	0.31173	400.433	293.189	105.504	92.072
0.3	0.00703	0.00999	142.253	140.840	0.00249	0.29613	401.393	298.513	105.045	95.708
0.4	0.00690	0.00998	145.019	143.578	0.00249	0.27957	401.602	303.655	106.545	97.484
0.5	0.00677	0.00998	147.662	146.195	0.00249	0.26421	401.839	308.537	107.965	99.175
0.6	0.00666	0.00998	150.085	148.594	0.00283	0.00997	353.925	350.414	105.342	104.296
0.7	0.00682	−0.06678	146.635	156.763	0.00249	0.23702	401.985	317.157	107.415	104.876
0.8	0.00680	−0.08753	147.020	160.468	0.00279	0.00997	358.642	355.083	104.256	110.508
0.9	0.00679	−0.10530	147.321	163.680	0.00278	0.00997	360.357	356.781	104.516	112.141
1	0.00678	−0.11980	147.546	166.324	0.00277	0.00995	361.607	358.028	104.739	113.506

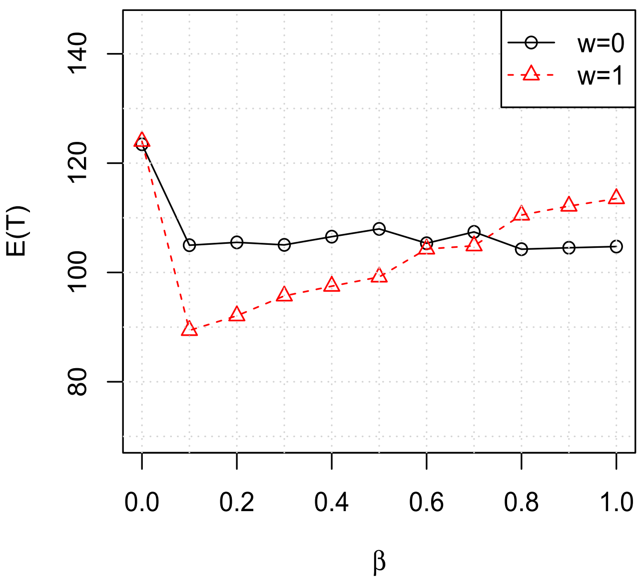

Fig. 1. MDPDEs of the mean lifetimes, for different values of the tuning parameter β, from the ED01 experiment.

TABLE V
NUMBER OF MICE SACRIFIED ($r = 0$) AND DIED WITHOUT TUMOR ($r = 1$) AND WITH TUMOR ($r = 2$) FROM THE BENZIDINE DIHYDROCHLORIDE DATA

		$r = 0$	$r = 1$	$r = 2$
$IT_1 = 9.37$	$w = 1$	70	2	0
	$w = 2$	22	3	0
$IT_2 = 14.07$	$w = 1$	48	1	0
	$w = 2$	14	4	17
$IT_3 = 18.7$	$w = 1$	35	4	7
	$w = 2$	1	1	9

tumor ($r = 1$) and died with tumor ($r = 2$), are shown, for two different doses of drug: 60 parts per million ($w = 1$) and 400 parts per million ($w = 2$). As in the previous example, we consider as "failures" the mice died due to cancer.

Table VI shows the MDPDEs of the model parameters and the corresponding estimates of mean lifetimes. Although some differences are observed in the results for different values of the tuning parameter, in all the cases, the mean lifetime shows a reduction when the carcinogenic drug is administered.

In order to have an idea of the behavior of the different MDPDEs, in relation to the efficiency as well as the robustness, we carry out an extensive Monte Carlo simulation study in the next section.

VII. SIMULATION STUDY

In this section, a simulation study is carried out to examine the behavior of the MDPDEs of the parameters of the one-shot device model, studied in Section III, as well as the Z-type tests, based on MDPDEs, detailed in Section IV. We pay special attention to the robustness issue here. It is interesting to note, in this context, the following. For each fixed time, t_j, under a fixed temperature, w_i, K devices are tested. In this sense, we can identify our data as a $I \times J$ contingency table with K observations in each cell. Hence, under this setting, we must consider "outlying cells" rather than "outlying observations". A cell which does not follow the one-shot device model will be called an outlying cell or outlier. The strong outliers may lead to reject a model fitting even if the rest of the cells fit the model properly. In other words, even though the cells seem to fit reasonably well the model, the outlying cells contribute to an increase in the values of the residuals as well as the divergence measure between the data and the fitted values according to the one-shot device model considered. Therefore, it is very important to have robust estimators as well as robust test statistics in order to avoid the undesirable effects of outliers in the data. The main purpose of this simulation study is to empirically illustrate that inside the family of MDPDEs developed here, some estimators may have better robust properties than the MLE, and the Z-type tests constructed from them can be at the same time more robustness than the classical Z-type test constructed through the MLEs.

A. The MDPDEs

In this section, we carry out a simulation study to compare the behavior of some MDPDEs with respect to the MLEs of

TABLE VI
MDPDEs OF THE PARAMETERS AND THE MEAN LIFETIMES OF THE BENZIDINE DIHYDROCHLORID EXPERIMENT

β	$\widehat{\alpha}_{10}$	$\widehat{\alpha}_{11}$	$E_{w=1}(T_1)$	$E_{w=2}(T_1)$	$\widehat{\alpha}_{20}$	$\widehat{\alpha}_{21}$	$E_{w=1}(T_2)$	$E_{w=2}(T_2)$	$E_{w=1}(T)$	$E_{w=2}(T)$
0	0.00114	1.03606	873.435	309.940	0.00029	2.41598	3433.466	306.539	699.300	154.778
0.1	0.00137	0.88718	731.822	301.376	0.00034	2.43535	2962.116	259.383	584.795	138.706
0.2	0.00141	0.86736	710.960	298.643	0.00036	2.41128	2792.303	250.469	564.971	135.684
0.3	0.00144	0.85118	695.808	297.046	0.00037	2.39392	2696.457	246.108	552.486	134.645
0.4	0.00148	0.83279	677.352	294.536	0.00039	2.37048	2580.905	241.149	534.759	131.968
0.5	0.00151	0.81685	662.421	292.673	0.00040	2.35318	2472.774	235.077	523.560	131.143
0.6	0.00154	0.80124	647.813	290.720	0.00042	2.33504	2387.518	231.128	510.204	128.697
0.7	0.00157	0.78793	636.020	289.252	0.00044	2.31249	2283.844	226.133	497.512	126.646
0.8	0.00160	0.77577	624.926	287.685	0.00045	2.29739	2208.798	222.030	487.805	125.749
0.9	0.00163	0.76485	615.166	286.300	0.00046	2.29251	2183.912	220.602	478.468	124.127
1	0.00165	0.7549	606.064	284.877	0.00047	2.27947	2121.063	217.066	471.698	123.412

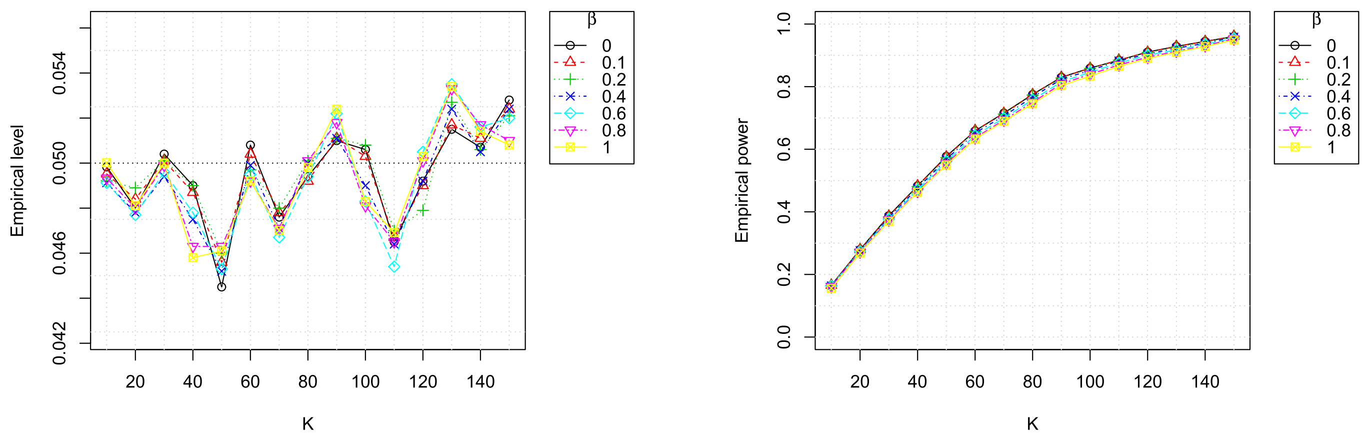

Fig. 2. Simulated levels (left) and powers (right) with no outliers in the data.

the parameters in the one-shot device model under the exponential distribution. In order to evaluate the performance of the proposed MDPDEs, as well as the MLEs, we consider the root of the mean square errors (RMSEs). We have considered a model in which, $I = J = 3$, $w \in \{35, 45, 55\}$, $t \in \{10, 20, 30\}$ and $K = 20$, as in the example in Table I, and the simulation experiment proposed by [1]. This model has been examined under three choices of $(\alpha_0, \alpha_1) = (0.005, 0.05)$, $(\alpha_0, \alpha_1) = (0.004, 0.05)$ and $(\alpha_0, \alpha_1) = (0.003, 0.05)$ for low-moderate, moderate and moderate-high reliability, respectively.

To evaluate the robustness of the MDPDEs, we have studied the behavior of this model under the consideration of an outlying cell for (w_1, t_1) in our contingency table, with $10,000$ replications and estimators corresponding to the tuning parameter $\beta \in \{0, 0.1, 0.2, 0.4, 0.6, 0.8, 1\}$. The reduction of each one of the parameters of the outlying cell, denoted by $\tilde{\alpha}_0$ or $\tilde{\alpha}_1$ ($\alpha_0 \geq \tilde{\alpha}_0$ or $\alpha_1 \geq \tilde{\alpha}_1$) increases the mean of its lifetime distribution function in (1). The results obtained by decreasing parameter α_0 are shown in Figure 3, while the results obtained by decreasing parameter α_1 are shown in Figure 3. In all the cases, we can see how the MLEs and the MDPDEs with small values of tuning parameter β present the smallest RMSEs for weak outliers, i.e., when $\tilde{\alpha}_0$ is close to α_0 ($1 - \tilde{\alpha}_0/\alpha_0$ is close to 0) or $\tilde{\alpha}_1$ is close to α_1 ($1 - \tilde{\alpha}_1/\alpha_1$ is close to 0). On the other hand, large values of tuning parameter β make the MDPDEs to present the smallest RMSEs, for medium and strong outliers, i.e., when $\tilde{\alpha}_0$ is not close to α_0 ($1 - \tilde{\alpha}_0/\alpha_0$ is not close to 0) or $\tilde{\alpha}_1$ is not close to α_1 ($1 - \tilde{\alpha}_1/\alpha_1$ is not close to 0). Therefore, the MLE of (α_0, α_1) is very efficient when there are no outliers, but highly non-robust when there are outliers. On the other hand, the MDPDEs with moderate values of the tuning parameter β exhibit a little loss of efficiency without outliers, but at the same time possess a considerable improvement in robustness in the presence of outliers. Actually, these values of the tuning parameter β are the most appropriate ones for the estimators of the parameters in the one-shot device model according to robustness theory: To improve in a considerable way the robustness of the estimators, a small amount of efficiency needs to be compromised.

