Metrika
https://doi.org/10.1007/s00184-019-00718-5

Robust estimators for one-shot device testing data under gamma lifetime model with an application to a tumor toxicological data

N. Balakrishnan[1] · E. Castilla[2] · N. Martín[3] · L. Pardo[2]

Received: 8 November 2018

Abstract
Due to its flexibility, gamma distribution is commonly used for lifetime data analysis in reliability and survival studies, and especially in one-shot device testing data. In the study of such data, inducing more failures by accelerated life tests is a common practice, to obtain more lifetime information within a relatively short period of time. In this paper, we develop weighted minimum density power divergence estimators, as a natural extension of the classical maximum likelihood estimator, in the analysis of one-shot device testing data, under accelerated life tests based on gamma lifetime distribution. Wald-type test statistics, based on these estimators, are also developed. Through a Monte Carlo simulation study, the suggested estimators and tests are shown to be robust alternatives to the maximum likelihood estimators and the classical Wald tests based on them. Finally, these procedures are applied to a mice tumor toxicological data for illustrative purpose.

Keywords Gamma distribution · Maximum likelihood estimator · Minimum density power divergence estimator · Multiple stresses · One-shot devices · Robustness · Tumor toxicological data · Wald-type tests

1 Introduction

Most manufactured products are of high quality these days and so they will usually have long lifetimes. Consequently, if the products are tested under normal conditions, the failure times of the products will be quite large resulting in a large

✉ E. Castilla
elecasti@ucm.es

[1] Department of Mathematics and Statistics, McMaster University, Hamilton, ON, Canada

[2] Department of Statistics and O.R. I, Complutense University of Madrid, Madrid, Spain

[3] Department of Financial and Actuarial Economics and Statistics, Complutense University of Madrid, Madrid, Spain

Published online: 22 May 2019

testing time. To reduce the experimental time and cost, therefore, accelerated life tests (ALTs) are commonly employed to evaluate the reliability of such products. An ALT shortens the life span of the products by increasing the levels of stress factors, such as temperature and humidity. After estimating the parameters from data collected under high-stress conditions, one usually extrapolates the life characteristics, such as mean lifetime and failure rates, from high- stress conditions to normal operating conditions; see Meeter and Meeker (1994) and Meeker et al. (1998).

Inferential methods based on one-shot devices from ALT data have been discussed extensively recently, mainly motivated by the work of Fan et al. (2009), who discussed the Bayesian approach for one-shot device testing along with an accelerating factor. One-shot devices, which can be used only once as the test devices get destroyed immediately after testing, yield only binary response data, since one can only observe the condition of the device at a specific inspection time, never the exact failure time. Balakrishnan and Ling (2012, 2013, 2014) used the EM algorithm for determining the maximum likelihood estimates (MLEs) of the model parameters, considering a multiple-stress relationship under exponential, Weibull and gamma distributions, respectively.

Gamma distribution is commonly used for fitting lifetime data in reliability and survival studies due to its flexibility. Its hazard function can be increasing, decreasing, and constant. When the hazard function of gamma distribution is a constant, it corresponds to the exponential distribution. In addition to the exponential distribution, the gamma distribution also includes the Chi-square distribution as a special case. The gamma distribution has found a number of applications in different fields. For example, Husak et al. (2007) used it to describe monthly rainfall in Africa for the management of water and agricultural resources, as well as food reserves. Kwon and Frangopol (2010) assessed and predicted bridge fatigue reliabilities of two existing bridges, the Neville Island Bridge and the Birmingham Bridge, based on long-term monitoring data. They made use of log-normal, Weibull, and gamma distributions to estimate the mean and standard deviation of the stress range. Tseng et al. (2009) proposed an optimal step-stress accelerated degradation testing plan for assessing the lifetime distribution of products with longer lifetime based on a gamma process.

For the estimation of parameters of one-shot devices under the gamma distribution, the weighted minimum density power divergence estimators (WMDPDEs) are developed in this work. WMDPDEs have been developed recently by Balakrishnan et al. (2019) for one-shot device testing data with exponential lifetimes. While the classical MLE has optimal asymptotic properties regarding efficiency, its serious lack of robustness against outlying observations poses a problem in practice. The WMDPDE provides a good robust alternative to the MLE.

In Sect. 2, we describe the model specification as well as the MLE of the model parameters. Based on the relation between the MLE and the Kullback–Leibler divergence, in Sect. 3, the WMDPDE is considered as a natural extension of the MLE and its asymptotic distribution is obtained. Wald-type test statistics are introduced in Sect. 4, based on these estimators. Their asymptotic and robustness properties are then studied. A simulation study is carried out in Sect. 5, while a numerical example is presented in Sect. 6, based on a data from a tumor toxicological study. Finally, some concluding remarks are made in Sect. 7.

2 Model specification

In accelerated life tests for one-shot devices, the devices are placed in I testing groups. Suppose, for $i = 1, 2, \dots, I$, K_i devices are subjected to J types of stress factors, x_{ij}, $j = 1, 2, \dots, J$, under inspection time IT_i. In the ith test group, the number of failures, n_i, is collected. The data thus observed can be summarized as in Table 1.

Let T be a gamma random variable with shape parameter $\alpha > 0$ and scale parameter $\lambda > 0$. The probability density function (pdf) and the cumulative distribution function (cdf) of T are given by

$$f(t) = \frac{t^{\alpha-1}}{\lambda^{\alpha}\Gamma(\alpha)} \exp\left(-\frac{t}{\lambda}\right),\ t > 0, \tag{1}$$

and

$$F(t) = \int_0^t \frac{y^{\alpha-1}}{\lambda^{\alpha}\Gamma(\alpha)} \exp\left(-\frac{y}{\lambda}\right)dy,\ t > 0. \tag{2}$$

We assume that the lifetimes of the units follow the gamma distribution and that, within each testing group, both shape and scale parameters are related to the stress factors in log-linear forms as

$$\alpha_i = \exp\left\{\sum_{j=0}^{J} a_j x_{ij}\right\} = \exp(\boldsymbol{x}_i^T \boldsymbol{a})$$

and

$$\lambda_i = \exp\left\{\sum_{j=0}^{J} b_j x_{ij}\right\} = \exp(\boldsymbol{x}_i^T \boldsymbol{b}),$$

Table 1 Typical form of one-shot device test data observed

Condition	Inspection time	Devices	Failures	Covariates		
				Stress 1	⋯	Stress J
1	IT_1	K_1	n_1	x_{11}	⋯	x_{1J}
2	IT_2	K_2	n_2	x_{21}	⋯	x_{2J}
⋮	⋮	⋮	⋮	⋮		⋮
I	IT_I	K_I	n_I	x_{I1}	⋯	x_{IJ}

where $x_{i0} = 1$, $\boldsymbol{x}_i = (x_{i0}, ..., x_{iJ})^T$, $\boldsymbol{a} = (a_0, a_1, \dots a_J)^T$ and $\boldsymbol{b} = (b_0, b_1, \dots b_J)^T$. Denoting $\boldsymbol{\theta} = \{a_j, b_j, j = 0, ..., J\}$ for the model parameters, the likelihood function based on this observed data is then given by

$$\mathcal{L}(\boldsymbol{\theta}) \propto \prod_{i=1}^{I} F\big(IT_i; \boldsymbol{x}_i, \boldsymbol{\theta}\big)^{n_i} \big(1 - F\big(IT_i; \boldsymbol{x}_i, \boldsymbol{\theta}\big)\big)^{K_i - n_i} \tag{3}$$

and the MLE, $\widehat{\boldsymbol{\theta}} = \left\{\widehat{a}_j, \widehat{b}_j, j = 1, ..., J\right\}$, is then given by

$$\widehat{\boldsymbol{\theta}} = \arg\max_{\boldsymbol{\theta}} \ \log \mathcal{L}(\boldsymbol{\theta}). \tag{4}$$

3 Minimum density power divergence estimator for the model parameters

In order to present the WMDPDEs for the model parameter, $\boldsymbol{\theta}$, we consider an alternative definition of the MLE, $\widehat{\boldsymbol{\theta}}$, on the basis of Kullback–Leibler divergence. We denote by $\widehat{p}_{i1} = \frac{n_i}{K_i}$ and $\widehat{p}_{i2} = 1 - \frac{n_i}{K_i}$ and then consider the following empirical probability vectors

$$\widehat{\boldsymbol{p}}_i = \big(\widehat{p}_{i1}, \widehat{p}_{i2}\big)^T, \quad i = 1, \dots, I. \tag{5}$$

