

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

N. Balakrishnan, E. Castilla, N. Martín, and L.Pardo

September 11, 2018

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 relation to 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 also in terms of robustness.

1 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, 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 only be used 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 left-censored (failure) or right-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 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 have more failures in the experiment. On the other hand, 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, Balakrishnan and Ling (2012) developed the EM algorithm for finding the maximum likelihood estimators of the model parameters. Fan et al. (2009) 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. Rodrigues et al. (1993) presented two approaches based on the likelihood ratio statistics and the posterior Bayes factor for comparing several exponential accelerated life models. Chimitova and Balakrishnan (2015) made a comparison of several goodness-of-fit tests for one-shot device testing.

In Section 2, we present a description of the one-shot device model as well as the maximum likelihood estimators for the model parameters. Section 3 develops the minimum density power divergence estimator as a natural extension of the maximum likelihood estimator, as well as its asymptotic distribution. In Section 4, Z -type test statistics are introduced in order to test some hypotheses about the parameters of the one-shot device model. Some numerical examples are presented in Section 5, with one of them relating to a reliability situation and the other two are real applications to tumorigenicity experiments. In Section 6, an extensive simulation study is presented in order to analyze the robustness of the MDPDEs, as well as the Z -type test introduced earlier. Finally, some concluding remarks are made in Section 7.

2 Model formulation and maximum likelihood estimator

Consider a reliability testing experiment in which at each time, t_j , $j = 1, 2, \dots, J$, K devices are placed in total under temperatures w_i , $i = 1, \dots, I$. Therefore, IJK devices are tested in total at temperatures w_i , $i = 1, \dots, I$, at times t_j , $j = 1, \dots, J$. 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 Balakrishnan and Ling (2012), an example is illustrated, 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 1. In this one-shot device testing experiment, there were in all 48 failures out a total of 90 tested devices.

We shall assume here, in accordance Balakrishnan and Ling (2012), that the true lifetimes T_{ijk} , where $i = 1, 2, \dots, I$, $j = 1, 2, \dots, J$, $k = 1, \dots, K$, are independent and identically distributed exponential random

Table 1: 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

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 , $j = 1, \dots, J$, 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 as

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

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

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

and the density function

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

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

We shall consider the theoretical probability vector $\mathbf{p}(\boldsymbol{\alpha})$ defined by

$$\mathbf{p}(\boldsymbol{\alpha}) = \left(\frac{F(t_1|\lambda_{w_1}(\boldsymbol{\alpha}))}{IJ}, \frac{1-F(t_1|\lambda_{w_1}(\boldsymbol{\alpha}))}{IJ}, \dots, \frac{F(t_J|\lambda_{w_I}(\boldsymbol{\alpha}))}{IJ}, \frac{1-F(t_J|\lambda_{w_I}(\boldsymbol{\alpha}))}{IJ} \right)^T,$$

as well as the observed probability vector

$$\hat{\mathbf{p}} = \left(\frac{n_{11}}{IJK}, \frac{K-n_{11}}{IJK}, \dots, \frac{n_{IJ}}{IJK}, \frac{K-n_{IJ}}{IJK} \right)^T,$$

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

given by

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

It is not difficult to establish the following result.

Theorem 1 *The likelihood function*

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

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

$$d_{KL}(\hat{\mathbf{p}}, \mathbf{p}(\boldsymbol{\alpha})) = \frac{1}{IJK} (s - \log \mathcal{L}(\boldsymbol{\alpha} | K, \mathbf{n}, \mathbf{t}, \mathbf{w})), \quad (3)$$

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

Based on the previous result, we have the following definition for the maximum likelihood estimators of α_0 and α_1 .

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

$$\hat{\boldsymbol{\alpha}} = \arg \min_{\boldsymbol{\alpha} \in \Theta} d_{KL}(\hat{\mathbf{p}}, \mathbf{p}(\boldsymbol{\alpha})), \quad (4)$$

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

3 Minimum density power divergence estimator

Based on expression (4), we can think of defining an estimator minimizing any distance or divergence between the probability vectors $\hat{\mathbf{p}}$ and $\mathbf{p}(\boldsymbol{\alpha})$. There are many different divergence measures (or distances) known in the literature, see, for instance, Pardo (2006) and Basu et al. (2011), and the natural question is if 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 maximum likelihood estimator is a BAN (Best Asymptotically Normal) estimator, but at the same time we know that the maximum likelihood estimator has a very poor behavior, in general, in relation to robustness. It is well-known that a gain in robustness leads to a loss of 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 Basu et al. (1998) has the required properties and has been studied for many different problems until now. For more details, see Ghosh et al. (2016), Basu et al. (2016) and the references therein.

Based on Ghosh and Basu (2013), the MDPDE of α 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} , $i = 1, 2, \dots, I$, $j = 1, 2, \dots, J$, $k = 1, \dots, K$, be a sequence of independent Bernoulli random variables, $y_{ijk} \stackrel{\text{ind}}{\sim} \text{Ber}(\pi_{ij}(\alpha))$, such that $\pi_{ij}(\alpha) = F(t_j | \lambda_{w_i}(\alpha))$ and $n_{ij} = \sum_{k=1}^K y_{ijk}$. The MDPDE of α , with tuning parameter $\beta \geq 0$, is given by

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

where

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

For more details about the interpretation of Definition 3, see formula 2.3 in Ghosh and Basu (2013), in which $\pi_{ij}^{y_{ijk}}(\alpha)(1 - \pi_{ij}(\alpha))^{1-y_{ijk}}$ plays the role of the density in our context. Notice that the expression to be minimized in (5) can be simplified as

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

The following result provides an alternative expression for $\hat{\alpha}_\beta$, 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 $\mathbf{p} = (p_1, \dots, p_M)^T$ and $\mathbf{q} = (q_1, \dots, q_M)^T$, the power density divergence measure between \mathbf{p} and \mathbf{q} , with

tuning parameter $\beta > 0$, is given by

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

and for $\beta = 0$,

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

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

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

and for $\beta = 0$

$$d_{\beta=0}(\hat{\mathbf{p}}, \mathbf{p}(\boldsymbol{\alpha})) = \lim_{\beta \rightarrow 0^+} d_\beta(\hat{\mathbf{p}}, \mathbf{p}(\boldsymbol{\alpha})) = d_{KL}(\hat{\mathbf{p}}, \mathbf{p}(\boldsymbol{\alpha})).$$