B. The Z-Type Tests Based on MDPDEs

We will study the performance, with respect to robustness, through simulation of the one-shot device model defined in Section II with the same values of I, J, t, w of the example of [1] given in Table I and for the same tuning parameter, β, as in Section VII-A. We are interested in testing the null hypothesis $H_0 : \alpha_1 = 0.05$ against the alternative

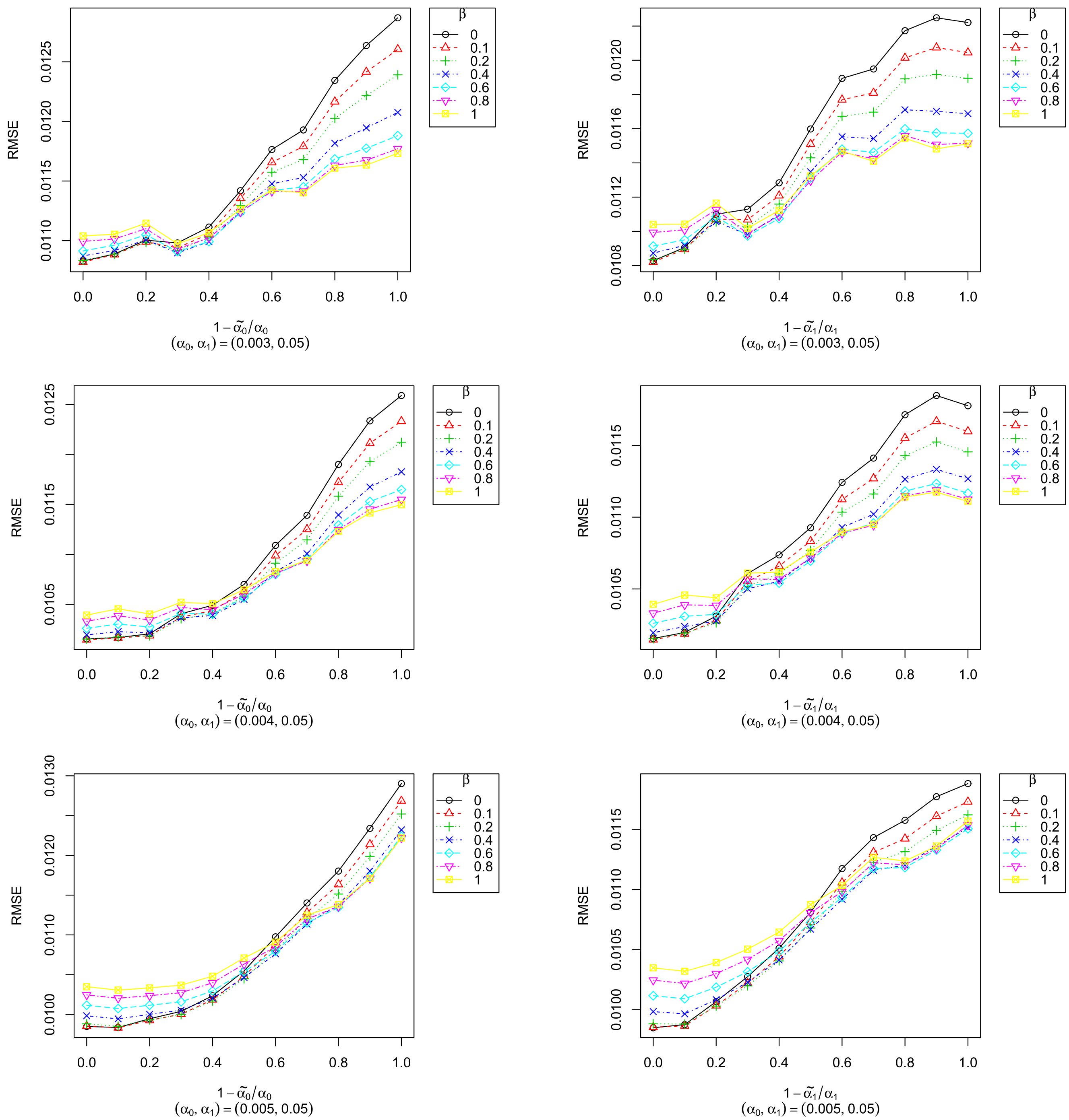

Fig. 3. RMSEs of MDPDEs for $\boldsymbol{\alpha}$. With an α_0-contaminated outlying cell in the data (left) or with an α_1-contaminated outlying cell in the data.

$H_1 : \alpha_1 \neq 0.05$, through the Z-type test statistics based on MDPDEs. Under the null hypothesis, we consider as true parameters $(\alpha_0, \alpha_1) = (0.004, 0.05)$, while under the alternative we consider as true parameters $(\alpha_0, \alpha_1) = (0.004, 0.02)$. In Figure 2, we present the empirical significance level (measured as the proportions of test statistics exceeding in absolute value the standard normal quantile critical value) based on $10,000$ replications. The empirical power (obtained in a similar manner) is also presented in the right hand side of Figure 2. Notice that, in all the cases, the observed levels are quite close to the nominal level of 0.05. The empirical power is similar for the different values of the tuning parameters β, a bit lower for large values of β, and closer to one as K or the sample size ($n = IJK$) increases.

To evaluate the robustness of the level and the power of the Z-type tests based on MDPDEs with an outlier placed on the first-row first-column cell, we perform the simulation for the same test and the same true values for the null and alternative hypotheses, in two different scenarios depending on the way the outlying cell is considered. In the first scenario, we keep α_1 the same and modify the true value of α_0 to be $\tilde{\alpha}_0 \leq \alpha_0$, and in the second one, we keep α_0 the same and modify the

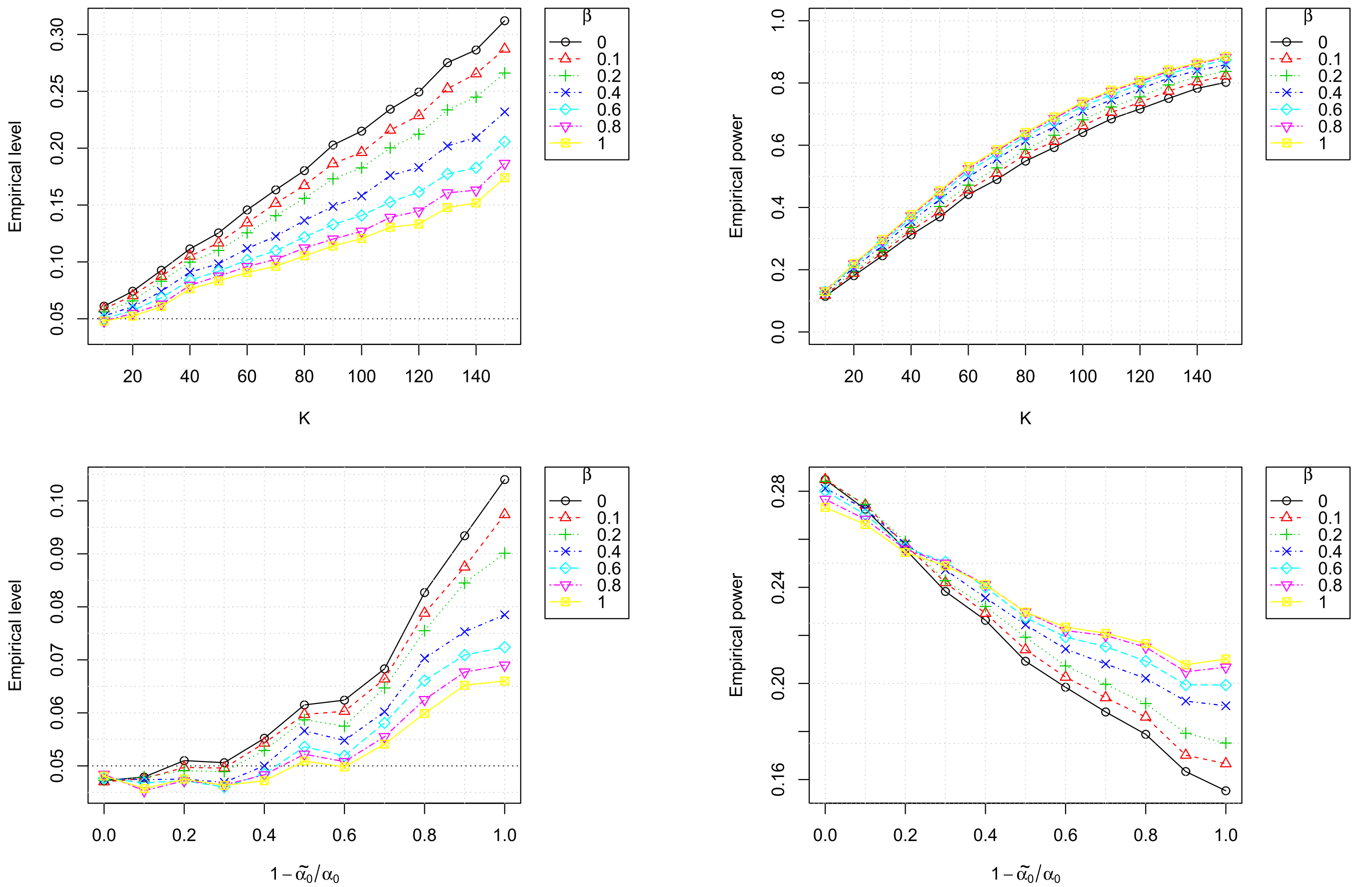

Fig. 4. Simulated levels (left) and powers (right) with an α_0-contaminated outlying cell in the data.

true value of α_1 to be $\tilde{\alpha}_1 \leq \alpha_1$. Both cases have been analyzed for different values of K and decreasing $\tilde{\alpha}_0$ in the first scenario (increasing $1-\tilde{\alpha}_0/\alpha_0$) or decreasing $\tilde{\alpha}_1$ in the second scenario (increasing $1-\tilde{\alpha}_1/\alpha_1$).