Denoting by $\pi_{i1}(\boldsymbol{\theta}) = F\big(IT_i; \boldsymbol{x}_i, \boldsymbol{\theta}\big)$ and $\pi_{i2}(\boldsymbol{\theta}) = 1 - F\big(IT_i; \boldsymbol{x}_i, \boldsymbol{\theta}\big)$, we now introduce the theoretical probability vectors

$$\boldsymbol{\pi}_i(\boldsymbol{\theta}) = (\pi_{i1}(\boldsymbol{\theta}), \pi_{i2}(\boldsymbol{\theta}))^T, \quad i = 1, \dots, I. \tag{6}$$

For each fixed i, the Kullback–Leibler divergence measure between the probability vectors $\widehat{\boldsymbol{p}}_i$ and $\boldsymbol{\pi}_i(\boldsymbol{\theta})$ is given by

$$\begin{aligned} d_{KL}(\widehat{\boldsymbol{p}}_i, \boldsymbol{\pi}_i(\boldsymbol{\theta})) = {} & \frac{n_i}{K_i} \log\left(\frac{\frac{n_i}{K_i}}{F\big(IT_i; \boldsymbol{x}_i, \boldsymbol{\theta}\big)}\right) \\ & + \left(1 - \frac{n_i}{K_i}\right) \log\left(\frac{1 - \frac{n_i}{K_i}}{1 - F\big(IT_i; \boldsymbol{x}_i, \boldsymbol{\theta}\big)}\right), \end{aligned} \tag{7}$$

and the weighted Kullback–Leibler divergence measure is given by

$$d_{KL}^{W}(\boldsymbol{\theta}) = \frac{1}{K} \sum_{i=1}^{I} \left[n_i \log\left(\frac{\frac{n_i}{K_i}}{F\big(IT_i; \boldsymbol{x}_i, \boldsymbol{\theta}\big)}\right) + (K_i - n_i) \log\left(\frac{\frac{K_i - n_i}{K_i}}{1 - F\big(IT_i; \boldsymbol{x}_i, \boldsymbol{\theta}\big)}\right)\right], \tag{8}$$

where $K = K_1 + \cdots + K_I$ is the total number of devices under test. Kullback–Leibler divergence measure was introduced and studied by Kullback and Leibler (1951) and some important statistical applications of this divergence measure have been given by Kullback (1959). See also Pardo (2006).

Theorem 1 *The likelihood function $\mathcal{L}(\boldsymbol{\theta})$, given in (3), is related to the Kullback–Leibler divergence between the probability vectors $\widehat{\boldsymbol{p}}$ and $\boldsymbol{\pi}(\boldsymbol{\theta})$ through*

$$d_{KL}^{W}(\boldsymbol{\theta}) = c - \frac{1}{K} \log \mathcal{L}(\boldsymbol{\theta}), \tag{9}$$

with c being a constant not dependent on $\boldsymbol{\theta}$.

Based on Theorem 1, we have the following definition for the MLE of $\boldsymbol{\theta}$.

Definition 2 The MLE of $\boldsymbol{\theta}$ can be defined as

$$\widehat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} d_{KL}^{W}(\boldsymbol{\theta}). \tag{10}$$

Definition 3 Given the probability vectors $\widehat{\boldsymbol{p}}_i$ and $\boldsymbol{\pi}_i(\boldsymbol{\theta})$, defined in (5) and (6), respectively, the density power divergence (DPD) between the two probability vectors is given by

$$\begin{aligned} d_{\beta}(\widehat{\boldsymbol{p}}_i, \boldsymbol{\pi}_i(\boldsymbol{\theta})) = &\left(\pi_{i1}^{\beta+1}(\boldsymbol{\theta}) + \pi_{i2}^{\beta+1}(\boldsymbol{\theta})\right) - \frac{\beta+1}{\beta}\left(\widehat{p}_{i1}\pi_{i1}^{\beta}(\boldsymbol{\theta}) + \widehat{p}_{i2}\pi_{i2}^{\beta}(\boldsymbol{\theta})\right) \\ &+ \frac{1}{\beta}\left(\widehat{p}_{i1}^{\beta+1} + \widehat{p}_{i2}^{\beta+1}\right), \quad \text{if } \beta > 0, \end{aligned} \tag{11}$$

and $d_{\beta=0}(\widehat{\boldsymbol{p}}_i, \boldsymbol{\pi}_i(\boldsymbol{\theta})) = \lim_{\beta \to 0^+} d_{\beta}(\widehat{\boldsymbol{p}}_i, \boldsymbol{\pi}_i(\boldsymbol{\theta})) = d_{KL}(\widehat{\boldsymbol{p}}_i, \boldsymbol{\pi}_i(\boldsymbol{\theta}))$, for $\beta = 0$.

Based on Definition 3, we can define the weighted density power divergence (WDPD) measure as below.

Definition 4 Given the probability vectors $\widehat{\boldsymbol{p}}_i$ and $\boldsymbol{\pi}_i(\boldsymbol{\theta}), i = 1, ..., I$, defined in (5) and (6), respectively, the WDPD measure is given by

$$\begin{aligned} d_{\beta}^{W}(\boldsymbol{\theta}) &= \sum_{i=1}^{I} \frac{K_i}{K} d_{\beta}(\widehat{\boldsymbol{p}}_i, \boldsymbol{\pi}_i(\boldsymbol{\theta})) \\ &= \sum_{i=1}^{I} \frac{K_i}{K}\left[\left(\pi_{i1}^{\beta+1}(\boldsymbol{\theta}) + \pi_{i2}^{\beta+1}(\boldsymbol{\theta})\right) - \frac{\beta+1}{\beta}\left(\widehat{p}_{i1}\pi_{i1}^{\beta}(\boldsymbol{\theta}) + \widehat{p}_{i2}\pi_{i2}^{\beta}(\boldsymbol{\theta})\right)\right. \\ &\quad \left. + \frac{1}{\beta}\left(\widehat{p}_{i1}^{\beta+1} + \widehat{p}_{i2}^{\beta+1}\right)\right]. \end{aligned} \tag{12}$$

We observe in (12) that the term

$$\frac{1}{\beta}\sum_{i=1}^{I}\frac{K_i}{K}\left(\widehat{p}_{i1}^{\beta+1}+\widehat{p}_{i2}^{\beta+1}\right)$$

does not have any role in the minimization of (12) with respect to $\boldsymbol{\theta}$. Therefore, we can consider the equivalent measure

$$^{*}d_{\beta}^{W}(\boldsymbol{\theta})=\sum_{i=1}^{I}\frac{K_i}{K}\left[\left(\pi_{i1}^{\beta+1}(\boldsymbol{\theta})+\pi_{i2}^{\beta+1}(\boldsymbol{\theta})\right)-\frac{\beta+1}{\beta}\left(\widehat{p}_{i1}\pi_{i1}^{\beta}(\boldsymbol{\theta})+\widehat{p}_{i2}\pi_{i2}^{\beta}(\boldsymbol{\theta})\right)\right]. \quad (13)$$

Definition 5 Based on (12) and (13), we can define the WMDPDEs for $\boldsymbol{\theta}$ as

$$\widehat{\boldsymbol{\theta}}_{\beta}=\arg\min_{\boldsymbol{\theta}}\,{}^{*}d_{\beta}^{W}(\widehat{\boldsymbol{p}},\boldsymbol{p}(\boldsymbol{\theta}))\quad\text{for }\beta>0$$

and for $\beta=0$, we have the MLE of $\boldsymbol{\theta}$.