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

$$\hat{\boldsymbol{\alpha}}_\beta = \arg \min_{\boldsymbol{\alpha} \in \Theta} d_\beta(\hat{\mathbf{p}}, \mathbf{p}(\boldsymbol{\alpha})), \quad (8)$$

where $d_\beta(\hat{\mathbf{p}}, \mathbf{p}(\boldsymbol{\alpha}))$ 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 $\beta \geq 0$, $\hat{\boldsymbol{\alpha}}_\beta$, can be obtained as the solution of equations (9) and (10).*

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

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

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

Theorem 6 *The asymptotic distribution of the MDPDE $\hat{\alpha}_\beta$ is given by*

$$\sqrt{K}(\hat{\alpha}_\beta - \alpha_0) \xrightarrow{K \rightarrow \infty} \mathcal{N}\left(\mathbf{0}, \bar{\mathbf{J}}_\beta^{-1}(\alpha_0) \bar{\mathbf{K}}_\beta(\alpha_0) \bar{\mathbf{J}}_\beta^{-1}(\alpha_0)\right),$$

where

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

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

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

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

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

the following result relates the asymptotic distribution of $\hat{\alpha}_{\beta=0}$ given previously in terms of $\bar{\mathbf{J}}_{\beta=0}(\alpha_0)$ and $\bar{\mathbf{K}}_{\beta=0}(\alpha_0)$, 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 α , $\hat{\alpha}_{\beta=0}$, is*

$$\sqrt{K}(\hat{\alpha}_\beta - \alpha_0) \xrightarrow{K \rightarrow \infty} \mathcal{N}\left(\mathbf{0}, \frac{1}{IJ} \mathbf{I}_F^{-1}(\alpha_0)\right),$$

where

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

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

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

4 Robust Z-type tests

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

$$H_0: \mathbf{m}^T \alpha = d, \quad (13)$$

where $\mathbf{m}^T = (m_0, m_1)$, it is important to know the asymptotic distribution of the MDPDE of α . In particular, in case we wish to test if the different temperatures do not affect the model of the one-shot devices, $\mathbf{m}^T = (m_0, m_1) = (0, 1)$ and $d = 0$ must be considered. In the following definition, we present Z-type test statistics based on $\hat{\alpha}_\beta$. 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 in Basu et al. (2016) and Ghosh et al. (2016).

Definition 8 Let $\hat{\alpha}_\beta = (\hat{\alpha}_{0,\beta}, \hat{\alpha}_{1,\beta})^T$ be the MDPDE of $\alpha = (\alpha_0, \alpha_1)^T$. The family of Z-type test statistics for testing (13) is given by

$$Z_K(\hat{\alpha}_\beta) = \sqrt{\frac{K}{\mathbf{m}^T \bar{\mathbf{J}}_\beta^{-1}(\hat{\alpha}_\beta) \bar{\mathbf{K}}_\beta(\hat{\alpha}_\beta) \bar{\mathbf{J}}_\beta^{-1}(\hat{\alpha}_\beta) \mathbf{m}}} (\mathbf{m}^T \hat{\alpha}_\beta - d). \quad (14)$$

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

Theorem 9 The asymptotic distribution of Z-type test statistics, $Z_K(\hat{\alpha}_\beta)$, defined in (14), is standard normal.

Based on the previous result, the null hypothesis given in (13) will be rejected, with significance level α , if $|Z_K(\hat{\alpha}_\beta)| > z_{\frac{\alpha}{2}}$, where $z_{\frac{\alpha}{2}}$ is a right hand side quantile of order $\frac{\alpha}{2}$ of a normal distribution. Now we are going to present a result in order to provide an approximation for the test statistic defined in (14).

Theorem 10 Let $\alpha^* \in \Theta$ be the true value of the parameter α so that

$$\hat{\alpha}_\beta \xrightarrow[K \rightarrow \infty]{\mathcal{P}} \alpha^* \in \Theta,$$

and $\mathbf{m}^T \alpha^* \neq d$. Then, the approximated power function of the test statistic in (14) at α^* is given by equation (15), where $\Phi(\cdot)$ is the standard normal distribution function.

$$\pi(\alpha^*) \simeq 2 \left(1 - \Phi \left(z_{\frac{\alpha}{2}} - \sqrt{\frac{K}{\mathbf{m}^T \bar{\mathbf{J}}_\beta^{-1}(\alpha^*) \bar{\mathbf{K}}_\beta(\alpha^*) \bar{\mathbf{J}}_\beta^{-1}(\alpha^*) \mathbf{m}}} (\mathbf{m}^T \alpha^* - d) \right) \right) \quad (15)$$

Remark 11 Based on the previous results, it is possible to establish an explicit expression of the number of devices

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

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

5 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 2, and the other two are real applications to tumorigenicity experiments considered earlier by other authors.

5.1 Example 1 (Reliability experiment)

Based on the example introduced in Section 2, 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 2, apart from the MDPDEs of α , the MDPDEs of the reliability function

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

are also presented at mission times (time points in the future at which we are interested in the reliability of the unit) $t \in \{10, 20, 30\}$, namely $R(10|\lambda_{w_0}(\hat{\alpha}_\beta))$, $R(20|\lambda_{w_0}(\hat{\alpha}_\beta))$, $R(30|\lambda_{w_0}(\hat{\alpha}_\beta))$, as well as the MDPDEs of the mean of the lifetime $T(\lambda_{w_0}(\alpha))$, namely,

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

under the normal operating temperature $w_0 = 25$.

Table 2 shows that the mean lifetime obtained by the maximum likelihood estimator ($\beta = 0$) is greater than that obtained from the alternative MPDPDEs.

5.2 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 Lyndsey and Ryan (1993) 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 tumour free (DNT) and with tumour (DWT), and sacrificed without tumour (SNT) and with tumour (SWT), summarized over three time intervals at 12, 18 and 33 months. A total of 3355 mice were involved in the experiment.