The results for the first scenario are presented in Figure 4. The empirical level for the one-shot device model with K from 10 to 150, true value $(\alpha_0, \alpha_1) = (0.004, 0.05)$ and $\tilde{\alpha}_0 = 0.001$ for the outlying cell is presented on the top left panel. Similarly, the empirical power for the one-shot device model with K from 10 to 150, true parameter $(\alpha_0, \alpha_1) = (0.004, 0.02)$ and $\tilde{\alpha}_0 = 0.001$ for the outlying cell is presented on the right top panel. In addition, the empirical level for the one-shot device model with $1-\tilde{\alpha}_0/\alpha_0$ from 0 to 1 for the outlying cell and true value $(\alpha_0, \alpha_1) = (0.004, 0.05)$ and $K = 20$ is presented on the bottom left panel. Similarly, the empirical power for the one-shot device model with $1-\tilde{\alpha}_0/\alpha_0$ from 0 to 1 for the outlying cell and true value and true parameter $(\alpha_0, \alpha_1) = (0.004, 0.02)$ is presented on the bottom right panel.

Notice that the outlying cell represents 1/9 of the total observations in the last plots. For large values of K (very large sample sizes, since $n = 9K$), there is a large inflation in the empirical level and shrinkage of the empirical power, but for the Z-type test statistic based on the MDPDEs with large values of the tuning parameter β, the effect of the outlying cell is weaker in comparison to those of smaller values of β, including the MLEs ($\beta = 0$). If $\tilde{\alpha}_0$ is separated from α_0 ($1-\tilde{\alpha}_0/\alpha_0$ increases from 0 to 1), the empirical level of the Z-type test statistics based on the MDPDEs is not stable around the nominal level, but being closer as the tuning parameter β becomes larger. If $\tilde{\alpha}_0$ is separated from α_0 ($1-\tilde{\alpha}_0/\alpha_0$ increases from 0 to 1), the empirical power of the Z-type test statistics based on the MDPDEs decreases, but being more slowly as the tuning parameter β becomes larger.

Figure 5 presents the results for the second scenario, in which $\tilde{\alpha}_1 = 0.01$ for the outlying cell on the top left panel and $\tilde{\alpha}_1 = -0.01$ for the outlying cell on the top right panel. Even though the outliers are, in the current scenario, slightly more pronunced with respect to the previous scenario, in general terms, we arrive at the same conclusions as in the previous scenario.

The results of the tests statistics presented here show again the poor behavior in robustness of the Z-type tests based on the MLEs of the parameters of the one-shot device model. Furthermore, the robustness properties of the Z-type test statistics based on the MDPDEs with large values of the tuning parameter β are often better as they maintain both level and power in a stable manner. Moreover, the comments made at the end of Section VII-A for the MDPDEs regarding moderate values of the tuning parameter β are valid for the Z-type test statistics based on the MDPDEs as well.

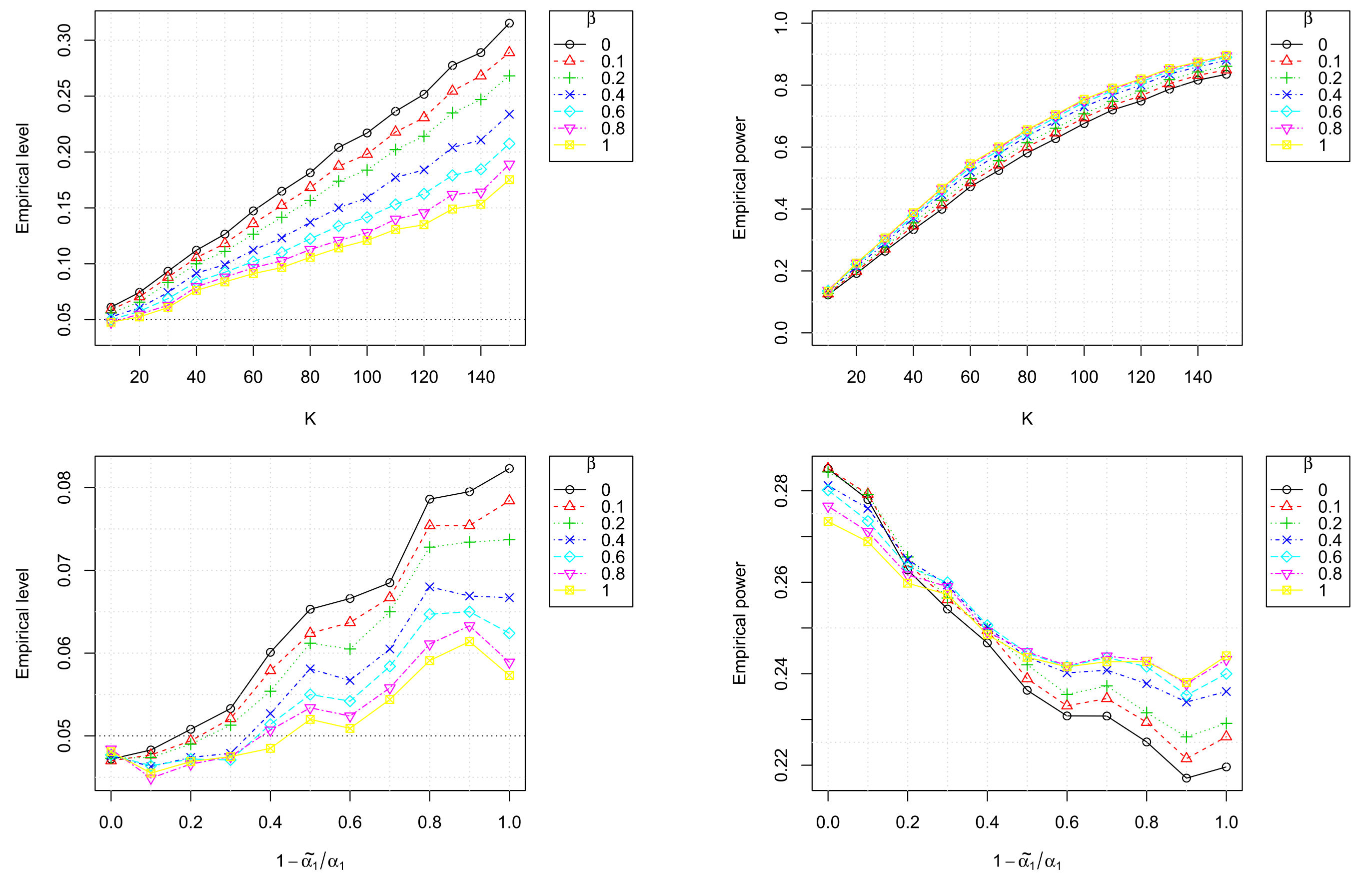

Fig. 5. Simulated levels (left) and powers (right) with an α_1-contaminated outlying cell in the data.

VIII. On the Choice of the Tuning Parameter

Throughout the previous sections, we have noted that the robustness of the proposed MDPDE seems to increase with increasing β; but, their pure data efficiency decrease slightly. From the results of our simulation study, a moderately large value of β is expected to provide the best trade-off for possibly contaminated data. Although a possible ad hoc choice of β may work quite well in practice, when working with real data, a data-driven choice of β would be better and convenient.

An useful procedure of the data-based selection of β for the MDPDE was proposed by Warwick and Jones [17]. It consists of minimizing the estimated mean squared error, an approach that requires pilot estimation of model parameters. We can adopt a similar approach to obtain a suitable data-driven β in our model. In this approach, we minimize an estimate of the asymptotic MSE of the MDPDE $\widehat{\boldsymbol{\alpha}}_\beta$ given by

$$\widehat{MSE}(\beta) = (\boldsymbol{\alpha}_\beta - \boldsymbol{\alpha}^*)^T(\boldsymbol{\alpha}_\beta - \boldsymbol{\alpha}^*) + \frac{1}{K}\text{trace}\left\{\overline{\boldsymbol{J}}_\beta^{-1}(\boldsymbol{\alpha}_\beta)\overline{\boldsymbol{K}}_\beta(\boldsymbol{\alpha}_\beta)\overline{\boldsymbol{J}}_\beta^{-1}(\boldsymbol{\alpha}_\beta)\right\}$$

However, as pointed out by [18], when dealing with the robustness issue, the estimation of the variance component should not assume the model to be true. So, following the general formulation of [11], model robust estimates of $\overline{\boldsymbol{J}}_\beta$ and $\overline{\boldsymbol{K}}_\beta$ can be obtained as

$$\begin{aligned}\widehat{\overline{\boldsymbol{J}}}_\beta^* &= (\beta+1)\overline{\boldsymbol{J}}_\beta(\boldsymbol{\alpha}) \\ &+ \sum_{i=1}^{I}\sum_{j=1}^{J}\left[F(t_j|\lambda_{w_i}(\boldsymbol{\alpha})) - \frac{n_{ij}}{K}\right] \\ &\times\left\{\boldsymbol{C}_1^{(ij)}(\boldsymbol{\alpha})\left[F^{\beta-1}(t_j|\lambda_{w_i}(\boldsymbol{\alpha})) + (1-F(t_j|\lambda_{w_i}(\boldsymbol{\alpha})))^{\beta-1}\right]\right. \\ &\quad\left.-\boldsymbol{C}_2^{(ij)}(\boldsymbol{\alpha})\left[F^{\beta-2}(t_j|\lambda_{w_i}(\boldsymbol{\alpha})) - (1-F(t_j|\lambda_{w_i}(\boldsymbol{\alpha})))^{\beta-2}\right]\right\} \\ &- \beta\sum_{i=1}^{I}\boldsymbol{\Delta}^{(i)}(\boldsymbol{\alpha})\sum_{j=1}^{J}t_j^2 f^2(t_j|\lambda_{w_i}(\boldsymbol{\alpha})) \\ &\times\left[\frac{n_{ij}}{K}F^{\beta-2}(t_j|\lambda_{w_i}(\boldsymbol{\alpha})) + \frac{K-n_{ij}}{K}(1-F(t_j|\lambda_{w_i}(\boldsymbol{\alpha})))^{\beta-2}\right],\end{aligned} \tag{23}$$

$$\begin{aligned}\widehat{\overline{\boldsymbol{K}}}_\beta^* &= \sum_{i=1}^{I}\boldsymbol{\Delta}^{(i)}(\boldsymbol{\alpha})\sum_{j=1}^{J}t_j^2 f^2(t_j|\lambda_{w_i}(\boldsymbol{\alpha})) \\ &\times\left\{\left[\frac{n_{ij}}{K}F^{2\beta-2}(t_j|\lambda_{w_i}(\boldsymbol{\alpha})) + \frac{K-n_{ij}}{K}(1-F(t_j|\lambda_{w_i}(\boldsymbol{\alpha})))^{2\beta-2}\right]\right. \\ &\left.-\left[\frac{n_{ij}}{K}F^{\beta-1}(t_j|\lambda_{w_i}(\boldsymbol{\alpha})) - \frac{K-n_{ij}}{K}(1-F(t_j|\lambda_{w_i}(\boldsymbol{\alpha})))^{\beta-1}\right]^2\right\},\end{aligned} \tag{24}$$

where

$$\boldsymbol{\Delta}^{(i)}(\boldsymbol{\alpha}) = \begin{pmatrix}\frac{1}{\alpha_0^2} & \frac{w_i}{\alpha_0} \\ \frac{w_i}{\alpha_0} & w_i^2\end{pmatrix},$$