Theorem 6 *For $\beta\geq 0$, the estimating equations are given by*

$$\begin{aligned}&\sum_{i=1}^{I}\left(K_i\pi_{i1}(\boldsymbol{\theta})-n_i\right)\left(F^{\beta-1}\left(IT_i;\boldsymbol{x}_i,\boldsymbol{\theta}\right)+\left(1-F\left(IT_i;\boldsymbol{x}_i,\boldsymbol{\theta}\right)\right)^{\beta-1}\right)\\&\quad f\left(IT_i;\boldsymbol{x}_i,\boldsymbol{\theta}\right)\frac{\partial f\left(IT_i;a_i,b_i\right)}{\partial\boldsymbol{a}}=\boldsymbol{0}_J,\\&\sum_{i=1}^{I}\left(K_i\pi_{i1}(\boldsymbol{\theta})-n_i\right)\left(F^{\beta-1}\left(IT_i;\boldsymbol{x}_i,\ \boldsymbol{\theta}\right)+\left(1-F\left(IT_i;\boldsymbol{x}_i,\ \boldsymbol{\theta}\right)\right)^{\beta-1}\right)\\&\quad f\left(IT_i;\boldsymbol{x}_i,\boldsymbol{\theta}\right)\frac{\partial f\left(IT_i;\boldsymbol{x}_i,\boldsymbol{\theta}\right)}{\partial\boldsymbol{b}}=\boldsymbol{0}_J,\end{aligned}$$

where $f(IT_i;\boldsymbol{a},\boldsymbol{b},\boldsymbol{w}_i)$ is as given in (1) and

$$\begin{aligned}\frac{\partial f\left(IT_i;\boldsymbol{x}_i,\boldsymbol{\theta}\right)}{\partial\boldsymbol{a}}&=\frac{\lambda_i^{\alpha_i}IT_i^{\alpha_i-1}}{\Gamma\left(\alpha_i\right)}\left\{-\log\lambda_i+\log IT_i-\Psi\left(\alpha_i\right)\right\}\boldsymbol{x}_i,\quad\frac{\partial f\left(IT_i;\boldsymbol{x}_i,\boldsymbol{\theta}\right)}{\partial\boldsymbol{b}}\\&=\frac{\lambda_i^{\alpha_i-1}IT_i^{\alpha_i-1}}{\Gamma\left(\alpha_i\right)}\left[-\alpha_i+\frac{IT_i}{\lambda_i}\right]\boldsymbol{x}_i,\end{aligned}$$

with Ψ being the digamma function.

In the following results, the asymptotic distribution of the WMDPDEs of $\boldsymbol{\theta}$, $\widehat{\boldsymbol{\theta}}_\beta$ is presented for one-shot device testing data under gamma lifetimes.

Theorem 7 *Let $\boldsymbol{\theta}_0$ be the true value of the parameter $\boldsymbol{\theta}$. The asymptotic distribution of the WMDPDEs, $\widehat{\boldsymbol{\theta}}_\beta$, is given by*

$$\sqrt{K}\left(\widehat{\boldsymbol{\theta}}_\beta - \boldsymbol{\theta}_0\right) \xrightarrow[K\to\infty]{\mathcal{L}} \mathcal{N}\left(\mathbf{0}_{2J}, \boldsymbol{J}_\beta^{-1}(\boldsymbol{\theta}_0)\boldsymbol{K}_\beta(\boldsymbol{\theta}_0)\boldsymbol{J}_\beta^{-1}(\boldsymbol{\theta}_0)\right),$$

where

$$\boldsymbol{J}_\beta(\boldsymbol{\theta}) = \sum_{i=1}^{I} \frac{K_i}{K}\boldsymbol{M}_i\left(F^{\beta-1}(IT_i;\boldsymbol{x}_i,\boldsymbol{\theta}) + \left(1 - F(IT_i;\boldsymbol{x}_i,\boldsymbol{\theta})\right)^{\beta-1}\right), \tag{14}$$

$$\begin{aligned}\boldsymbol{K}_\beta(\boldsymbol{\theta}) = \sum_{i=1}^{I} \frac{K_i}{K}&\boldsymbol{M}_i F(IT_i;\boldsymbol{x}_i,\boldsymbol{\theta})\left(1 - F(IT_i;\boldsymbol{x}_i,\boldsymbol{\theta})\right)\\ &\left(F^{\beta-1}(IT_i;\boldsymbol{x}_i,\boldsymbol{\theta}) + \left(1 - F(IT_i;\boldsymbol{x}_i,\boldsymbol{\theta})\right)^{\beta-1}\right)^2,\end{aligned} \tag{15}$$

$$\begin{aligned}\boldsymbol{M}_i &= \begin{pmatrix} l_i^2\boldsymbol{x}_i\boldsymbol{x}_i^T & l_is_i\boldsymbol{x}_i\boldsymbol{x}_i^T \\ l_is_i\boldsymbol{x}_i\boldsymbol{x}_i^T & s_i^2\boldsymbol{x}_i\boldsymbol{x}_i^T \end{pmatrix},\\ l_i &= \alpha_i\left\{-\Psi\left(\alpha_i\right)\pi_{i1}(\boldsymbol{\theta}) + \log\left(\frac{IT_i}{\lambda_i}\right)\pi_{i1}(\boldsymbol{\theta}) - \frac{\left(\frac{IT_i}{\lambda_i}\right)^{\alpha_i}}{\alpha_i^2\Gamma(\alpha_i)}{}_2\right.\\ &\qquad \left. F_2\left(\alpha_i,\alpha_i;1+\alpha_i,1+\alpha_i;-\frac{IT_i}{\lambda_i}\right)\right\}\end{aligned} \tag{16}$$

and

$$s_i = -f\left(IT_i;\boldsymbol{x}_i,\boldsymbol{\theta}\right)IT_i. \tag{17}$$

Here, ${}_nF_m(a_1,\dots,a_n;b_1,\dots,b_m;z)$ denotes the Gaussian hypergeometric function. For more details about the Gaussian hypergeometric function, one may refer to Seabom (1991).

4 Wald-type test statistics

In this section, we present a new family of test statistics for testing different hypotheses concerning the parameter $\boldsymbol{\theta} = \left\{a_j, b_j, j = 0, 1, ..., J\right\}$ in the model under consideration. The new family of test statistics is based on the WMDPDEs, and we shall refer to them as the Wald-type test statistics.

In the following, we consider the composite null hypotheses of the type

$$\boldsymbol{m}(\boldsymbol{\theta}) = \mathbf{0}_r, \tag{18}$$

where $\boldsymbol{m}$ is a function, $\boldsymbol{m} : \mathbb{R}^{2J} \longrightarrow \mathbb{R}^r$, where $r \leq 2J$.

We assume that the $2J \times r$ matrix

$$\boldsymbol{M}(\boldsymbol{\theta}) = \frac{\partial \boldsymbol{m}(\boldsymbol{\theta})^T}{\partial \boldsymbol{\theta}}$$

exists and is continuous in "$\boldsymbol{\theta}$" and with rank $\boldsymbol{M}(\boldsymbol{\theta}) = r$.

Definition 8 For testing

$$H_0 : \boldsymbol{\theta} \in \boldsymbol{\Theta}_0 \text{ against } H_1 : \boldsymbol{\theta} \notin \boldsymbol{\Theta}_0, \tag{19}$$

where $\boldsymbol{\Theta}_0 = \left\{\boldsymbol{\theta} \in \mathbb{R}^{J+1} : \boldsymbol{m}(\boldsymbol{\theta}) = \mathbf{0}\right\}$, we consider the Wald-type test statistic

$$W_K\left(\widehat{\boldsymbol{\theta}}_\beta\right) = K\boldsymbol{m}\left(\widehat{\boldsymbol{\theta}}_\beta\right)^T \left(\boldsymbol{M}\left(\widehat{\boldsymbol{\theta}}_\beta\right)^T \boldsymbol{\Sigma}\left(\widehat{\boldsymbol{\theta}}_\beta\right)\boldsymbol{M}\left(\widehat{\boldsymbol{\theta}}_\beta\right)\right)^{-1} \boldsymbol{m}\left(\widehat{\boldsymbol{\theta}}_\beta\right),$$

where

$$\boldsymbol{\Sigma}\left(\widehat{\boldsymbol{\theta}}_\beta\right) = \left(\boldsymbol{J}_\beta\left(\widehat{\boldsymbol{\theta}}_\beta\right)\right)^{-1} \boldsymbol{K}_\beta\left(\widehat{\boldsymbol{\theta}}_\beta\right)\left(\boldsymbol{J}_\beta\left(\widehat{\boldsymbol{\theta}}_\beta\right)\right)^{-1}$$

and $\boldsymbol{J}_\beta(\boldsymbol{\theta})$, $\boldsymbol{K}_\beta(\boldsymbol{\theta})$ are as in (14) and (15), respectively.

4.1 Asymptotic properties

Theorem 9 *We have*

$$W_K\left(\widehat{\boldsymbol{\theta}}_\beta\right) \xrightarrow[K\to\infty]{\mathcal{L}} \chi^2_r,$$

where χ^2_r is the upper α percentage point of χ^2_r distribution.