Balakrishnan et al. (2016a) 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 3, while MDPDEs of the model parameters and the corresponding estimates of mean lifetimes are presented in Table 4. Here $w = 0$ refers to control group and $w = 1$ is the

Table 2: MDPDEs of the parameters, the reliability function at times $t \in \{10, 20, 30\}$, and mean of lifetime at normal temperature of $25^\circ C$ in one-shot device testing experiment considered by Balakrishnan and Ling (2012).

β	$\hat{\alpha}_{0,\beta}$	$\hat{\alpha}_{1,\beta}$	$R(10 \lambda_{25}(\hat{\alpha}_\beta))$	$R(20 \lambda_{25}(\hat{\alpha}_\beta))$	$R(30 \lambda_{25}(\hat{\alpha}_\beta))$	$E[T(\lambda_{25}(\hat{\alpha}_\beta))]$
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

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 4, some MDPDEs of α_{11} are seen to be negative. As pointed out in Balakrishnan et al. (2016), 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, 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

Table 3: Number of mice sacrificed ($r = 0$) and died without tumour ($r = 1$) and with tumour ($r = 2$) 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

Table 4: MDPDEs of the parameters and the mean lifetimes of the ED01 experiment

β	$\hat{\alpha}_{10}$	$\hat{\alpha}_{11}$	$E_{w=0}(T_1)$	$E_{w=1}(T_1)$	$\hat{\alpha}_{20}$	$\hat{\alpha}_{21}$	$E_{w=0}(T_2)$	$E_{w=1}(T_2)$	$E_{w=0}(T)$	$E_{w=1}(T)$
0	0.00617	-0.12790	162.233	184.165	0.00236	0.25620	426.425	331.582	117.447	118.299
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

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 maximum likelihood estimator ($\beta = 0$). The simulation study presented in this paper will prove how, in a general case, MDPDEs with these tuning parameters will

also present a better behaviour in terms of robustness.

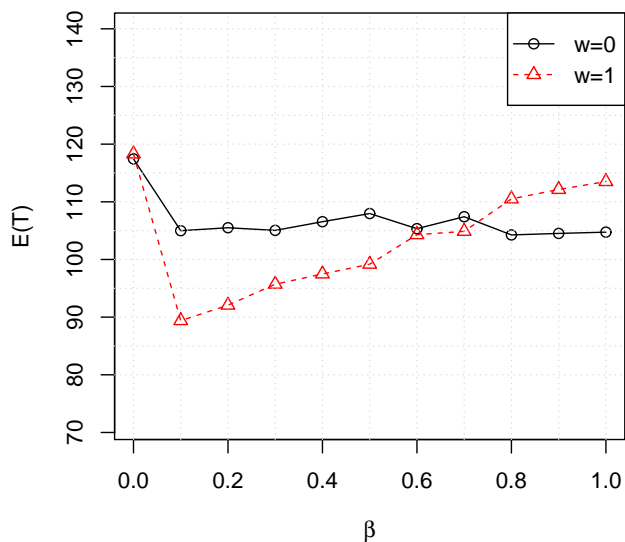


Figure 1: MDPDEs of the mean lifetimes, for different values of the tuning parameter β , from the ED01 experiment

5.3 Example 3 (Benzidine Dihydrochloride Data)

Table 5: Number of mice sacrificed ($r = 0$) and died without tumour ($r = 1$) and with tumour ($r = 2$) 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

The benzidine dihydrochloride experiment was also conducted at the National Center for Toxicological Research to examine the incidence in mice of liver tumours induced by the drug, and studied by Lyndsey and

Ryan (1993) and Balakrishnan et al. (2016b). The inspection times used on the mice were 9.37, 14.07 and 18.7 months. In Table 6, the numbers of mice sacrificed ($r = 0$), died without tumour ($r = 1$) and died with tumour ($r = 2$), 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 6 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.

Table 6: MDPDEs of the parameters and the mean lifetimes of the Benzidine Dihydrochlorid experiment

β	$\hat{\alpha}_{10}$	$\hat{\alpha}_{11}$	$E_{w=0}(T_1)$	$E_{w=1}(T_1)$	$\hat{\alpha}_{20}$	$\hat{\alpha}_{21}$	$E_{w=1}(T_2)$	$E_{w=2}(T_2)$	$E_{w=1}(T)$	$E_{w=2}(T)$
0	0.00074	1.08665	1342.580	452.912	0.00018	2.49999	5472.201	449.190	1081.274	227.233
0.1	0.00093	0.87121	1072.790	448.905	0.00022	2.45781	4459.410	381.825	867.943	208.460
0.2	0.00097	0.84038	1032.863	445.729	0.00024	2.42125	4110.686	365.071	827.690	202.187
0.3	0.00101	0.81098	994.958	442.182	0.00026	2.39084	3867.836	354.112	790.471	196.024
0.4	0.00104	0.78168	958.766	438.766	0.00029	2.34614	3507.841	335.834	750.183	188.387
0.5	0.00109	0.75071	920.459	434.483	0.00029	2.33901	3449.648	332.624	726.525	188.353
0.6	0.00112	0.72656	893.899	432.261	0.00032	2.29717	3168.017	318.521	695.074	181.946
0.7	0.00115	0.70252	866.492	429.206	0.00032	2.28271	3078.308	314.009	678.390	182.918
0.8	0.00118	0.68232	845.366	427.285	0.00033	2.27346	3011.326	310.030	660.973	180.322
0.9	0.00121	0.66476	826.449	425.122	0.00034	2.25372	2902.649	304.799	645.163	178.887
1	0.00124	0.64796	807.541	422.432	0.00035	2.23942	2823.643	300.774	629.593	176.897

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 simulation study in the next section.

6 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 3, as well as the Z -type tests, based on MDPDEs, detailed in Section 4. We pay special attention to the robustness issue. 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 cases, 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 the outliers in the data. The main purpose of this simulation study is to show that inside the family of MDPDEs, developed here, there are estimators with better robust properties than the MLE, and the Z -type tests constructed from them are at the same time more robust than the classical Z -type test, constructed through the MLEs.