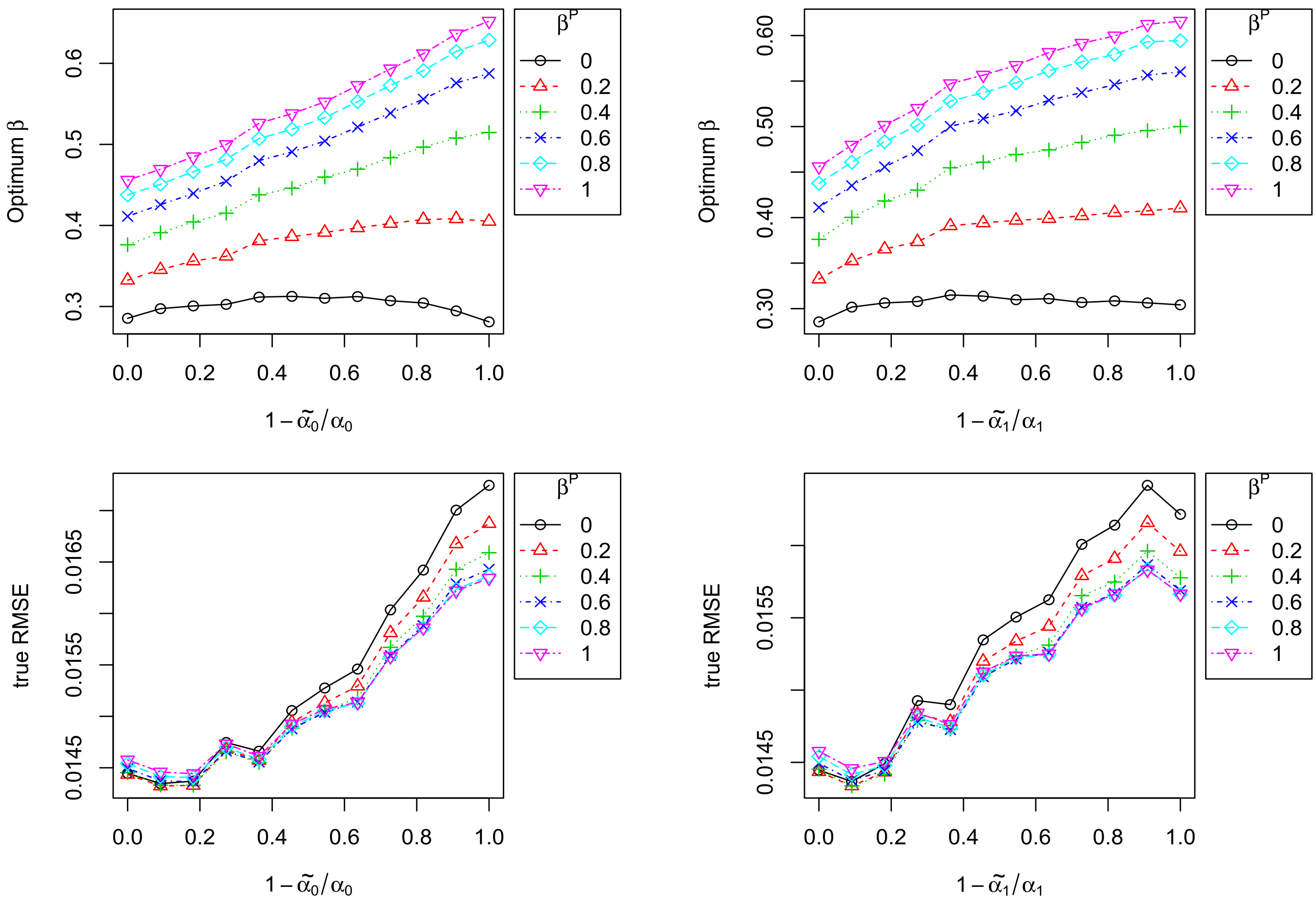

Fig. 6. Average optimal values of β for different values of the pilot estimators (above) and their corresponding RMSEs.

TABLE VII

OPTIMAL VALUES OF β FOR DIFFERENT DATASETS, AND THEIR CORRESPONDING ESTIMATED PARAMETERS

Example	β_{opt}	$\widehat{\alpha}_0$	$\widehat{\alpha}_1$	-	-	-	-	-
Balakrishnan and Ling	0.62	0.0049	0.04696	-	-	-	-	-
Example	β_{opt}	$\widehat{\alpha}_{10}$	$\widehat{\alpha}_{11}$	β_{opt}	$\widehat{\alpha}_{20}$	$\widehat{\alpha}_{21}$	$\widehat{E}_{w=0}(T)$	$\widehat{E}_{w=1}(T)$
ED01 Data	0.24	0.00711	0.00998	0	0.00216	0.27620	107.827	99.681
Example	β_{opt}	$\widehat{\alpha}_{10}$	$\widehat{\alpha}_{11}$	β_{opt}	$\widehat{\alpha}_{20}$	$\widehat{\alpha}_{21}$	$\widehat{E}_{w=1}(T)$	$\widehat{E}_{w=2}(T)$
Benzidine Data	0.30	0.00143	0.85118	0	0.00029	2.41598	578.560	150.859

$$\boldsymbol{C}_1^{(ij)}(\boldsymbol{\alpha}) = \begin{pmatrix} -\frac{1}{\alpha_0^2} & 0 \\ 0 & 0 \end{pmatrix} t_j^2 f(t_j|\lambda_{w_i}(\boldsymbol{\alpha})) + \boldsymbol{\Delta}^{(i)}(\boldsymbol{\alpha}) t_j f(t_j|\lambda_{w_i}(\boldsymbol{\alpha})) F(t_j|\lambda_{w_i}(\boldsymbol{\alpha})), \quad (25)$$

$$\boldsymbol{C}_2^{(ij)}(\boldsymbol{\alpha}) = \boldsymbol{\Delta}^{(i)}(\boldsymbol{\alpha}) t_j^2 f^2(t_j|\lambda_{w_i}(\boldsymbol{\alpha})). \quad (26)$$

Reference [17] suggested to use a suitable pilot estimator $\boldsymbol{\alpha}^P$ in place of $\boldsymbol{\alpha}^*$. The choice of $\boldsymbol{\alpha}^P$ will be empirically discussed, as the overall procedure depends on this choice.

Let us reconsider the previous simulation study with $(\alpha_0, \alpha_1) = (0.004, 0.05)$, but now we perform the selection of β following the above proposal for each iteration with different possible pilot estimators. Let us consider as potential pilot parameters $\beta^P = \{0, 0.3, 0.6, 0.9\}$. The selection of β is done through a grid search of $[0, 1]$ with spacing 0.01 and 10,000 samples. In Figure 6, we show the simulated true RMSEs for this scenario and the average optimal values of β for this same scenario. We can observe how the use of pilot estimators leads us to different optimal values of β, but, in general cases, optimal values of β are higher when a higher degree of contamination is considered, as expected. It seems that the best trade-off between the efficiency in pure data and the robustness under contaminated data is provided by the pilot choice $\beta^P = 0.4$ and so we suggest it as our pilot estimator.

We now apply the corresponding procedure to our previous data sets. Optimal β and corresponding $\widehat{\boldsymbol{\alpha}}$ are presented in Table VII. Here, MLE is not always the best choice, which justifies the need for MDPDEs as an alternative to classical MLE.

Is is also important to notice that this ad-hoc choice of β does not depend on the results of data analysis and expert knowledge. In this sense, and with respect to the ED01 Data, we see that, once the optimal values of the parameters are obtained, expected lifetimes in control group are seen to be higher than in the group to which the carcinogen is applied, which is a result that is consistent in the context of the experiment studied.

IX. Concluding Remarks

In this paper, we have introduced and studied the MDPDEs for one-shot device testing with an accelerating factor of temperature. Based on these estimators, we have also introduced a family of Wald-type test statistics. Since the MLE is a particular estimator in the family of MDPDEs developed here, the classical Wald test is also taken into account for comparison. The results obtained in the simulation study suggest that some MDPDEs are considerably better for the estimation of the model parameters when outliers are present in the data and at the same time not incurring much loss of efficiency when outliers are not present. Similar results are obtained for some Wald-type test statistics in terms of stability with respect to level and power. These proposed estimators also give a more meaningful result in the case of ED01 tumorigenicity experiment data than the MLEs.

Appendix
Proofs of Results

A. Proof of Theorem 1:

We have

$$\begin{aligned} d_{KL}(\widehat{\boldsymbol{p}}, \boldsymbol{p}(\boldsymbol{\alpha})) &= \frac{1}{IJK}\left(s - \sum_{i=1}^{I}\sum_{j=1}^{J}\log\left(F(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))\right)^{n_{ij}}\right. \\ &\quad \left. + \sum_{i=1}^{I}\sum_{j=1}^{J}\log\left((1-F(t_j|\lambda_{w_i}(\boldsymbol{\alpha})))\right)^{K-n_{ij}}\right) \\ &= \frac{1}{IJK}\left(s - \log\prod_{i=1}^{I}\prod_{j=1}^{J}F(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))^{n_{ij}}\right. \\ &\quad \left. \times\left(1-F(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))\right)^{K-n_{ij}}\right) \\ &= \frac{1}{IJK}\left(s - \log\mathcal{L}(\boldsymbol{\alpha}\,|K, \boldsymbol{n}, \boldsymbol{t}, \boldsymbol{w})\right), \end{aligned}$$

with

$$s = \sum_{i=1}^{I}\sum_{j=1}^{J} n_{ij}\log\frac{n_{ij}}{K} + \sum_{i=1}^{I}\sum_{j=1}^{J}(K-n_{ij})\log\frac{K-n_{ij}}{K},$$

as required.