Based on Theorem 9, we shall reject the null hypothesis in (19) if

$$W_K\left(\widehat{\boldsymbol{\theta}}_\beta\right) > \chi^2_{r,\alpha}.$$

In many cases, the power function of this test procedure cannot be derived explicitly. In the following theorem, we present an useful asymptotic result for approximating the power function of the Wald-type test statistic in (19). We shall assume that $\boldsymbol{\theta}^* \notin \boldsymbol{\Theta}_0$ is the true value of the parameter such that

$$\widehat{\boldsymbol{\theta}}_\beta \xrightarrow[K\to\infty]{P} \boldsymbol{\theta}^*,$$

and we denote

$$\ell_\beta(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2) = \boldsymbol{m}^T(\boldsymbol{\theta}_1)\left(\boldsymbol{M}^T(\boldsymbol{\theta}_2)\boldsymbol{\Sigma}_\beta(\boldsymbol{\theta}_2)\boldsymbol{M}(\boldsymbol{\theta}_2)\right)^{-1}\boldsymbol{m}(\boldsymbol{\theta}_1).$$

We then have the following result.

Theorem 10 *We have*

$$\sqrt{K}\left(\ell_\beta\left(\widehat{\boldsymbol{\theta}}_1, \widehat{\boldsymbol{\theta}}_2\right) - \ell_\beta(\boldsymbol{\theta}^*, \boldsymbol{\theta}^*)\right) \xrightarrow[K\to\infty]{\mathcal{L}} \mathcal{N}(0, \sigma^2_{W_K,\beta}(\boldsymbol{\theta}^*)),$$

where

$$\sigma^2_{W_K,\beta}(\boldsymbol{\theta}^*) = \left.\frac{\partial \ell_\beta(\boldsymbol{\theta}, \boldsymbol{\theta}^*)}{\partial \boldsymbol{\theta}^T}\right|_{\boldsymbol{\theta}=\boldsymbol{\theta}^*} \boldsymbol{\Sigma}_\beta(\boldsymbol{\theta}^*) \left.\frac{\partial \ell_\beta(\boldsymbol{\theta}, \boldsymbol{\theta}^*)}{\partial \boldsymbol{\theta}}\right|_{\boldsymbol{\theta}=\boldsymbol{\theta}^*}.$$

Remark 11 Based on Theorem 10, we can give an approximation of the power function of the Wald-type test statistic in $\boldsymbol{\theta}^*$ to be

$$\begin{aligned}
\pi_{W,K}(\boldsymbol{\theta}^*) &= \Pr\left(W_K\left(\widehat{\boldsymbol{\theta}}_\beta\right) > \chi^2_{r,\alpha}\right) \\
&= \Pr\left(K\left(\ell_\beta\left(\widehat{\boldsymbol{\theta}}_\beta, \boldsymbol{\theta}^*\right) - \ell_\beta(\boldsymbol{\theta}^*, \boldsymbol{\theta}^*)\right) > \chi^2_{r,\alpha} - K\ell_\beta(\boldsymbol{\theta}^*, \boldsymbol{\theta}^*)\right) \\
&= \Pr\left(\frac{\sqrt{K}\left(\ell_\beta\left(\widehat{\boldsymbol{\theta}}_\beta, \boldsymbol{\theta}^*\right) - \ell_\beta(\boldsymbol{\theta}^*, \boldsymbol{\theta}^*)\right)}{\sigma_{W_K,\beta}(\boldsymbol{\theta}^*)}\right. \\
&\quad \left. > \frac{1}{\sigma_{W_K,\beta}(\boldsymbol{\theta}^*)}\left(\frac{\chi^2_{r,\alpha}}{\sqrt{K}} - \sqrt{K}\ell_\beta(\boldsymbol{\theta}^*, \boldsymbol{\theta}^*)\right)\right) \\
&= 1 - \Phi_K\left(\frac{1}{\sigma_{W_K,\beta}(\boldsymbol{\theta}^*)}\left(\frac{\chi^2_{r,\alpha}}{\sqrt{K}} - \sqrt{K}\ell_\beta(\boldsymbol{\theta}^*, \boldsymbol{\theta}^*)\right)\right),
\end{aligned}$$

for a sequence of distributions functions, $\Phi_K(x)$, tending uniformly to the standard normal distribution $\Phi(x)$. It is clear that

$$\lim_{K\to\infty} \pi_{W,K}\left(\boldsymbol{\theta}^*\right) = 1,$$

i.e., the Wald-type test statistics are consistent.

The above approximation of the power function of the Wald-type test statistic can be used to obtain the sample size, K, necessary to achieve a prefixed power $\pi_{W,K}\left(\boldsymbol{\theta}^*\right) = \pi_0$. To do that, it will be necessary to solve the equation

$$\pi_0 = 1 - \Phi_K\left(\frac{1}{\sigma_{W_K,\beta}\left(\boldsymbol{\theta}^*\right)}\left(\frac{\chi^2_{r,\alpha}}{\sqrt{K}} - \sqrt{K}\ell_\beta\left(\boldsymbol{\theta}^*, \boldsymbol{\theta}^*\right)\right)\right). \tag{20}$$

The solution, in K, of (20) is

$$\widehat{K}_\beta = \left[\widehat{K}^*_\beta\right] + 1,$$

where

$$\widehat{K}^*_\beta = \frac{\widehat{A}_\beta + \widehat{B}_\beta + \sqrt{\widehat{A}_\beta(\widehat{A}_\beta + 2\widehat{B}_\beta)}}{2\ell^2_\beta\left(\boldsymbol{\theta}^*, \boldsymbol{\theta}^*\right)},$$

with

$$\widehat{A}_\beta = \sigma^2_{W_K,\beta}\left(\boldsymbol{\theta}^*\right)\left(\Phi^{-1}\left(1-\pi_0\right)\right)^2 \text{ and } \widehat{B}_\beta = 2\ell_\beta\left(\boldsymbol{\theta}^*, \boldsymbol{\theta}^*\right)\chi^2_{r,\alpha}.$$

We may also find an approximation of the power of $W_K\left(\widehat{\boldsymbol{\theta}}_\beta\right)$ at an alternative close to the null hypothesis. Let $\boldsymbol{\theta}_n \in \Theta - \Theta_0$ be a given alternative, and let $\boldsymbol{\theta}_0$ be the element in Θ_0 closest to $\boldsymbol{\theta}_n$ in the sense of Euclidean distance. The first possibility to introduce contiguous alternative hypotheses is to consider a fixed $\boldsymbol{d} \in \mathbb{R}^p$ and to allow $\boldsymbol{\theta}_n$ to move toward $\boldsymbol{\theta}_0$ as n increases as follows:

$$H_{1,n} : \boldsymbol{\theta}_n = \boldsymbol{\theta}_0 + n^{-1/2}\boldsymbol{d}. \tag{21}$$

The second approach is to relax the condition $\boldsymbol{m}(\boldsymbol{\theta}) = \boldsymbol{0}_r$ defining Θ_0. Let $\boldsymbol{d}^* \in \mathbb{R}^r$ and consider the following sequence, $\left\{\boldsymbol{\theta}_n\right\}$, of parameters moving toward $\boldsymbol{\theta}_0$ according to

$$H^*_{1,n} : \boldsymbol{m}(\boldsymbol{\theta}_n) = n^{-1/2}\boldsymbol{d}^*. \tag{22}$$

Note that a Taylor series expansion of $\boldsymbol{m}(\boldsymbol{\theta}_n)$ around $\boldsymbol{\theta}_0$ yields

$$\boldsymbol{m}(\boldsymbol{\theta}_n) = \boldsymbol{m}(\boldsymbol{\theta}_0) + \boldsymbol{M}^T(\boldsymbol{\theta}_0)\big(\boldsymbol{\theta}_n - \boldsymbol{\theta}_0\big) + o\big(\|\boldsymbol{\theta}_n - \boldsymbol{\theta}_0\|\big). \tag{23}$$

Upon substituting $\boldsymbol{\theta}_n = \boldsymbol{\theta}_0 + n^{-1/2}\boldsymbol{d}$ in (23) and taking into account that $\boldsymbol{m}(\boldsymbol{\theta}_0) = \boldsymbol{0}_r$, we get

$$\boldsymbol{m}(\boldsymbol{\theta}_n) = n^{-1/2}\boldsymbol{M}^T(\boldsymbol{\theta}_0)\boldsymbol{d} + o\big(\|\boldsymbol{\theta}_n - \boldsymbol{\theta}_0\|\big),$$

so that the equivalence in the limit is obtained for $\boldsymbol{d}^* = \boldsymbol{M}^T\ (\boldsymbol{\theta}_0)\boldsymbol{d}$.