6.1 The MDPDEs

In this section, we carry out a simulation study to compare the behavior of some MDPDEs with respect to the MLEs of 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 1, and the simulation experiment proposed by Ling (2012). 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 3a, while the results obtained by decreasing parameter α_1 are shown in Figure 3b. 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 turn 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 a considerable improvement of robustness

with 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.

6.2 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 2 with the same values of I, J, t, w of the example of Balakrishnan and Ling (2012) given in Table 1 and for the same tuning parameter, β , as in Section 6.1. We are interested in testing the null hypothesis $H_0 : \alpha_1 = 0.05$ against the alternative $H_1 : \alpha_1 \neq 0.05$, 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) with 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.

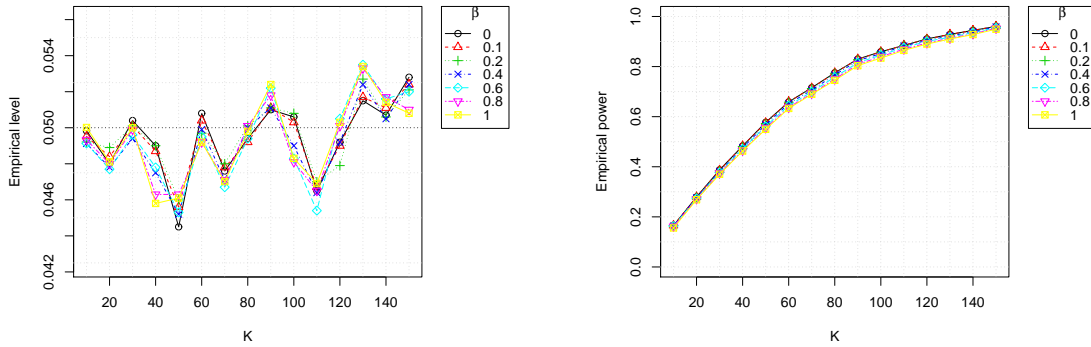


Figure 2: Simulated levels (left) and powers (right) with no outliers in the data.

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 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 left and top 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 left bottom 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 right bottom 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 , included 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, being however 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, being however 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 left top panel and $\tilde{\alpha}_1 = -0.01$ for the outlying cell on the right top panel. Even though the outliers are, in the current scenario, slightly more pronounced with respect to the previous scenario, in general terms, we arrive at the same conclusions as in the previous scenario.

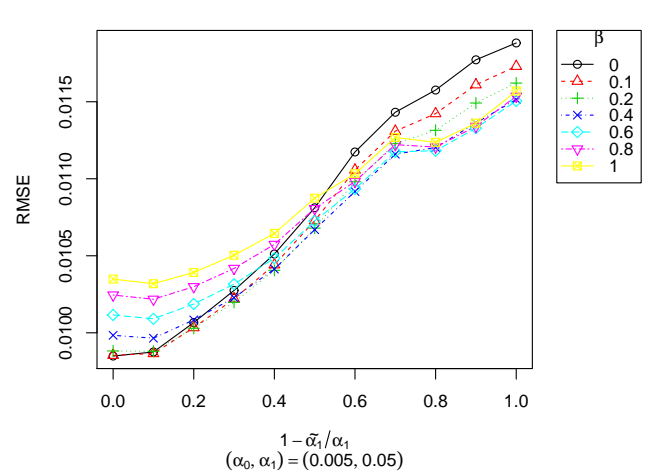
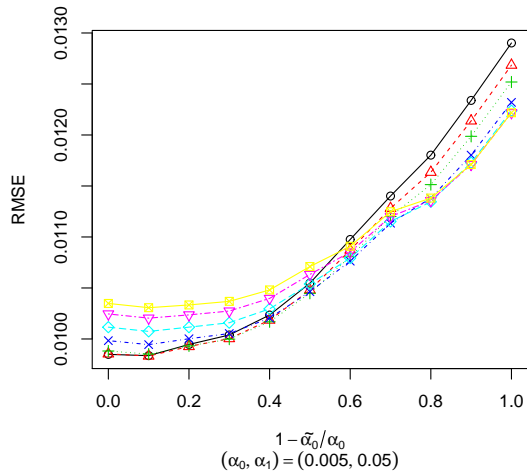
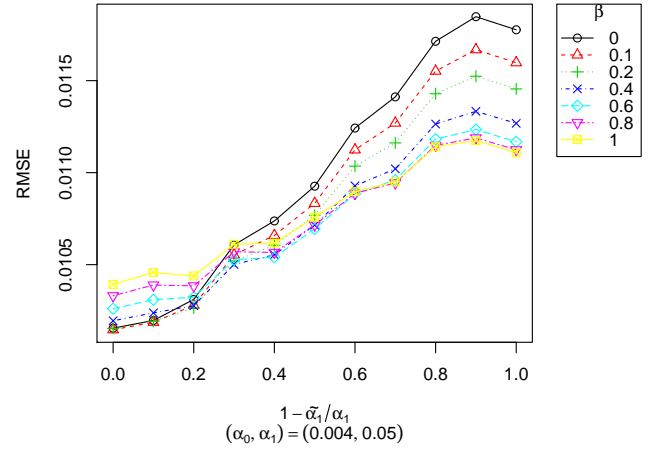
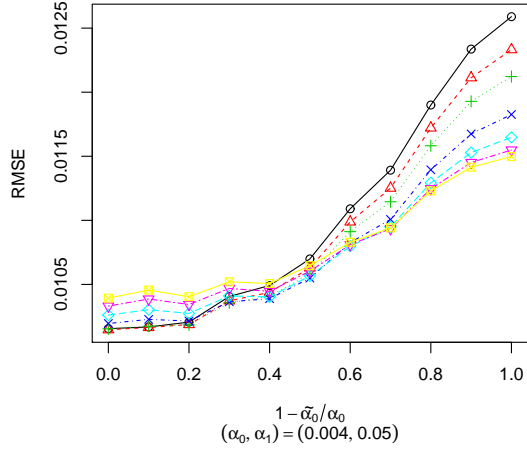
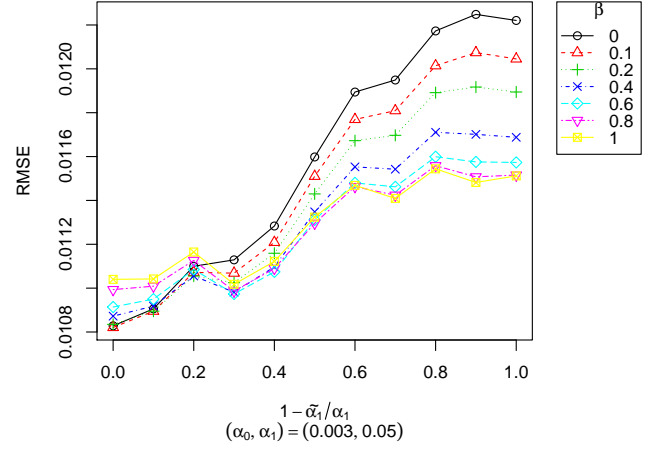
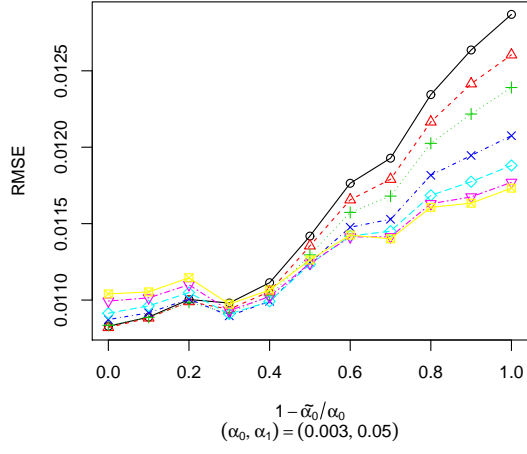
The results of the tests statistics presented here show again the poor behavior in robustness of the Z -type tests based on the MLE of the parameters of the one-shot device model. Furthermore, the robustness properties of the Z -type test statistics based on the MDPDEs with large values of the tuning 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 6.1 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.