B. Proof of Theorem 4:

The relationship between (6) and $d_\beta(\widehat{\boldsymbol{p}}, \boldsymbol{p}(\boldsymbol{\alpha}))$ defined in (7) is given by

$$\begin{aligned} &\frac{1}{IJ}\sum_{i=1}^{I}\sum_{j=1}^{J}\left\{\pi_{ij}^{\beta+1}(\boldsymbol{\alpha}) + (1-\pi_{ij}(\boldsymbol{\alpha}))^{\beta+1}\right. \\ &\left. -\frac{1+\beta}{\beta}\frac{n_{ij}}{K}\pi_{ij}^{\beta}(\boldsymbol{\alpha}) - \frac{1+\beta}{\beta}\frac{K-n_{ij}}{K}(1-\pi_{ij}(\boldsymbol{\alpha}))^{\beta}\right\} \\ &= (IJ)^{\beta+1}d_\beta(\widehat{\boldsymbol{p}}, \boldsymbol{p}(\boldsymbol{\alpha})) + c, \end{aligned}$$

where c is a constant not dependent on $\boldsymbol{\alpha}$, and so $\widehat{\boldsymbol{\alpha}}_\beta$ is the same for both cases. Hence, the result.

C. Proof of Theorem 5:

We have

$$\begin{aligned} \frac{\partial F(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))}{\partial\alpha_0} &= \exp\left\{-\alpha_0\exp(\alpha_1 w_i)\,t_j\right\}\exp\{\alpha_1 w_i\}\,t_j \\ &= f(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))\frac{t_j}{\alpha_0} \end{aligned} \tag{27}$$

and

$$\begin{aligned} \frac{\partial F(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))}{\partial\alpha_1} &= \exp\left\{-\alpha_0\exp(\alpha_1 w_i)\,t_j\right\} \\ &\quad\times\exp\{\alpha_1 w_i\}\,\alpha_0 t_j w_i \\ &= f(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))t_j w_i. \end{aligned} \tag{28}$$

We denote

$$d_\beta(\widehat{\boldsymbol{p}}, \boldsymbol{p}(\boldsymbol{\alpha})) = \mathcal{T}_{1,\beta}(\boldsymbol{\alpha}) + \mathcal{T}_{2,\beta}(\boldsymbol{\alpha}),$$

where $\mathcal{T}_{1,\beta}(\boldsymbol{\alpha})$ and $\mathcal{T}_{2,\beta}(\boldsymbol{\alpha})$ are as follows, for $\beta > 0$:

$$\begin{aligned} \mathcal{T}_{1,\beta}(\boldsymbol{\alpha}) = \sum_{i=1}^{I}\sum_{j=1}^{J}&\left\{\left(\frac{F\left(t_j|\lambda_{w_i}(\boldsymbol{\alpha})\right)}{IJ}\right)^{1+\beta}\right. \\ &-(1+\tfrac{1}{\beta})\left(\frac{F\left(t_j|\lambda_{w_i}(\boldsymbol{\alpha})\right)}{IJ}\right)^{\beta}\frac{n_{ij}}{IJK} \\ &\left.+\tfrac{1}{\beta}\left(\frac{n_{ij}}{IJK}\right)^{1+\beta}\right\} \end{aligned} \tag{29}$$

$$\begin{aligned} \mathcal{T}_{2,\beta}(\boldsymbol{\alpha}) = \sum_{i=1}^{I}\sum_{j=1}^{J}&\left\{\left(\frac{1-F\left(t_j|\lambda_{w_i}(\boldsymbol{\alpha})\right)}{IJ}\right)^{1+\beta}\right. \\ &-(1+\tfrac{1}{\beta})\left(\frac{1-F\left(t_j|\lambda_{w_i}(\boldsymbol{\alpha})\right)}{IJ}\right)^{\beta}\frac{K-n_{ij}}{IJK} \\ &\left.+\tfrac{1}{\beta}\left(\frac{K-n_{ij}}{IJK}\right)^{1+\beta}\right\} \end{aligned} \tag{30}$$

Based on (27), we have

$$\begin{aligned} \frac{\partial\mathcal{T}_{1,\beta}(\boldsymbol{\alpha})}{\partial\alpha_0} &= \frac{\beta+1}{(IJ)^{\beta+1}}\sum_{i=1}^{I}\sum_{j=1}^{J}\left(F(t_j|\lambda_{w_i}(\boldsymbol{\alpha})) - \frac{n_{ij}}{K}\right) \\ &\quad\times f(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))\frac{t_j}{\alpha_0}F^{\beta-1}(t_j|\lambda_{w_i}(\boldsymbol{\alpha})) \end{aligned}$$

and

$$\begin{aligned} \frac{\partial\mathcal{T}_{2,\beta}(\boldsymbol{\alpha})}{\partial\alpha_0} &= \frac{\beta+1}{(IJ)^{\beta+1}}\sum_{i=1}^{I}\sum_{j=1}^{J}\left(F(t_j|\lambda_{w_i}(\boldsymbol{\alpha})) - \frac{n_{ij}}{K}\right) \\ &\quad\times f(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))\frac{t_j}{\alpha_0}\left(1-F(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))\right)^{\beta-1}. \end{aligned}$$

On the other hand, by (28), we have

$$\begin{aligned} \frac{\partial\mathcal{T}_{1,\beta}(\boldsymbol{\alpha})}{\partial\alpha_1} &= \frac{\beta+1}{(IJ)^{\beta+1}}\sum_{i=1}^{I}\sum_{j=1}^{J}\left(F(t_j|\lambda_{w_i}(\boldsymbol{\alpha})) - \frac{n_{ij}}{K}\right) \\ &\quad\times f(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))t_j w_i F^{\beta-1}(t_j|\lambda_{w_i}(\boldsymbol{\alpha})) \end{aligned}$$

and

$$\frac{\partial \mathcal{T}_{2,\beta}(\boldsymbol{\alpha})}{\partial \alpha_1} = \frac{\beta+1}{(IJ)^{\beta+1}} \sum_{i=1}^{I}\sum_{j=1}^{J} \left(F(t_j|\lambda_{w_i}(\boldsymbol{\alpha})) - \frac{n_{ij}}{K} \right) \times f(t_j|\lambda_{w_i}(\boldsymbol{\alpha})) t_j w_i \left(1 - F(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))\right)^{\beta-1}.$$

Finally, the system of equations is given by

$$\frac{(IJ)^{\beta+1}}{\beta+1}\left(\frac{\partial \mathcal{T}_{1,\beta}(\boldsymbol{\alpha})}{\partial \alpha_0} + \frac{\partial \mathcal{T}_{2,\beta}(\boldsymbol{\alpha})}{\partial \alpha_0} \right) = 0,$$
$$\frac{(IJ)^{\beta+1}}{\beta+1}\left(\frac{\partial \mathcal{T}_{1,\beta}(\boldsymbol{\alpha})}{\partial \alpha_1} + \frac{\partial \mathcal{T}_{2,\beta}(\boldsymbol{\alpha})}{\partial \alpha_1} \right) = 0.$$

If we consider $\beta = 0$ in (9) and (10), we get the system needed to solve for getting the MLE. Hence, the previous system of equations is valid not only for tuning parameters $\beta > 0$, but also for $\beta = 0$.

D. Proof of Theorem 6:

Based on [11] and also on Definition 3, we have

$$\sqrt{IJK}\left(\widehat{\boldsymbol{\alpha}}_\beta - \boldsymbol{\alpha}_0\right) \xrightarrow[IJK\to\infty]{\mathcal{L}} \mathcal{N}\left(\mathbf{0}, \boldsymbol{J}_\beta^{-1}(\boldsymbol{\alpha}_0)\boldsymbol{K}_\beta(\boldsymbol{\alpha}_0)\boldsymbol{J}_\beta^{-1}(\boldsymbol{\alpha}_0)\right),$$

where

$$\begin{aligned}
\boldsymbol{J}_\beta(\boldsymbol{\alpha}) &= \frac{1}{IJK}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} \boldsymbol{J}_{ij,\beta}(\boldsymbol{\alpha}) \\
&= \frac{1}{IJ}\sum_{i=1}^{I}\sum_{j=1}^{J} \boldsymbol{J}_{ij,\beta}(\boldsymbol{\alpha}), \\
\boldsymbol{J}_{ij,\beta}(\boldsymbol{\alpha}) &= \boldsymbol{u}_{ij}(\boldsymbol{\alpha})\boldsymbol{u}_{ij}^T(\boldsymbol{\alpha}) F^{\beta+1}(t_j|\lambda_{w_i}(\boldsymbol{\alpha})) \\
&\quad + \boldsymbol{v}_{ij}(\boldsymbol{\alpha})\boldsymbol{v}_{ij}^T(\boldsymbol{\alpha})(1 - F(t_j|\lambda_{w_i}(\boldsymbol{\alpha})))^{\beta+1} \\
&= t_j^2 f^2(t_j|\lambda_{w_i}(\boldsymbol{\alpha})) \begin{pmatrix} \frac{1}{\alpha_0^2} & \frac{w_i}{\alpha_0} \\ \frac{w_i}{\alpha_0} & w_i^2 \end{pmatrix} \\
&\quad \times \Big[F^{\beta-1}(t_j|\lambda_{w_i}(\boldsymbol{\alpha})) \\
&\quad + (1 - F(t_j|\lambda_{w_i}(\boldsymbol{\alpha})))^{\beta-1} \Big], \\
\boldsymbol{u}_{ij}(\boldsymbol{\alpha}) &= \frac{\partial \log F(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))}{\partial \boldsymbol{\alpha}} \\
&= \frac{1}{F(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))}\frac{\partial}{\partial \boldsymbol{\alpha}} F(t_j|\lambda_{w_i}(\boldsymbol{\alpha})), \\
\boldsymbol{v}_{ij}(\boldsymbol{\alpha}) &= \frac{\partial \log\left[1 - F(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))\right]}{\partial \boldsymbol{\alpha}} \\
&= \frac{1}{1 - F(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))}\frac{\partial}{\partial \boldsymbol{\alpha}} F(t_j|\lambda_{w_i}(\boldsymbol{\alpha})), \\
\frac{\partial}{\partial \boldsymbol{\alpha}} F(t_j|\lambda_{w_i}(\boldsymbol{\alpha})) &= -\frac{\partial}{\partial \boldsymbol{\alpha}} \exp\left\{-\alpha_0 \exp\{\alpha_1 w_i\} t_j\right\} \\
&= \begin{pmatrix} \frac{1}{\alpha_0} \\ w_i \end{pmatrix} t_j f(t_j|\lambda_{w_i}(\boldsymbol{\alpha})),
\end{aligned}$$