Theorem 12 *We have the following results under the two forms of the alternative hypothesis:*

i) $W_K\left(\widehat{\boldsymbol{\theta}}_\beta\right) \xrightarrow[n\to\infty]{\mathcal{L}} \chi^2_r(a)$ under $H_{1,n}$ in (21),
ii) $W_K\left(\widehat{\boldsymbol{\theta}}_\beta\right) \xrightarrow[n\to\infty]{\mathcal{L}} \chi^2_r(b)$ under $H^*_{1,n}$ in (22),

where a and b are the non-centrality parameters given by

$$a = \boldsymbol{d}^T\boldsymbol{M}(\boldsymbol{\theta}_0)\big(\boldsymbol{M}^T(\boldsymbol{\theta}_0)\boldsymbol{\Sigma}_\beta(\boldsymbol{\theta}_0)\boldsymbol{M}(\boldsymbol{\theta}_0)\big)^{-1}\boldsymbol{M}^T(\boldsymbol{\theta}_0)\boldsymbol{d},$$
$$b = \boldsymbol{d}^{*T}\big(\boldsymbol{M}^T(\boldsymbol{\theta}_0)\boldsymbol{\Sigma}_\beta(\boldsymbol{\theta}_0)\boldsymbol{M}(\boldsymbol{\theta}_0)\big)^{-1}\boldsymbol{d}^*.$$

Based on Theorem 12, we can give an approximation to the power function at $\boldsymbol{\theta}^*$ because $\boldsymbol{\theta}^* = \boldsymbol{\theta}_0 + n^{-1/2}\boldsymbol{d}$ with $d = \sqrt{n}(\boldsymbol{\theta}^* - \boldsymbol{\theta}_0)$.

4.2 Robustness properties

In order to be able to write in a convenient way the influence function for the WMDPDEs as well as for the Wald-type test statistics, it is necessary to give a new interpretation for ${}^*d^W_\beta(\widehat{\boldsymbol{p}}_i, \boldsymbol{\pi}_i(\boldsymbol{\theta}))$. Let $F_{i,\boldsymbol{\theta}}$ represent the distribution function for each individual from the ith group of K_i individuals, $i = 1, ..., I$, associated with the model, with probability mass function $f_{i,\boldsymbol{\theta}}$. Denoting by $Y_i^{(j)}$ $(j = 1, ..., K_i)$ the random variable having a Bernoulli distribution with probability of success $\pi_{i1}(\boldsymbol{\theta})$, we have

$$f_{i,\boldsymbol{\theta}}(y) = y\pi_{i1}(\boldsymbol{\theta}) + (1-y)\pi_{i2}(\boldsymbol{\theta}), \text{ for } y \in \{0, 1\}.$$

Let G_i represent the true distribution function for each individual from the i th group of K_i individuals, $i = 1, ..., I$, having mass function g_i. In vector notation, we let

$$\underline{K} = (K_1, ..., K_I)^T,$$
$$\boldsymbol{G}_{\underline{K}} = (G_1 \otimes \boldsymbol{1}^T_{K_1}, ..., G_I \otimes \boldsymbol{1}^T_{K_I})^T,$$
$$\boldsymbol{F}_{\underline{K},\boldsymbol{\theta}} = (F_{1,\boldsymbol{\theta}} \otimes \boldsymbol{1}^T_{K_1}, ..., F_{I,\boldsymbol{\theta}} \otimes \boldsymbol{1}^T_{K_I})^T.$$

We first need to define the statistical functional $\boldsymbol{U}_\beta(\boldsymbol{G}_{\underline{K}})$ corresponding to the WMDPDEs as the minimizer of the weighted sum of DPDs between the true and model densities. This is defined as the minimizer of

$$\sum_{i=1}^{I} \frac{K_i}{K} d_\beta(\boldsymbol{g}_i, \boldsymbol{f}_{i,\boldsymbol{\theta}}) \propto \sum_{i=1}^{I} \frac{K_i}{K} \left[\pi_{i1}^{\beta+1}(\boldsymbol{\theta}) + \pi_{i2}^{\beta+1}(\boldsymbol{\theta}) - \frac{\beta+1}{\beta} \left(\frac{N_{i1}}{K_i} \pi_{i1}^{\beta}(\boldsymbol{\theta}) + \frac{N_{i2}}{K_i} \pi_{i2}^{\beta}(\boldsymbol{\theta}) \right) \right], \tag{24}$$

whenever it exists. When the assumption of the model holds with true parameter $\boldsymbol{\theta}_0$, we have $\boldsymbol{g}_i = \boldsymbol{f}_{i,\boldsymbol{\theta}_0}$ and (24) is minimized at $\boldsymbol{\theta} = \boldsymbol{\theta}_0$, implying the Fisher consistency of the WMDPDEs functional $\boldsymbol{U}_\beta(\boldsymbol{G}_{\underline{K}})$ in our model.

The influence function of $\widehat{\boldsymbol{\theta}}_\beta = \boldsymbol{U}_\beta(\widehat{\boldsymbol{G}}_{\underline{K}})$, with respect to the j_0th observation of the i_0th group of observations, is then given by

$$\begin{aligned} \mathcal{IF}\left(i_0, j_0, x, \boldsymbol{U}_\beta, \boldsymbol{F}_{\underline{K},\boldsymbol{\theta}}\right) &= \left.\frac{\partial \boldsymbol{U}_\beta\left(\boldsymbol{F}_{\underline{K},\boldsymbol{\theta},x,\varepsilon}\right)}{\partial \varepsilon}\right|_{\varepsilon=0^+} \\ &= \boldsymbol{J}_\beta^{-1}(\boldsymbol{\theta})(F(IT_i;\boldsymbol{x}_i,\boldsymbol{\theta}) - x) f(IT_i;\boldsymbol{x}_i,\boldsymbol{\theta}) IT_i \boldsymbol{w}_i \\ &\quad \times \left[F^{\beta-1}(IT_i;\boldsymbol{x}_i,\boldsymbol{\theta}) + R^{\beta-1}(IT_i;\boldsymbol{x}_i,\boldsymbol{\theta})\right], \end{aligned} \tag{25}$$

where $\boldsymbol{F}_{\underline{K},\boldsymbol{\theta},x,\varepsilon} = (1-\varepsilon)\boldsymbol{F}_{\underline{K},\boldsymbol{\theta}} + \varepsilon \Delta_x \boldsymbol{e}_{i_0,j_0}$, $\Delta_x(y) = I(x \le y)$ and $\boldsymbol{e}_{i_0,j_0}$ is a vector with 1 in the position (i_0, j_0) in lexicographical order.

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

$$\begin{aligned} \mathcal{IF}\left(\underline{x}, \boldsymbol{U}_\beta, \boldsymbol{F}_{\underline{K},\boldsymbol{\theta}_0}\right) &= \left.\frac{\partial \boldsymbol{U}_\beta\left(\boldsymbol{F}_{\underline{K},\boldsymbol{\theta}_0,\underline{x},\varepsilon}\right)}{\partial \varepsilon}\right|_{\varepsilon=0^+} \\ &= \boldsymbol{J}_\beta^{-1}\left(\boldsymbol{\theta}_0\right) \sum_{i=1}^{I} (K_i F(IT_i;\boldsymbol{x}_i,\boldsymbol{\theta}_0) - x_i) f(IT_i;\boldsymbol{x}_i,\boldsymbol{\theta}) IT_i \boldsymbol{w}_i \\ &\quad \times \left[F^{\beta-1}(IT_i;\boldsymbol{x}_i,\boldsymbol{\theta}_0) + R^{\beta-1}(IT_i;\boldsymbol{x}_i,\boldsymbol{\theta}_0)\right], \end{aligned} \tag{26}$$

where $\boldsymbol{F}_{\underline{K},\boldsymbol{\theta}_0,\underline{x},\varepsilon} = (1-\varepsilon)\boldsymbol{F}_{\underline{K},\boldsymbol{\theta}_0} + \varepsilon \sum_{i=1}^{I} \sum_{j=1}^{K_i} \Delta_{x_{ij}} \boldsymbol{e}_{i,j}$, $\Delta_{x_{ij}}(y) = I(x_{ij} \le y)$ and $x_i = \sum_{j=1}^{K_i} x_{ij}$.