$$\mathcal{T}_{1,\beta}(\boldsymbol{\alpha}) = \sum_{i=1}^I \sum_{j=1}^J \left\{ \left(\frac{F(t_j | \lambda_{w_i}(\boldsymbol{\alpha}))}{IJ} \right)^{1+\beta} - \left(1 + \frac{1}{\beta} \right) \left(\frac{F(t_j | \lambda_{w_i}(\boldsymbol{\alpha}))}{IJ} \right)^{\beta} \frac{n_{ij}}{IJK} + \frac{1}{\beta} \left(\frac{n_{ij}}{IJK} \right)^{1+\beta} \right\} \quad (16)$$

$$\mathcal{T}_{2,\beta}(\boldsymbol{\alpha}) = \sum_{i=1}^I \sum_{j=1}^J \left\{ \left(\frac{1 - F(t_j | \lambda_{w_i}(\boldsymbol{\alpha}))}{IJ} \right)^{1+\beta} - \left(1 + \frac{1}{\beta} \right) \left(\frac{1 - F(t_j | \lambda_{w_i}(\boldsymbol{\alpha}))}{IJ} \right)^{\beta} \frac{K - n_{ij}}{IJK} + \frac{1}{\beta} \left(\frac{K - n_{ij}}{IJK} \right)^{1+\beta} \right\} \quad (17)$$

7 Concluding Remarks

In this paper, we have introduced and studied the minimum density power divergence estimators for one-shot device testing with an accelerating factor of temperature. Based on these estimators, we have also introduced a Wald-type test statistic family. Since the maximum likelihood estimator is a particular estimator in the family of minimum density power divergence estimators developed here, the classical Wald test is also taken into account for comparison. The results obtained in the simulation study suggest that some minimum density power divergence estimators are considerably better for the estimation of the model parameters when outliers are present in the data and at the same time not facing 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 maximum likelihood estimators.



(a) with an α_0 -contaminated outlying cell.

(b) with an α_1 -contaminated outlying cell.

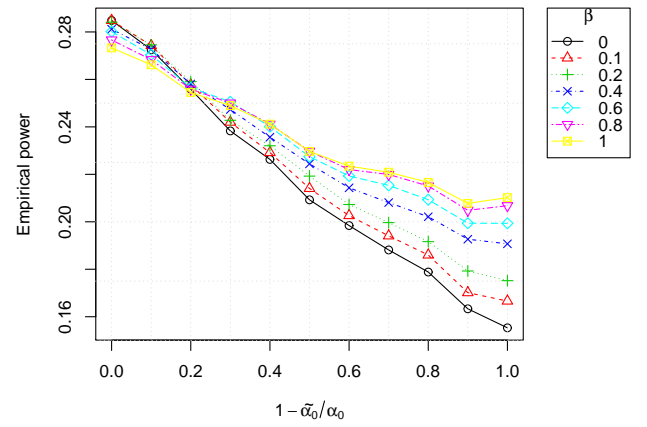
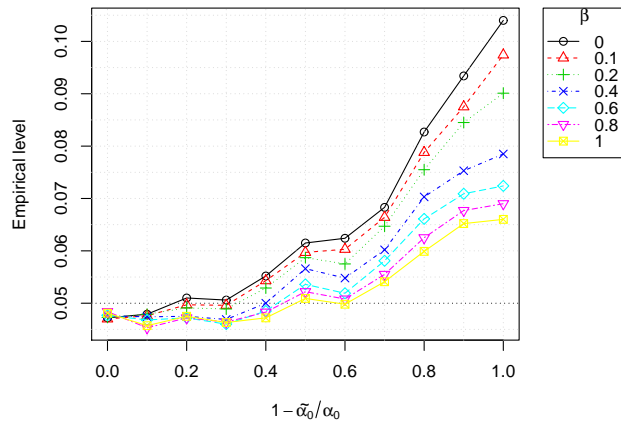
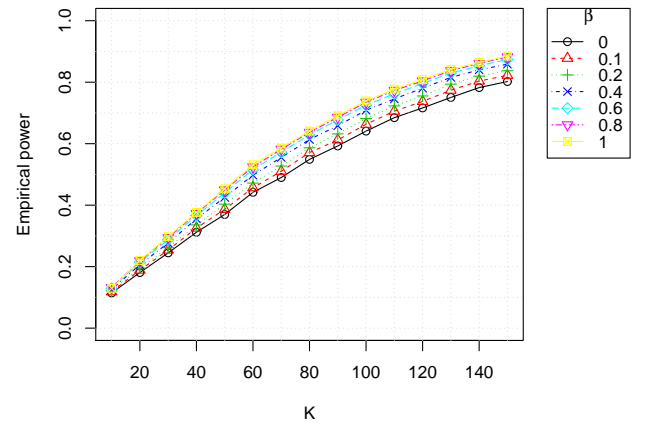
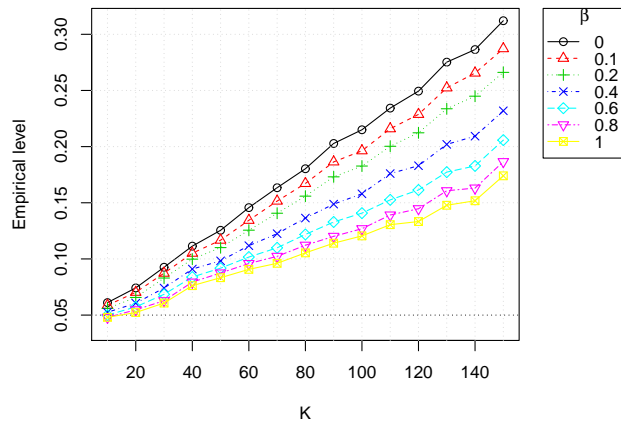