and

$$\begin{aligned}
\boldsymbol{K}_\beta(\boldsymbol{\alpha}) &= \frac{1}{IJK}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} \boldsymbol{K}_{ij,\beta}(\boldsymbol{\alpha}) \\
&= \frac{1}{IJ}\sum_{i=1}^{I}\sum_{j=1}^{J} \boldsymbol{K}_{ij,\beta}(\boldsymbol{\alpha}), \\
\boldsymbol{K}_{ij,\beta}(\boldsymbol{\alpha}) &= \boldsymbol{S}_{ij,\beta}(\boldsymbol{\alpha}) - \boldsymbol{\xi}_{ij,\beta}(\boldsymbol{\alpha})\boldsymbol{\xi}_{ij,\beta}^T(\boldsymbol{\alpha}), \\
\boldsymbol{S}_{ij,\beta}(\boldsymbol{\alpha}) &= \boldsymbol{u}_{ij}(\boldsymbol{\alpha})\boldsymbol{u}_{ij}^T(\boldsymbol{\alpha}) F^{2\beta+1}(t_j|\lambda_{w_i}(\boldsymbol{\alpha})) \\
&\quad + \boldsymbol{v}_{ij}(\boldsymbol{\alpha})\boldsymbol{v}_{ij}^T(\boldsymbol{\alpha})(1 - F(t_j|\lambda_{w_i}(\boldsymbol{\alpha})))^{2\beta+1} \\
&= t_j^2 f^2(t_j|\lambda_{w_i}(\boldsymbol{\alpha})) \begin{pmatrix} \frac{1}{\alpha_0^2} & \frac{w_i}{\alpha_0} \\ \frac{w_i}{\alpha_0} & w_i^2 \end{pmatrix} \Big[F^{2\beta-1}(t_j|\lambda_{w_i}(\boldsymbol{\alpha})) \\
&\quad + (1 - F(t_j|\lambda_{w_i}(\boldsymbol{\alpha})))^{2\beta-1} \Big], \\
\boldsymbol{\xi}_{ij,\beta}(\boldsymbol{\alpha}) &= \boldsymbol{u}_{ij}(\boldsymbol{\alpha}) F^{\beta+1}(t_j|\lambda_{w_i}(\boldsymbol{\alpha})) \\
&\quad + \boldsymbol{v}_{ij}(\boldsymbol{\alpha})(1 - F(t_j|\lambda_{w_i}(\boldsymbol{\alpha})))^{\beta+1} \\
&= \begin{pmatrix} \frac{1}{\alpha_0} \\ w_i \end{pmatrix} t_j f(t_j|\lambda_{w_i}(\boldsymbol{\alpha})) \Big[F^{\beta}(t_j|\lambda_{w_i}(\boldsymbol{\alpha})) \\
&\quad - (1 - F(t_j|\lambda_{w_i}(\boldsymbol{\alpha})))^{\beta} \Big].
\end{aligned}$$

Since I, J are fixed and $IJK \to \infty$, it follows that $K \to \infty$ and

$$\sqrt{K}\left(\widehat{\boldsymbol{\alpha}}_\beta - \boldsymbol{\alpha}_0\right) \xrightarrow[K\to\infty]{\mathcal{L}} \mathcal{N}\left(\mathbf{0}, \bar{\boldsymbol{J}}_\beta^{-1}(\boldsymbol{\alpha}_0)\bar{\boldsymbol{K}}_\beta(\boldsymbol{\alpha}_0)\bar{\boldsymbol{J}}_\beta^{-1}(\boldsymbol{\alpha}_0)\right),$$

where

$$\begin{aligned}
\bar{\boldsymbol{J}}_\beta^{-1}(\boldsymbol{\alpha}_0)\bar{\boldsymbol{K}}_\beta(\boldsymbol{\alpha}_0)\bar{\boldsymbol{J}}_\beta^{-1}(\boldsymbol{\alpha}_0) &= \frac{1}{IJ}\boldsymbol{J}_\beta^{-1}(\boldsymbol{\alpha}_0)\boldsymbol{K}_\beta(\boldsymbol{\alpha}_0)\boldsymbol{J}_\beta^{-1}(\boldsymbol{\alpha}_0), \\
\bar{\boldsymbol{J}}_\beta(\boldsymbol{\alpha}_0) &= (IJ)\boldsymbol{J}_\beta(\boldsymbol{\alpha}_0), \\
\bar{\boldsymbol{K}}_\beta(\boldsymbol{\alpha}_0) &= (IJ)\boldsymbol{K}_\beta(\boldsymbol{\alpha}_0).
\end{aligned}$$

E. Proof of Theorem 7:

The Fisher information matrix for IJK observations is

$$\boldsymbol{I}_{IJK,F}(\boldsymbol{\alpha}) = E\left[-\frac{\partial \boldsymbol{v}^T(\boldsymbol{\alpha}|K, \boldsymbol{n}, \boldsymbol{t}, \boldsymbol{w})}{\partial \boldsymbol{\alpha}} \right],$$

where

$$\boldsymbol{v}(\boldsymbol{\alpha}|K, \boldsymbol{n}, \boldsymbol{t}, \boldsymbol{w}) = \frac{\partial \log \mathcal{L}(\boldsymbol{\alpha}|K, \boldsymbol{n}, \boldsymbol{t}, \boldsymbol{w})}{\partial \boldsymbol{\alpha}}.$$

From (3), we have

$$\begin{aligned}
\boldsymbol{I}_{IJK,F}(\boldsymbol{\alpha}) &= IJK\ E\left[\frac{\partial^2 d_{KL}(\widehat{\boldsymbol{p}}, \boldsymbol{p}(\boldsymbol{\alpha}))}{\partial \boldsymbol{\alpha}\partial \boldsymbol{\alpha}^T} \right] \\
&= IJK\ E\left[\frac{\partial \boldsymbol{u}^T(\boldsymbol{\alpha}|K, \boldsymbol{n}, \boldsymbol{t}, \boldsymbol{w})}{\partial \boldsymbol{\alpha}} \right],
\end{aligned}$$

where

$$\begin{aligned}
\boldsymbol{u}(\boldsymbol{\alpha}|K, \boldsymbol{n}, \boldsymbol{t}, \boldsymbol{w}) &= \frac{\partial d_{KL}(\widehat{\boldsymbol{p}}, \boldsymbol{p}(\boldsymbol{\alpha}))}{\partial \boldsymbol{\alpha}} \\
&= \frac{\partial \mathcal{T}_{1,\beta=0}(\boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}} + \frac{\partial \mathcal{T}_{2,\beta=0}(\boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}}.
\end{aligned}$$

The Fisher information matrix for a single observation, i.e., the Fisher information matrix for the one-shot device model is

$$\boldsymbol{I}_F(\boldsymbol{\alpha}) = \frac{1}{IJK}\boldsymbol{I}_{IJK,F}(\boldsymbol{\alpha}) = E\left[\frac{\partial \boldsymbol{u}^T(\boldsymbol{\alpha}|K, \boldsymbol{n}, \boldsymbol{t}, \boldsymbol{w})}{\partial \boldsymbol{\alpha}} \right].$$

From Result 5, the first and second components of $\boldsymbol{u}(\boldsymbol{\alpha}|K, \boldsymbol{n}, \boldsymbol{t}, \boldsymbol{w})$ are

$$\begin{aligned}
u_1(\boldsymbol{\alpha}|K, \boldsymbol{n}, \boldsymbol{t}, \boldsymbol{w}) &= \frac{\partial \mathcal{T}_{1,\beta=0}(\boldsymbol{\alpha})}{\partial \alpha_0} + \frac{\partial \mathcal{T}_{2,\beta=0}(\boldsymbol{\alpha})}{\partial \alpha_0} \\
&= \frac{1}{IJK}\sum_{i=1}^{I}\sum_{j=1}^{J} \left(K\ F(t_j|\lambda_{w_i}(\boldsymbol{\alpha})) - n_{ij}\right) \\
&\quad \times f(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))\frac{t_j}{\alpha_0}\left[F^{-1}(t_j|\lambda_{w_i}(\boldsymbol{\alpha})) + \left(1 - F(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))\right)^{-1} \right] \\
&= \frac{1}{IJK}\sum_{i=1}^{I}\sum_{j=1}^{J} \frac{K\ F(t_j|\lambda_{w_i}(\boldsymbol{\alpha})) - n_{ij}}{F(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))\left(1 - F(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))\right)} f(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))\frac{t_j}{\alpha_0}
\end{aligned}$$

and

$$u_2(\boldsymbol{\alpha}\,|K,\boldsymbol{n},\boldsymbol{t},\boldsymbol{w}) = \frac{\partial \mathcal{T}_{1,\beta=0}(\boldsymbol{\alpha})}{\partial \alpha_1} + \frac{\partial \mathcal{T}_{2,\beta=0}(\boldsymbol{\alpha})}{\partial \alpha_1}$$
$$= \frac{1}{IJK}\sum_{i=1}^{I}\sum_{j=1}^{J}\left(K\ F(t_j|\lambda_{w_i}(\boldsymbol{\alpha})) - n_{ij}\right)$$
$$\times f(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))t_j w_i\left[F^{-1}(t_j|\lambda_{w_i}(\boldsymbol{\alpha})) + \left(1-F(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))\right)^{-1}\right]$$
$$= \frac{1}{IJK}\sum_{i=1}^{I}\sum_{j=1}^{J}\frac{K\ F(t_j|\lambda_{w_i}(\boldsymbol{\alpha})) - n_{ij}}{F(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))\left(1-F(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))\right)} f(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))t_j w_i,$$

respectively. The $(1, 1)$th term of $\boldsymbol{I}_F(\boldsymbol{\alpha})$ is the expectation of

$$\frac{\partial u_1(\boldsymbol{\alpha}\,|K,\boldsymbol{n},\boldsymbol{t},\boldsymbol{w})}{\partial \alpha_0}$$
$$= \frac{1}{IJK}\sum_{i=1}^{I}\sum_{j=1}^{J}\left\{-\frac{t_j}{\alpha_0^2}\frac{K\ F(t_j|\lambda_{w_i}(\boldsymbol{\alpha})) - n_{ij}}{F(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))\left(1-F(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))\right)} f(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))\right.$$
$$+ \frac{\partial f(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))}{\partial \alpha_0}\frac{t_j}{\alpha_0}\frac{K\ F(t_j|\lambda_{w_i}(\boldsymbol{\alpha})) - n_{ij}}{F(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))\left(1-F(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))\right)}$$
$$\left.+ \frac{\partial}{\partial \alpha_0}\left(\frac{K\ F(t_j|\lambda_{w_i}(\boldsymbol{\alpha})) - n_{ij}}{F(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))\left(1-F(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))\right)}\right)\frac{t_j}{\alpha_0} f(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))\right\}.$$