Let us consider the Wald-type test statistic $W_K(\widehat{\boldsymbol{\theta}}_\beta)$ for testing the composite null hypothesis in (18). The influence function of the functional associated with the

Wald-type test statistic $W_K(\widehat{\boldsymbol{\theta}}_\beta) = W_K(\boldsymbol{U}_\beta(\widehat{\boldsymbol{G}}_{\underline{K}})) = T_\beta(\widehat{\boldsymbol{G}}_{\underline{K}})$, with respect to the j_0th observation of the i_0th group of observations, is defined as

$$\mathcal{IF}\left(i_0,j_0,x,T_\beta,\boldsymbol{F}_{\underline{K},\boldsymbol{\theta}_0}\right) = \left.\frac{\partial T_\beta\left(\boldsymbol{F}_{\underline{K},\boldsymbol{\theta}_0,x,\varepsilon}\right)}{\partial\varepsilon}\right|_{\varepsilon=0^+} = 0.$$

Similarly, for all the indices, we have

$$\mathcal{IF}\left(\underline{x},T_\beta,\boldsymbol{F}_{\underline{K},\boldsymbol{\theta}_0}\right) = \left.\frac{\partial T_\beta\left(\boldsymbol{F}_{\underline{K},\boldsymbol{\theta}_0,\underline{x},\varepsilon}\right)}{\partial\varepsilon}\right|_{\varepsilon=0^+} = 0.$$

It is, therefore, necessary to consider the second-order influence function, as presented in the following theorem.

Theorem 13 *The second-order influence function of the functional associated with the Wald-type test statistic $W_K(\widehat{\boldsymbol{\theta}}_\beta) = W_K(\boldsymbol{U}_\beta(\widehat{\boldsymbol{G}}_{\underline{K}})) = T_\beta(\widehat{\boldsymbol{G}}_{\underline{K}})$, with respect to the j_0th observation of the i_0th group of observations, is given by*

$$\begin{aligned}\mathcal{IF}_2\left(i_0,j_0,x,T_\beta,\boldsymbol{F}_{\underline{K},\boldsymbol{\theta}_0}\right) &= \left.\frac{\partial^2 T_\beta\left(\boldsymbol{F}_{\underline{K},\boldsymbol{\theta}_0,x,\varepsilon}\right)}{\partial\varepsilon^2}\right|_{\varepsilon=0^+}\\ &= 2\mathcal{IF}^T\left(i_0,j_0,x,\boldsymbol{U}_\beta,\boldsymbol{F}_{\underline{K},\boldsymbol{\theta}_0}\right)\boldsymbol{M}(\boldsymbol{\theta}_0)\left(\boldsymbol{M}^T(\boldsymbol{\theta}_0)\boldsymbol{\Sigma}(\boldsymbol{\theta}_0)\boldsymbol{M}(\boldsymbol{\theta}_0)\right)^{-1}\\ &\quad\times\boldsymbol{M}^T(\boldsymbol{\theta}_0)\mathcal{IF}\left(i_0,j_0,x,\boldsymbol{U}_\beta,\boldsymbol{F}_{\underline{K},\boldsymbol{\theta}_0}\right),\end{aligned}$$

with $\mathcal{IF}\left(i_0,j_0,x,\boldsymbol{U}_\beta,\boldsymbol{F}_{\underline{K},\boldsymbol{\theta}_0}\right)$ being as in (25). Similarly, for all the indices, we have

$$\begin{aligned}\mathcal{IF}_2\left(\underline{x},T_\beta,\boldsymbol{F}_{\underline{K},\boldsymbol{\theta}_0}\right) &= \left.\frac{\partial^2 T_\beta\left(\boldsymbol{F}_{\underline{K},\boldsymbol{\theta}_0,\underline{x},\varepsilon}\right)}{\partial\varepsilon^2}\right|_{\varepsilon=0^+}\\ &= 2\mathcal{IF}^T\left(\underline{x},\boldsymbol{U}_\beta,\boldsymbol{F}_{\underline{K},\boldsymbol{\theta}_0}\right)\boldsymbol{M}(\boldsymbol{\theta}_0)\left(\boldsymbol{M}^T(\boldsymbol{\theta}_0)\boldsymbol{\Sigma}(\boldsymbol{\theta}_0)\boldsymbol{M}(\boldsymbol{\theta}_0)\right)^{-1}\\ &\quad\times\boldsymbol{M}^T(\boldsymbol{\theta}_0)\mathcal{IF}\left(\underline{x},\boldsymbol{U}_\beta,\boldsymbol{F}_{\underline{K},\boldsymbol{\theta}_0}\right),\end{aligned}$$

with $\mathcal{IF}\left(\underline{x},\boldsymbol{U}_\beta,\boldsymbol{F}_{\underline{K},\boldsymbol{\theta}_0}\right)$ being as in (26).

5 Simulation study

In this section, Monte Carlo simulations of size 2.500 are carried out to examine the behavior of the WMDPDEs discussed in the preceding sections.

Based on the simulation experiment proposed by Balakrishnan and Ling (2014), we consider the devices to have gamma lifetimes, under *four* different conditions with *two* stress factors at *two* levels, taken to be $\{(30,40), (40, 40), (30, 50), (40,50)\}$. Then, all devices under each condition are tested at *three* different inspection times, depending on the reliability considered. The model parameters were set as $(a_1, a_2, b_0, b_1, b_2) = (-0.06, -0.06, -0.36, 0.04, -0.01)$ while $a_0 = 6.5, 7$ or 7.5, corresponding to low, moderate, and high reliability, respectively. In order to study the robustness of the WMDPDEs, we consider a contaminated scheme; wherein, the first "cell" is generated under $\tilde{a}_1 = -0.035$. This will result in a lower number of devices failed in this "outlying cell" than expected under the corresponding model. This idea is similar to the principle of inflated models in distribution theory (see Lambert (1992) and Heilbron (1994)) and was also adopted in the study of one-shot devices by Balakrishnan et al. (2019).

Bias of estimates of reliabilities at normal conditions and different times, as well as the root-mean-square error (RMSE) of the parameter estimates, is computed with

Table 2 Bias of the estimates of reliabilities for pure and contaminated data in the case of low reliability

Low reliability		Pure data				Contaminated data			
	True value	$\beta = 0$	$\beta = 0.2$	$\beta = 0.4$	$\beta = 0.6$	$\beta = 0$	$\beta = 0.2$	$\beta = 0.4$	$\beta = 0.6$
$k = 50$									
$R(10; (25, 30))$	0.8197	– 0.0101	– 0.0085	−0.0135	– 0.3130	0.1216	0.1137	0.1141	– 0.0572
$R(20; (25, 30))$	0.6168	– 0.0037	−0.0058	– 0.0064	−0.2287	0.1162	0.1027	0.0993	– 0.0294
$R(30; (25, 30))$	0.4497	– 0.009	−0.0134	– 0.0114	−0.1688	0.0037	– 0.0009	−0.0063	– 0.0788
$R(40; (25, 30))$	0.3220	– 0.0091	−0.0145	– 0.0109	−0.1214	– 0.0677	−0.0655	– 0.0704	−0.0106
$R(50; (25, 30))$	0.2278	– 0.0037	−0.0092	– 0.0048	−0.0830	– 0.0849	−0.0815	– 0.0852	−0.1013
RMSE(θ)	–	0.9933	0.9737	1.0207	2.2000	1.8496	1.7226	1.7652	1.9626
$k = 100$									
$R(10; (25, 30))$	0.8197	– 0.0055	−0.003	– 0.0056	−0.2749	0.1291	0.1193	0.1216	0.0100
$R(20; (25, 30))$	0.6168	– 0.0027	−0.0035	– 0.0031	−0.2016	0.1229	0.1084	0.106	0.0194
$R(30; (25, 30))$	0.4497	– 0.0055	−0.0088	– 0.0061	−0.1456	0.0099	0.0064	– 0.0007	−0.0498
$R(40; (25, 30))$	0.3220	– 0.0057	−0.0105	– 0.0063	−0.1022	– 0.0700	−0.0648	– 0.0728	−0.0949
$R(50; (25, 30))$	0.2278	– 0.0029	−0.0082	– 0.0033	−0.0686	– 0.0940	−0.0867	– 0.0938	−0.1018
RMSE(θ)	–	0.706	0.6919	0.7174	1.4967	1.7763	1.6122	1.6848	1.6592
$k = 150$									
$R(10; (25, 30))$	0.8197	– 0.0055	−0.0021	– 0.0061	−0.2563	0.1317	0.1214	0.1234	0.0697
$R(20; (25, 30))$	0.6168	– 0.0028	−0.0024	– 0.0034	−0.1887	0.1237	0.1092	0.1063	0.0632
$R(30; (25, 30))$	0.4497	– 0.0039	−0.0064	– 0.0042	−0.1357	0.0120	0.0095	0.0017	– 0.0215
$R(40; (25, 30))$	0.3220	– 0.0037	−0.0081	– 0.0039	−0.0950	– 0.0710	−0.0638	– 0.0730	−0.0810
$R(50; (25, 30))$	0.2278	– 0.0019	−0.0070	– 0.0018	−0.0640	– 0.0991	−0.0893	– 0.0977	−0.0982
RMSE(θ)	–	0.5714	0.5754	0.5785	1.1034	1.7675	1.5892	1.6629	1.5933