Figure 4: Simulated levels (left) and powers (right) with an α_0 -contaminated outlying cell.

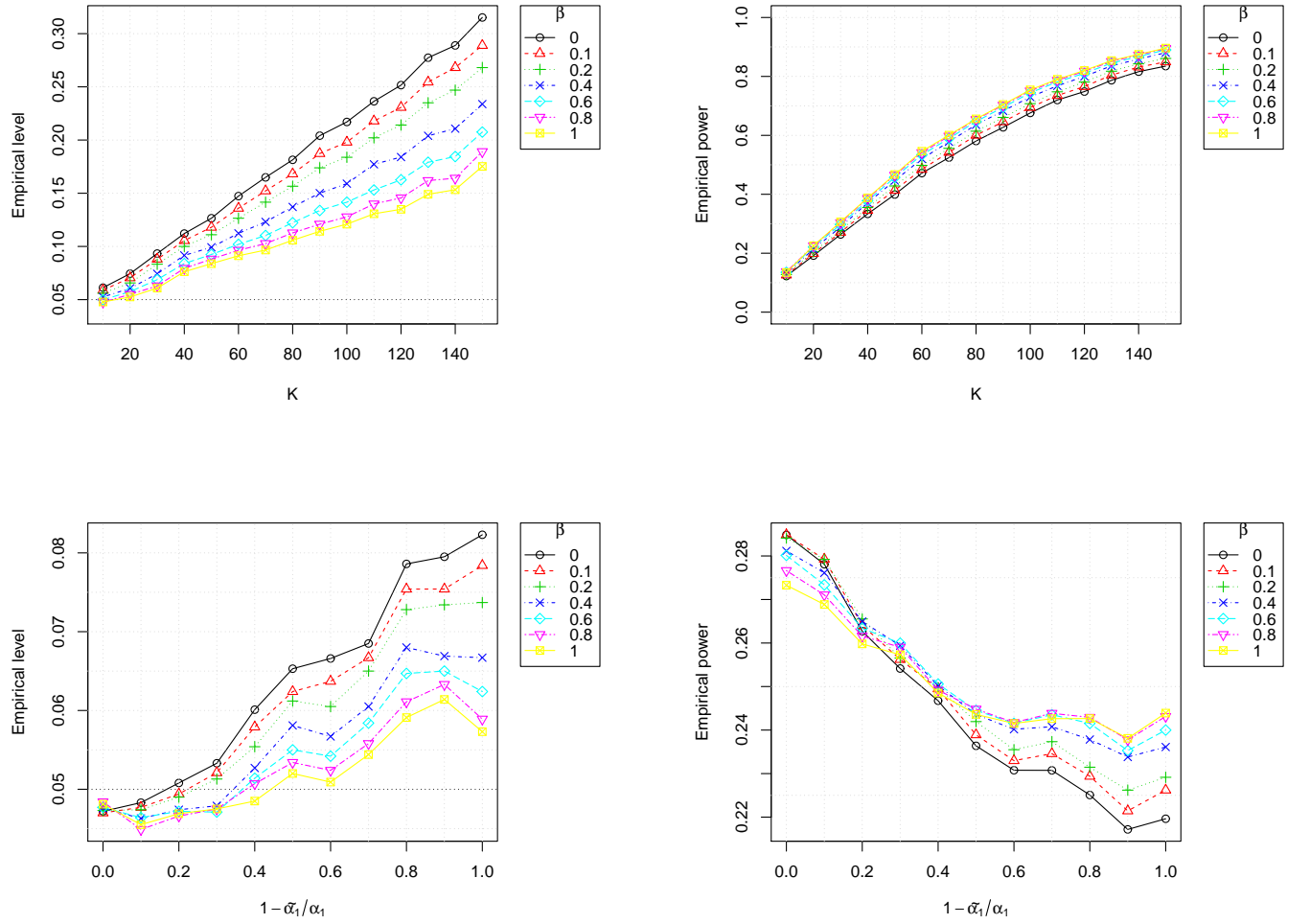


Figure 5: Simulated levels (left) and powers (right) with an α_1 -contaminated outlying cell in the data.

A Proofs of Results

A.1 Proof of Result 1:

We have

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

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.