Since the expectation of the first two summands of $\partial u_1(\boldsymbol{\alpha}\,|K,\boldsymbol{n},\boldsymbol{t},\boldsymbol{w})/\partial \alpha_0$ are zero, the interest is on the expectation of L_{ij}, which is as follows:

$$L_{ij} = \frac{\partial}{\partial \alpha_0}\left(\frac{K\ F(t_j|\lambda_{w_i}(\boldsymbol{\alpha})) - n_{ij}}{F(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))\left(1-F(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))\right)}\right)\frac{t_j}{\alpha_0} f(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))$$
$$= \frac{K\frac{t_j}{\alpha_0} f(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))F(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))\left(1-F(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))\right)}{F(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))^2\left(1-F(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))\right)^2}$$
$$\times \frac{t_j}{\alpha_0} f(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))$$
$$- \frac{\frac{\partial}{\partial \alpha_0}\left[F(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))\left(1-F(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))\right)\right]\left(K\ F(t_j|\lambda_{w_i}(\boldsymbol{\alpha})) - n_{ij}\right)}{F(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))^2\left(1-F(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))\right)^2}$$
$$\times \frac{t_j}{\alpha_0} f(t_j|\lambda_{w_i}(\boldsymbol{\alpha})). \tag{31}$$

The expectation of the second summand of L_{ij} is zero and hence

$$E[L_{ij}] = \frac{K\left(\frac{t_j}{\alpha_0}\right)^2 f^2(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))}{F(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))\left(1-F(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))\right)}.$$

These finally yield the $(1, 1)$th term of $\boldsymbol{I}_F(\boldsymbol{\alpha})$ as

$$E\left[\frac{\partial u_1(\boldsymbol{\alpha}\,|K,\boldsymbol{n},\boldsymbol{t},\boldsymbol{w})}{\partial \alpha_0}\right]$$
$$= \frac{K}{IJK}\sum_{i=1}^{I}\sum_{j=1}^{J}\frac{\left(\frac{t_j}{\alpha_0}\right)^2 f^2(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))}{F(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))\left(1-F(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))\right)}$$
$$= \frac{1}{IJ}\sum_{i=1}^{I}\sum_{j=1}^{J}\frac{\left(\frac{t_j}{\alpha_0}\right)^2 f^2(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))}{F(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))\left(1-F(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))\right)}.$$

The rest of the terms of $\boldsymbol{I}_F(\boldsymbol{\alpha})$ can be obtained in a similar manner. On the other hand, from Theorem 5, sustituting $\beta = 0$ into $\boldsymbol{J}_\beta(\boldsymbol{\alpha}) = \frac{1}{IJ}\bar{\boldsymbol{J}}_\beta(\boldsymbol{\alpha})$ and $\boldsymbol{K}_\beta(\boldsymbol{\alpha}) = \frac{1}{IJ}\bar{\boldsymbol{K}}_\beta(\boldsymbol{\alpha})$, we simply obtain $\boldsymbol{J}_{\beta=0}(\boldsymbol{\alpha}) = \boldsymbol{K}_{\beta=0}(\boldsymbol{\alpha}) = \boldsymbol{I}_F(\boldsymbol{\alpha})$.

F. Proof of Theorem 10:

Let $\boldsymbol{\alpha}_0$ be the true value of parameter $\boldsymbol{\alpha}$. It is clear that under (14), we have

$$\boldsymbol{m}^T\widehat{\boldsymbol{\alpha}}_\beta - d = \boldsymbol{m}^T(\widehat{\boldsymbol{\alpha}}_\beta - \boldsymbol{\alpha}_0)$$

and we know that

$$\sqrt{K}(\widehat{\boldsymbol{\alpha}}_\beta - \boldsymbol{\alpha}_0) \xrightarrow[K\to\infty]{\mathcal{L}} \mathcal{N}(\boldsymbol{0}, \bar{\boldsymbol{J}}_\beta^{-1}(\boldsymbol{\alpha}_0)\bar{\boldsymbol{K}}_\beta(\boldsymbol{\alpha}_0)\bar{\boldsymbol{J}}_\beta^{-1}(\boldsymbol{\alpha}_0)),$$

from which it follows that

$$\sqrt{K}(\boldsymbol{m}^T\widehat{\boldsymbol{\alpha}}_\beta - d) \xrightarrow[K\to\infty]{\mathcal{L}} \mathcal{N}(0, \boldsymbol{m}^T\bar{\boldsymbol{J}}_\beta^{-1}(\boldsymbol{\alpha}_0)\bar{\boldsymbol{K}}_\beta(\boldsymbol{\alpha}_0)\bar{\boldsymbol{J}}_\beta^{-1}(\boldsymbol{\alpha}_0)\boldsymbol{m}).$$

Dividing the left hand side by

$$\sqrt{\boldsymbol{m}^T\bar{\boldsymbol{J}}_\beta^{-1}(\widehat{\boldsymbol{\alpha}}_\beta)\bar{\boldsymbol{K}}_\beta(\widehat{\boldsymbol{\alpha}}_\beta)\bar{\boldsymbol{J}}_\beta^{-1}(\widehat{\boldsymbol{\alpha}}_\beta)\boldsymbol{m}},$$

since $\boldsymbol{m}^T\bar{\boldsymbol{J}}_\beta^{-1}(\widehat{\boldsymbol{\alpha}}_\beta)\bar{\boldsymbol{K}}_\beta(\widehat{\boldsymbol{\alpha}}_\beta)\bar{\boldsymbol{J}}_\beta^{-1}(\widehat{\boldsymbol{\alpha}}_\beta)\boldsymbol{m}$ is a consistent estimator of $\boldsymbol{m}^T\bar{\boldsymbol{J}}_\beta^{-1}(\boldsymbol{\alpha}_0)\bar{\boldsymbol{K}}_\beta(\boldsymbol{\alpha}_0)\bar{\boldsymbol{J}}_\beta^{-1}(\boldsymbol{\alpha}_0)\boldsymbol{m}$, the desired result is obtained.

G. Proof of Theorem 11:

The power function at $\boldsymbol{\alpha}^*$ of $Z_K(\widehat{\boldsymbol{\alpha}}_\beta)$ can be obtained as follows:

$$\pi(\boldsymbol{\alpha}^*) = \Pr\left(\left|Z_K(\widehat{\boldsymbol{\alpha}}_\beta)\right| > z_{\frac{\alpha}{2}}|\boldsymbol{\alpha} = \boldsymbol{\alpha}^*\right) \tag{32}$$
$$= 2\Pr\left(Z_K(\widehat{\boldsymbol{\alpha}}_\beta) > z_{\frac{\alpha}{2}}|\boldsymbol{\alpha} = \boldsymbol{\alpha}^*\right)$$
$$= 2\Pr\left(\sqrt{\frac{K}{\boldsymbol{m}^T\bar{\boldsymbol{J}}_\beta^{-1}(\widehat{\boldsymbol{\alpha}}_\beta)\bar{\boldsymbol{K}}_\beta(\widehat{\boldsymbol{\alpha}}_\beta)\bar{\boldsymbol{J}}_\beta^{-1}(\widehat{\boldsymbol{\alpha}}_\beta)\boldsymbol{m}}}\right.$$
$$\left.\times(\boldsymbol{m}^T\widehat{\boldsymbol{\alpha}}_\beta - d) > z_{\frac{\alpha}{2}}|\boldsymbol{\alpha} = \boldsymbol{\alpha}^*\right)$$
$$= 2\Pr\left(\sqrt{\frac{K}{\boldsymbol{m}^T\bar{\boldsymbol{J}}_\beta^{-1}(\widehat{\boldsymbol{\alpha}}_\beta)\bar{\boldsymbol{K}}_\beta(\widehat{\boldsymbol{\alpha}}_\beta)\bar{\boldsymbol{J}}_\beta^{-1}(\widehat{\boldsymbol{\alpha}}_\beta)\boldsymbol{m}}}\boldsymbol{m}^T(\widehat{\boldsymbol{\alpha}}_\beta - \boldsymbol{\alpha}^*) > \right.$$
$$\left. z_{\frac{\alpha}{2}} - \sqrt{\frac{K}{\boldsymbol{m}^T\bar{\boldsymbol{J}}_\beta^{-1}(\widehat{\boldsymbol{\alpha}}_\beta)\bar{\boldsymbol{K}}_\beta(\widehat{\boldsymbol{\alpha}}_\beta)\bar{\boldsymbol{J}}_\beta^{-1}(\widehat{\boldsymbol{\alpha}}_\beta)\boldsymbol{m}}}(\boldsymbol{m}^T\boldsymbol{\alpha}^* - d)\right).$$

Finally, since $\boldsymbol{m}^T\bar{\boldsymbol{J}}_\beta^{-1}(\widehat{\boldsymbol{\alpha}}_\beta)\bar{\boldsymbol{K}}_\beta(\widehat{\boldsymbol{\alpha}}_\beta)\bar{\boldsymbol{J}}_\beta^{-1}(\widehat{\boldsymbol{\alpha}}_\beta)\boldsymbol{m}$ is a consistent estimator of $\boldsymbol{m}^T\bar{\boldsymbol{J}}_\beta^{-1}(\boldsymbol{\alpha}^*)\bar{\boldsymbol{K}}_\beta(\boldsymbol{\alpha}^*)\bar{\boldsymbol{J}}_\beta^{-1}(\boldsymbol{\alpha}^*)\boldsymbol{m}$ and

$$\boldsymbol{m}^T\sqrt{K}(\widehat{\boldsymbol{\alpha}}_\beta - \boldsymbol{\alpha}^*) \xrightarrow[K\to\infty]{\mathcal{L}} \mathcal{N}(0, \boldsymbol{m}^T\bar{\boldsymbol{J}}_\beta^{-1}(\boldsymbol{\alpha}^*)\bar{\boldsymbol{K}}_\beta(\boldsymbol{\alpha}^*)\bar{\boldsymbol{J}}_\beta^{-1}(\boldsymbol{\alpha}^*)\boldsymbol{m}),$$

the desired result follows.