Table 3 Bias of the estimates of reliabilities for pure and contaminated data in the case of moderate reliability

Moderate reliability		Pure data				Contaminated data			
	True value	$\beta = 0$	$\beta = 0.2$	$\beta = 0.4$	$\beta = 0.6$	$\beta = 0$	$\beta = 0.2$	$\beta = 0.4$	$\beta = 0.6$
$k = 50$									
$R(40; (25, 30))$	0.5406	− 0.0322	− 0.0331	− 0.0345	− 0.0356	− 0.0965	− 0.0859	− 0.0764	− 0.0716
$R(50; (25, 30))$	0.4449	− 0.0317	− 0.0318	− 0.0335	− 0.0342	− 0.1242	− 0.1108	− 0.0989	− 0.0917
$R(60; (25, 30))$	0.3638	− 0.0255	− 0.0248	− 0.0266	− 0.0269	− 0.1282	− 0.1145	− 0.1024	− 0.0942
$R(70; (25, 30))$	0.2960	− 0.0165	− 0.0151	− 0.0169	− 0.0169	− 0.1195	− 0.1067	− 0.0955	− 0.0872
$R(80; (25, 30))$	0.2399	− 0.0066	− 0.0048	− 0.0064	− 0.0063	− 0.1053	− 0.0936	− 0.0837	− 0.0758
RMSE(θ)	–	1.1827	1.1857	1.2093	1.2265	1.7034	1.5738	1.4878	1.4367
$k = 100$									
$R(40; (25, 30))$	0.5406	− 0.0169	− 0.0177	− 0.0174	− 0.0184	− 0.0774	− 0.0666	− 0.0552	− 0.0512
$R(50; (25, 30))$	0.4449	− 0.0189	− 0.0193	− 0.0191	− 0.0208	− 0.1177	− 0.1028	− 0.0875	− 0.0811
$R(60; (25, 30))$	0.3638	− 0.0173	− 0.0174	− 0.0171	− 0.0191	− 0.1330	− 0.1175	− 0.1016	− 0.0940
$R(70; (25, 30))$	0.2960	− 0.0132	− 0.0130	− 0.0127	− 0.0147	− 0.1319	− 0.1177	− 0.1028	− 0.0951
$R(80; (25, 30))$	0.2399	− 0.0080	− 0.0074	− 0.0072	− 0.0091	− 0.1220	− 0.1096	− 0.0965	− 0.0892
RMSE(θ)	–	0.8056	0.8074	0.8144	0.8344	1.4918	1.3356	1.2141	1.1558
$k = 150$									
$R(40; (25, 30))$	0.5406	− 0.0128	− 0.0128	− 0.0131	− 0.0132	− 0.0713	− 0.0592	− 0.0483	− 0.0432
$R(50; (25, 30))$	0.4449	− 0.0143	− 0.0138	− 0.0147	− 0.0149	− 0.1160	− 0.0988	− 0.0836	− 0.0754
$R(60; (25, 30))$	0.3638	− 0.0133	− 0.0124	− 0.0138	− 0.0139	− 0.1356	− 0.1175	− 0.1012	− 0.0917
$R(70; (25, 30))$	0.2960	− 0.0105	− 0.0092	− 0.0109	− 0.0109	− 0.1374	− 0.1209	− 0.1053	− 0.0959
$R(80; (25, 30))$	0.2399	− 0.0067	− 0.0052	− 0.0069	− 0.0068	− 0.1289	− 0.1148	− 0.1010	− 0.0922
RMSE(θ)	–	0.6769	0.6788	0.6903	0.7028	1.4402	1.2655	1.1330	1.0542

the same sample size for each condition $K = \{50, 100, 150\}$, and those are presented in Tables 2, 3, and 4.

It can be seen that, while for the non-contaminated scheme, the MLE generally possesses the best behavior, WMDPDEs with medium β are a better option in the contamination scenario. This robustness is in accordance with the earlier finding of Balakrishnan et al. (2019) for the case of one-shot device testing based on exponential lifetimes.

5.1 Wald-type tests

Let us now empirically evaluate the robustness of the WMDPDEs-based Wald-type tests developed in Sect. 4. The simulation is performed under the low-reliability model described before.

We first study the observed level (measured as the proportion of test statistics exceeding the corresponding Chi-square critical value) of the test under the true null hypothesis $H_0 : a_1 = -0.06$ against the alternative $H_1 : a_1 \neq -0.06$. In the top of Figure 1, these levels are plotted for different values of the samples

Table 4 Bias of the estimates of reliabilities for pure and contaminated data in the case of high reliability

High reliability		Pure data				Contaminated data			
	True value	$\beta=0$	$\beta=0.2$	$\beta=0.4$	$\beta=0.6$	$\beta=0$	$\beta=0.2$	$\beta=0.4$	$\beta=0.6$
$k=50$									
$R(70;(25,30))$	0.5157	– 0.0587	– 0.0551	– 0.0580	– 0.0623	– 0.1591	– 0.1374	– 0.1259	– 0.1101
$R(80;(25,30))$	0.4581	– 0.0493	– 0.0458	– 0.0496	– 0.0554	– 0.1530	– 0.1326	– 0.1214	– 0.1067
$R(90;(25,30))$	0.4059	– 0.0385	– 0.0349	– 0.0395	– 0.0465	– 0.1431	– 0.1240	– 0.1133	– 0.1000
$R(100;(25,30))$	0.3590	– 0.0270	– 0.0233	– 0.0284	– 0.0364	– 0.1311	– 0.1131	– 0.1030	– 0.0912
$R(110;(25,30))$	0.3169	– 0.0154	– 0.0116	– 0.0170	– 0.0258	– 0.1181	– 0.1012	– 0.0915	– 0.0811
RMSE(θ)	–	1.7033	1.7030	1.7033	1.6846	1.8756	1.7491	1.7106	1.5983
$k=100$									
$R(70;(25,30))$	0.5157	– 0.0451	– 0.0463	– 0.0492	– 0.0535	– 0.1652	– 0.1405	– 0.1243	– 0.1107
$R(80;(25,30))$	0.4581	– 0.0421	– 0.0441	– 0.0475	– 0.0525	– 0.1677	– 0.1438	– 0.1278	– 0.1157
$R(90;(25,30))$	0.4059	– 0.0369	– 0.0396	– 0.0433	– 0.0489	– 0.1643	– 0.1416	– 0.1263	– 0.1158
$R(100;(25,30))$	0.3590	– 0.0304	– 0.0335	– 0.0374	– 0.0434	– 0.1569	– 0.1356	– 0.1213	– 0.1123
$R(110;(25,30))$	0.3169	– 0.0231	– 0.0265	– 0.0305	– 0.0367	– 0.1471	– 0.1273	– 0.1141	– 0.1063
RMSE(θ)	–	1.1432	1.1406	1.1653	1.1651	1.5383	1.3921	1.3102	1.2219
$k=150$									
$R(70;(25,30))$	0.5157	– 0.0376	– 0.0393	– 0.0419	– 0.0485	– 0.1639	– 0.1379	– 0.1194	– 0.1080
$R(80;(25,30))$	0.4581	– 0.0369	– 0.0391	– 0.0420	– 0.0497	– 0.1705	– 0.1458	– 0.1273	– 0.1164
$R(90;(25,30))$	0.4059	– 0.0341	– 0.0367	– 0.0398	– 0.0483	– 0.1702	– 0.1474	– 0.1295	– 0.1195
$R(100;(25,30))$	0.3590	– 0.0298	– 0.0326	– 0.0359	– 0.0448	– 0.1651	– 0.1444	– 0.1275	– 0.1184
$R(110;(25,30))$	0.3169	– 0.0246	– 0.0275	– 0.0307	– 0.0400	– 0.1569	– 0.1382	– 0.1225	– 0.1144
RMSE(θ)	–	0.9471	0.9525	0.9749	0.9772	1.4014	1.2390	1.1308	1.0681

sizes, pure data (left), and contaminated data ($\tilde{a}_1 = -0.035$, right). Notice that in the case of pure data considered, all the observed levels are close to the nominal level of 0.05. In the case of contaminated data, the level of the classical Wald test (at $\beta = 0$) displays a lack of robustness, while the WMDPDEs-based Wald-type tests for moderate and large positive β possess levels closer to the nominal level.