A.2 Proof of Result 4:

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

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

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

A.3 Proof of Result 5:

We have

$$\begin{aligned}
\frac{\partial F(t_j | \lambda_{w_i}(\boldsymbol{\alpha}))}{\partial \alpha_0} &= \exp \{-\alpha_0 \exp(\alpha_1 w_i) t_j\} \exp \{\alpha_1 w_i\} t_j \\
&= f(t_j | \lambda_{w_i}(\boldsymbol{\alpha})) \frac{t_j}{\alpha_0}
\end{aligned} \tag{18}$$

and

$$\begin{aligned}
\frac{\partial F(t_j|\lambda_{w_i}(\boldsymbol{\alpha}))}{\partial \alpha_1} &= \exp \{ -\alpha_0 \exp(\alpha_1 w_i) t_j \} \\
&\quad \times \exp \{ \alpha_1 w_i \} \alpha_0 t_j w_i \\
&= f(t_j|\lambda_{w_i}(\boldsymbol{\alpha})) t_j w_i.
\end{aligned} \tag{19}$$

We denote

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

where $\mathcal{T}_{1,\beta}(\boldsymbol{\alpha})$ and $\mathcal{T}_{2,\beta}(\boldsymbol{\alpha})$ are given by (16) and (17), respectively, for $\beta > 0$.

Based on (18), 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

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

On the other hand, by (19), 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

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

Finally, the system of equations is given by

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

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

A.4 Proof of Result 6:

Based on Ghosh and Basu (2013) and also on Definition 3, we have

$$\sqrt{IJK}(\hat{\alpha}_\beta - \alpha_0) \xrightarrow{IJK \rightarrow \infty} \mathcal{N}\left(\mathbf{0}, \mathbf{J}_\beta^{-1}(\alpha_0) \mathbf{K}_\beta(\alpha_0) \mathbf{J}_\beta^{-1}(\alpha_0)\right),$$

where

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

$$\begin{aligned} \mathbf{u}_{ij}(\alpha) &= \frac{\partial \log F(t_j | \lambda_{w_i}(\alpha))}{\partial \alpha} \\ &= \frac{1}{F(t_j | \lambda_{w_i}(\alpha))} \frac{\partial}{\partial \alpha} F(t_j | \lambda_{w_i}(\alpha)), \\ \mathbf{v}_{ij}(\alpha) &= \frac{\partial \log [1 - F(t_j | \lambda_{w_i}(\alpha))]}{\partial \alpha} \\ &= - \frac{1}{1 - F(t_j | \lambda_{w_i}(\alpha))} \frac{\partial}{\partial \alpha} F(t_j | \lambda_{w_i}(\alpha)), \\ \frac{\partial}{\partial \alpha} F(t_j | \lambda_{w_i}(\alpha)) &= - \frac{\partial}{\partial \alpha} \exp \{-\alpha_0 \exp \{\alpha_1 w_i\} t_j\} \\ &= \begin{pmatrix} \frac{1}{\alpha_0} \\ w_i \end{pmatrix} t_j f(t_j | \lambda_{w_i}(\alpha)), \end{aligned}$$

and

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

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

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

where

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

A.5 Proof of Result 7:

The Fisher information matrix for IJK observations is

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

where

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

From (3),

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

where

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

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

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

From Result 6, the first and second components of $\mathbf{u}(\boldsymbol{\alpha} | K, \mathbf{n}, \mathbf{t}, \mathbf{w})$ are

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

and

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

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

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

Since the expectation of the first two summands of $\partial u_1(\boldsymbol{\alpha} | K, \mathbf{n}, \mathbf{t}, \mathbf{w}) / \partial \alpha_0$ are zero, the interest is on the expectation of L_{ij} which is given in (20).

$$\begin{aligned}
L_{ij} &= \frac{\partial}{\partial \alpha_0} \left(\frac{K F(t_j | \lambda_{w_i}(\alpha)) - n_{ij}}{F(t_j | \lambda_{w_i}(\alpha)) (1 - F(t_j | \lambda_{w_i}(\alpha)))} \right) \frac{t_j}{\alpha_0} f(t_j | \lambda_{w_i}(\alpha)) \\
&= \frac{K \frac{t_j}{\alpha_0} f(t_j | \lambda_{w_i}(\alpha)) F(t_j | \lambda_{w_i}(\alpha)) (1 - F(t_j | \lambda_{w_i}(\alpha)))}{F(t_j | \lambda_{w_i}(\alpha))^2 (1 - F(t_j | \lambda_{w_i}(\alpha)))^2} \frac{t_j}{\alpha_0} f(t_j | \lambda_{w_i}(\alpha)) \\
&\quad - \frac{\frac{\partial}{\partial \alpha_0} [F(t_j | \lambda_{w_i}(\alpha)) (1 - F(t_j | \lambda_{w_i}(\alpha)))] (K F(t_j | \lambda_{w_i}(\alpha)) - n_{ij})}{F(t_j | \lambda_{w_i}(\alpha))^2 (1 - F(t_j | \lambda_{w_i}(\alpha)))^2} \frac{t_j}{\alpha_0} f(t_j | \lambda_{w_i}(\alpha)). \tag{20}
\end{aligned}$$

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

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

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

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

The rest of the terms of $\mathbf{I}_F(\alpha)$ can be obtained in a similar manner. On the other hand, from Theorem 6, substituting $\beta = 0$ into $\mathbf{J}_\beta(\alpha) = \frac{1}{IJ} \bar{\mathbf{J}}_\beta(\alpha)$ and $\mathbf{K}_\beta(\alpha) = \frac{1}{IJ} \bar{\mathbf{K}}_\beta(\alpha)$, we simply obtain $\mathbf{J}_{\beta=0}(\alpha) = \mathbf{K}_{\beta=0}(\alpha) = \mathbf{I}_F(\alpha)$.