H. Proof of Equations (23) and (24)

Following [19], we have

$$\widehat{\overline{\boldsymbol{J}}}_\beta^* = (\beta+1)\overline{\boldsymbol{J}}_\beta(\boldsymbol{\alpha}) + \sum_{i=1}^{I}\sum_{j=1}^{J}\left\{\frac{\partial \boldsymbol{u}_{ij}(\boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}}F^{\beta+1}(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))\right.$$
$$\left.+\frac{\partial \boldsymbol{v}_{ij}(\boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}}(1-F(t_j|\lambda_{w_i}(\boldsymbol{\alpha})))^{\beta+1}\right\}$$
$$-\beta\sum_{i=1}^{I}\sum_{j=1}^{J}\left\{\boldsymbol{u}_{ij}(\boldsymbol{\alpha})\boldsymbol{u}_{ij}^T(\boldsymbol{\alpha})\frac{n_{ij}}{K}F^\beta(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))\right.$$
$$\left.+\boldsymbol{v}_{ij}(\boldsymbol{\alpha})\boldsymbol{v}_{ij}^T(\boldsymbol{\alpha})\frac{K-n_{ij}}{K}(1-F(t_j|\lambda_{w_i}(\boldsymbol{\alpha})))^\beta\right\}$$
$$-\sum_{i=1}^{I}\sum_{j=1}^{J}\left\{\frac{\partial \boldsymbol{u}_{ij}(\boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}}\frac{n_{ij}}{K}F^\beta(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))\right.$$
$$\left.+\frac{\partial \boldsymbol{v}_{ij}(\boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}}\frac{K-n_{ij}}{K}(1-F(t_j|\lambda_{w_i}(\boldsymbol{\alpha})))^\beta\right\}$$

and

$$\widehat{\overline{\boldsymbol{K}}}_{\beta}^{*} = \sum_{i=1}^{I}\sum_{j=1}^{J}\Big\{\boldsymbol{u}_{ij}(\boldsymbol{\alpha})\boldsymbol{u}_{ij}^{T}(\boldsymbol{\alpha})\frac{n_{ij}}{K}F^{2\beta}(t_j|\lambda_{w_i}(\boldsymbol{\alpha})) + \boldsymbol{v}_{ij}(\boldsymbol{\alpha})\boldsymbol{v}_{ij}^{T}(\boldsymbol{\alpha})\frac{K-n_{ij}}{K}(1-F(t_j|\lambda_{w_i}(\boldsymbol{\alpha})))^{2\beta}\Big\} - \sum_{i=1}^{I}\sum_{j=1}^{J}\boldsymbol{\xi}_{ij,\beta}^{*}(\boldsymbol{\alpha})\boldsymbol{\xi}_{ij,\beta}^{*T}(\boldsymbol{\alpha})$$

with

$$\boldsymbol{\xi}_{ij,\beta}^{*}(\boldsymbol{\alpha}) = \boldsymbol{u}_{ij}(\boldsymbol{\alpha})\frac{n_{ij}}{K}F^{\beta}(t_j|\lambda_{w_i}(\boldsymbol{\alpha})) + \boldsymbol{v}_{ij}(\boldsymbol{\alpha})\frac{K-n_{ij}}{K}(1-F(t_j|\lambda_{w_i}(\boldsymbol{\alpha})))^{\beta}.$$

The required result follows taking into account that

$$\frac{\partial \boldsymbol{u}_{ij}(\boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}} = \frac{1}{F(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))}\boldsymbol{C}_1^{(ij)}(\boldsymbol{\alpha}) + \frac{1}{F^2(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))}\boldsymbol{C}_2^{(ij)}(\boldsymbol{\alpha})$$

$$\frac{\partial \boldsymbol{v}_{ij}(\boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}} = -\frac{1}{(1-F(t_j|\lambda_{w_i}(\boldsymbol{\alpha})))}\boldsymbol{C}_1^{(ij)}(\boldsymbol{\alpha}) + \frac{1}{(1-F(t_j|\lambda_{w_i}(\boldsymbol{\alpha})))^2}\boldsymbol{C}_2^{(ij)}(\boldsymbol{\alpha}),$$

where $\boldsymbol{C}_1^{(ij)}(\boldsymbol{\alpha})$ and $\boldsymbol{C}_2^{(ij)}(\boldsymbol{\alpha})$ are as given in (25) and (26), respectively.

REFERENCES

[1] N. Balakrishnan and M. H. Ling, "EM algorithm for one-shot device testing under the exponential distribution," *Comput. Statist. Data Anal.*, vol. 56, no. 3, pp. 502–509, Mar. 2012.

[2] T. H. Fan, N. Balakrishnan, and C. C. Chang, "The Bayesian approach for highly reliable electro-explosive devices using one-shot device testing," *J. Statist. Comput. Simul.*, vol. 79, no. 9, pp. 1143–1154, Aug. 2009.

[3] J. Rodrigues, H. Bolfarine, and F. Louzada-Nieto, "Comparing several accelerated life methods," *Commun. Statist.-Theory Methods* vol. 22, no. 8, pp. 2297–2308, 1993.

[4] E. V. Chimitova and N. Balakrishnan, "Goodness-of-fittests for one-shot device testing data," in *Advanced Mathematical and Computational Tools in Metrology and Testing X*, F. Pavese, W. Bremser, A. Chunovkina, N. Fischer and A. B. Forbe, Eds. Singapore: World Scientific, 2015. pp. 124–133.

[5] N. Balakrishnan, H. Y. So, and M. H. Ling, "Bayesian approach for one-shot device testing with exponential lifetimes under competing risks," *IEEE Trans. Rel.*, vol. 65, no. 1, pp. 469–485, Mar. 2016.

[6] L. Pardo, *Statistical Inference Based on Divergence Measures*. Boston, MA, USA: Chapman & Hall, 2006.

[7] A. Basu, H. Shioya, and C. Park, *Statistical Inference: The Minimum Distance Approach*. Boca Raton, FL, USA: CRC Press, 2011.

[8] A. Basu, I. R. Harris, N. L. Hjort, and M. C. Jones "Robust and efficient estimation by minimising a density power divergence," *Biometrika*, vol. 85, no. 3, pp. 549–559, Sep. 1998.

[9] A. Ghosh, A. Mandal, N. Martín, and L. Pardo, "Influence analysis of robust wald-type tests," *J. Multivariate Anal.*, vol. 147, pp. 102–126, May 2016.

[10] A. Basu, A. Mandal, N. Martin, and L. Pardo, "Generalized wald-type tests based on minimum density power divergence estimators," *Statistics*, vol. 50, no. 1, pp. 1–26, 2016.

[11] A. Ghosh and A. Basu, "Robust estimation for independent non-homogeneous observations using density power divergence with applications to linear regression," *Electron. J. Statist.*, vol. 7, pp. 2420–2456, 2013.

[12] S. M. Schennach, "Point estimation with exponentially tilted empirical likelihood," *Ann. Statist.*, vol. 35, no. 2, pp. 634–672, 2007.

[13] N. Balakrishnan, N. Martin, and L. Pardo, "Empirical phi-divergence test statistics for the difference of means of two populations," *AStA Adv. Statist. Anal.*, vol. 101, no. 2, pp. 199–226, Apr. 2017.

[14] A. Felipe, N. Martín, P. Miranda, and L. Pardo, "Testing with exponentially tilted empirical likelihood," *Methodol. Comput. Appl. Probab.*, vol. 20, no. 4, pp. 1319–1358, Dec. 2018.

[15] J. Lindsey and L. A. Ryan, "three-state multiplicative model for rodent tumorigenicity experiments," *J. Roy. Statist. Soc. C, Appl. Statist.*, vol. 42, no. 2, pp. 283–300, 1993.

[16] M. H. Ling, H. Y. So, and N. Balakrishnan, "Likelihood inference under proportional hazards model for one-shot device testing," *IEEE Trans. Rel.*, vol. 65, no. 1, pp. 446–458, Mar. 2016.

[17] J. Warwick, and M. C. Jones, "Choosing a robustness tuning parameter," *J. Statist. Comput. Simul.*, vol. 75, no. 5, pp. 581–588, 2005.

[18] A. Basu, A. Ghosh, A. Mandal, N. Martin, and L. Pardo, "A wald-type test statistic for testing linear hypothesis in logistic regression models based on minimum density power divergence estimator," *Electron. J. Statist.*, vol. 11, no. 2, pp. 2741–2772, 2017.

[19] A. Ghosh, and A. Basu, "Robust estimation for non-homogeneous data and the selection of the optimal tuning parameter: The density power divergence approach," *J. Appl. Statist.*, vol. 42, no. 9, pp. 2056–2072, 2015.

Narayanaswamy Balakrishnan received his B.Sc. and M.Sc. degrees in Statistics from the University of Madras, India, in 1976 and 1978, respectively. He finished his Ph.D. in Statistics from the Indian Institute of Technology, Kanpur, India, in 1981. He is a Distinguished University Professor at McMaster University, Hamilton, ON, Canada. His research interests include distribution theory, ordered data analysis, censoring methodology, reliability, survival analysis, nonparametric inference, and statistical quality control. Prof. Balakrishnan is a Fellow of the American Statistical Association and a Fellow of the Institute of Mathematical Statistics. He is currently the Editor-in-Chief of *Communications in Statistics*.

Elena Castilla is a doctoral student in the Department of Statistics and Operational Research, Complutense University of Madrid (Spain), working on her Ph.D. thesis under the supervision of Professor L. Pardo and N. Martin. Her research interests include reliability analysis and efficiency and robustness for logistic regression models with complex design.

Nirian Martín is an Associate Professor in the department of Financial and Actuarial Economics and Statistics, Complutense University of Madrid (Spain). Her research interests include categorical data analysis, change points, empirical likelihood, generalized linear models, order restricted statistical inference, overdispersion models, regression diagnostics, robust statistics.

Leandro Pardo is a Full Professor in the Department of Statistics and Operational Research, Faculty of Mathematics, Complutense University of Madrid (Spain) since 1993. He holds a B. Sc. and a Ph.D. in Mathematics from Complutense University of Madrid in 1976 and 1980, respectively. He was elected "Distinguished Eugene Lukacs Professor" in the Booling Green University (Booling Green, Ohio) in 2004. Prof. Leandro Pardo was President of the Spanish Society of Statistics and Operations Research (SEIO) berween 2013-2016. He was an editor in Chief for *TEST* between 2005 and 2008.