A.6 Proof of Result 9:

Let α_0 be the true value of parameter α . It is clear that under (13)

$$\mathbf{m}^T \hat{\alpha}_\beta - d = \mathbf{m}^T (\hat{\alpha}_\beta - \alpha_0)$$

and we know

$$\sqrt{K}(\hat{\alpha}_\beta - \alpha_0) \xrightarrow{K \rightarrow \infty} \mathcal{N}(\mathbf{0}, \bar{\mathbf{J}}_\beta^{-1}(\alpha_0) \bar{\mathbf{K}}_\beta(\alpha_0) \bar{\mathbf{J}}_\beta^{-1}(\alpha_0)),$$

from which it follows that

$$\sqrt{K}(\mathbf{m}^T \hat{\alpha}_\beta - d) \xrightarrow{K \rightarrow \infty} \mathcal{N}(0, \mathbf{m}^T \bar{\mathbf{J}}_\beta^{-1}(\alpha_0) \bar{\mathbf{K}}_\beta(\alpha_0) \bar{\mathbf{J}}_\beta^{-1}(\alpha_0) \mathbf{m}).$$

Dividing the left hand side by

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

since $\mathbf{m}^T \bar{\mathbf{J}}_\beta^{-1}(\hat{\alpha}_\beta) \bar{\mathbf{K}}_\beta(\hat{\alpha}_\beta) \bar{\mathbf{J}}_\beta^{-1}(\hat{\alpha}_\beta) \mathbf{m}$ is a consistent estimator of $\mathbf{m}^T \bar{\mathbf{J}}_\beta^{-1}(\alpha_0) \bar{\mathbf{K}}_\beta(\alpha_0) \bar{\mathbf{J}}_\beta^{-1}(\alpha_0) \mathbf{m}$, the desired result is obtained.

A.7 Proof of Result 10:

The power function at α^* of $Z_K(\hat{\alpha}_\beta)$ is given by equation (21).

$$\begin{aligned}
\pi(\alpha^*) &= \Pr(|Z_K(\hat{\alpha}_\beta)| > z_{\frac{\alpha}{2}} | \alpha = \alpha^*) \\
&= 2 \Pr(Z_K(\hat{\alpha}_\beta) > z_{\frac{\alpha}{2}} | \alpha = \alpha^*) \\
&= 2 \Pr\left(\sqrt{\frac{K}{\mathbf{m}^T \bar{\mathbf{J}}_\beta^{-1}(\hat{\alpha}_\beta) \bar{\mathbf{K}}_\beta(\hat{\alpha}_\beta) \bar{\mathbf{J}}_\beta^{-1}(\hat{\alpha}_\beta) \mathbf{m}}} (\mathbf{m}^T \hat{\alpha}_\beta - d) > z_{\frac{\alpha}{2}} | \alpha = \alpha^*\right) \\
&= 2 \Pr\left(\sqrt{\frac{K}{\mathbf{m}^T \bar{\mathbf{J}}_\beta^{-1}(\hat{\alpha}_\beta) \bar{\mathbf{K}}_\beta(\hat{\alpha}_\beta) \bar{\mathbf{J}}_\beta^{-1}(\hat{\alpha}_\beta) \mathbf{m}}} \mathbf{m}^T (\hat{\alpha}_\beta - \alpha^*) > \right. \\
&\quad \left. z_{\frac{\alpha}{2}} - \sqrt{\frac{K}{\mathbf{m}^T \bar{\mathbf{J}}_\beta^{-1}(\hat{\alpha}_\beta) \bar{\mathbf{K}}_\beta(\hat{\alpha}_\beta) \bar{\mathbf{J}}_\beta^{-1}(\hat{\alpha}_\beta) \mathbf{m}}} (\mathbf{m}^T \alpha^* - d)\right). \tag{21}
\end{aligned}$$

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

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

the desired result follows.

References

- [1] Basu, A., Harris, I. R., Hjort, N. L. and Jones, M. C. (1998). Robust and efficient estimation by minimizing a density power divergence. *Biometrika*, **85**, 549–559.
- [2] Basu, A., Shioya, H. and Park, C. (2011). *Statistical Inference: The Minimum Distance Approach*. Chapman & Hall/CRC Press, Boca Raton, Florida.
- [3] Basu, A., Mandal, A., Martin, N. and Pardo, L. (2016). Generalized Wald-type tests based on minimum density power divergence estimators. *Statistics*, **50**, 1–26.
- [4] Balakrishnan, N. and Ling, M. H. (2012). EM algorithm for one-shot device testing under the exponential distribution. *Computational Statistics & Data Analysis*, **56**, 502–509.
- [5] Balakrishnan, N. So, H.Y. and Ling, M. H. (2016a). A Bayesian approach for one-shot device testing with exponential lifetimes under competing risks. *IEEE Transactions on Reliability*, **65**, 469–485.
- [6] Balakrishnan, N. So, H.Y. and Ling, M. H. (2016b). Likelihood inference under proportional hazards model for one-shot device testing. *IEEE Transactions on Reliability*, **65**, 446–458.

- [7] Chimitova, E. V. and Balakrishnan, N. (2015). Goodness-of-fit tests for one-shot device testing data. In *Advanced Mathematical and Computational Tools in Metrology and Testing X*. Edited by F. Pavese, W. Bremser, A. Chunovkina, N. Fischer and A. B. Forbe, World Scientific Publishing, pp. 124–133, Singapore.
- [8] Fan, T. H., Balakrishnan, N. and Chang, C. C. (2009). Bayesian approach for highly reliable electro-explosive devices using one-shot device testing. *Journal of Statistical Computation and Simulation*, **79**, 1143–1154.
- [9] Ghosh, A. and Basu, A. (2013). Robust estimation for independent non-homogeneous observations nusing power density divergence with applications to linear regression. *Electronic Journal of Statistics*, **7**, 240–2456.
- [10] Ghosh, A., Mandal, A., Martín, N. and Pardo, L. (2016). Influence analysis of robust Wald-type tests. *Journal of Multivariate Analysis*, **147**, 102–126.
- [11] Lindsey, J. and Ryan, L. (1993). A three-state multiplicative model for rodent tumorigenicity experiments *Journal of the Royal Statistical Society, Series C (Applied Statistics)*, **42**, 283–300.
- [12] Pardo, L. (2005). *Statistical Inference Based on Divergence Measures*. *Statistics*: Chapman & Hall/CRC Press, New York.
- [13] Rodrigues, J., Bolfarine, H. and Louzada-Nieto, F. (1993). Comparing several accelerated life methods. *Communications in Statistics Theory and Methods*, **22**, 2297–2308.