To investigate the power of these tests (obtained in a similar manner), we change the true data generating parameters value to $\boldsymbol{\theta} = (6.5, -0.06, -0.035, -0.36, 0.04, -0.01)$, and $\tilde{a}_1 = -0.45$ in a contaminated scenario, nearer to the null hypothesis. The resulting empirical powers are plotted in the bottom of Figure 1. When there are no outliers in the data, the classical Wald test (at $\beta = 0$) is quite similar, not even the most powerful, to other tests. On the other hand, when there are outliers in the data, the Wald-type test with larger $\beta > 0$ provides a significantly better power.

5.2 On the choice of the tuning parameter

Although the robustness of the proposed WMDPDEs seems to increase with increasing β, their efficiency in case of pure data decreases slightly. Same happens with the proposed Wald-type tests. A moderate value of β is expected to

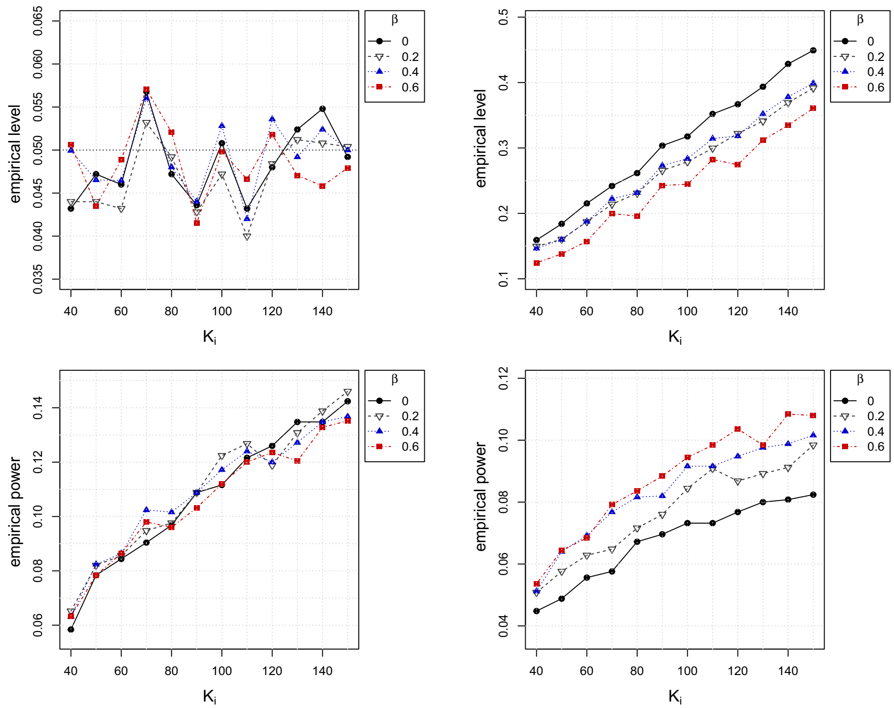

Fig. 1 Levels and powers of WMDPDEs for pure (left) and contaminated data (right)

provide the best trade-off for possibly contaminated data, but a data-driven choice of β would be more convenient. An useful procedure for the data-based selection of β was proposed by Warwick and Jones (2005). It consists of minimizing the estimated mean-squared error (MSE), and an approach that requires pilot estimation of model parameters. In this approach, we minimize an estimate of the asymptotic MSE of the WMDPDEs $\widehat{\boldsymbol{\alpha}}_\beta$ given by

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

Warwick and Jones (2005) 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.

As proposed by Basu et al. (2017), when dealing with the robustness issue, the estimation of the variance component should not assume the model to be true. See, for example, Balakrishnan et al. (2019), where model robust estimates of the variance components are computed for the exponential distribution under simple stress model. However, in the case of gamma distribution with multiple stresses, where the reliability function does not have a closed-form and require numerical integration, the corresponding formulas become tedious and computationally difficult (see equations (24), (25), and (26) of Balakrishnan and Ling 2014).

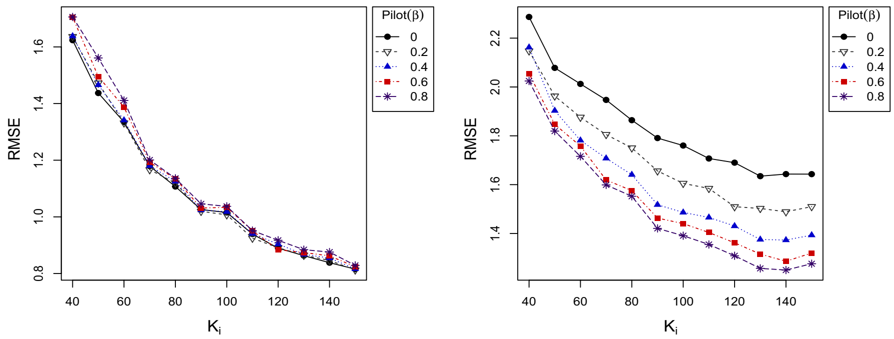

Fig. 2 Simulated MSEs of the WMDPDEs at the optimally chosen β, starting from different pilot estimators under the pure data (left) and the contaminated data (right)

Let us now consider the previous simulation study for moderate reliability and pure and contaminated schemes. We now perform the selection of β following the above proposal for each iteration with different possible pilot estimators. Let us consider as potential pilot parameters $\beta^P = \{0.0.2, 0.4, 0.6, 0.8\}$. The selection of β is done through a grid search of [0, 1] with length 50. In Figure 2, we show the simulated true RMSEs for these pilot parameters. The best trade-off between the efficiency in pure data and the robustness under contaminated data is provided by the pilot choice $\beta^P = 0.6$, so we suggest to use the pilot choice $\boldsymbol{\alpha}^P = \widehat{\boldsymbol{\alpha}}_{0.6}$ when using this procedure.

6 Application to a tumor toxicological data

Survival analysis usually faces problems associated with interval censoring. One extreme situation is the one in which the only available information on a survival variable is whether or not it exceeds a monitoring time. This form of censoring, known as current status data, can be seen as one-shot device testing data, and so we can apply the methods developed in the preceding sections to a real current status data from a tumor toxicological experiment.

The data considered, taken from the National Center for Toxicological Research, were originally reported by Kodell and Nelson (1980) and recently analyzed by Balakrishnan and Ling (2013, 2014) using MLE under a one-shot device model. Balakrishnan et al. (2019) analyzed these data using density power divergences, but under the assumption of exponential lifetimes. However, the gamma distribution is a better lifetime model for these data (Balakrishnan and Ling 2014). These data consist of 1816 mice, of which 553 had tumors, involving the strain of offspring (F1 or F2), gender (females or males), and concentration of benzidine dihydrochloride (60 ppm, 120 ppm, 200 ppm, or 400 ppm) as the stress factors. For each testing condition, the numbers of mice tested and the numbers of mice having tumors were all recorded.

Let a_1, a_2, and a_3 denote the parameters corresponding to the covariates of strain of offspring, gender, and square root of concentration of the chemical of benzidine

Table 5 WMDPDEs of the model parameters

β	a_0	a_1	a_2	a_3	b_0	b_1	b_2	b_3	$\mathrm{MAB}(p(\hat{\boldsymbol{\theta}}))$	$\mathrm{RMSE}(p(\hat{\boldsymbol{\theta}}))$
0	2.4066	−0.1875	−1.0099	0.0359	0.8730	0.2419	1.5545	−0.0901	0.2758	0.3950
0.1	2.8958	−0.1743	−1.2198	0.0136	0.3678	0.2196	1.7680	−0.0670	0.2701	0.3931
0.2	2.8644	−0.1477	−1.3185	0.0219	0.4050	0.1885	1.8742	−0.0758	0.2667	0.3926
0.3	2.7834	−0.1375	−1.3920	0.0332	0.4935	0.1756	1.9535	−0.0877	0.2638	0.3922
0.4	2.6980	−0.1275	−1.4569	0.0443	0.5847	0.1635	2.0217	−0.0994	0.2616	0.3919
0.5	2.6343	−0.1205	−1.5071	0.0529	0.6517	0.1549	2.0732	−0.1082	0.2601	0.3917
0.6	2.5965	−0.1189	−1.5404	0.0585	0.6912	0.1525	2.1067	−0.1139	0.2593	0.3917
0.7	2.5758	−0.1219	−1.5603	0.0619	0.7126	0.1551	2.1263	−0.1173	0.2590	0.3918
0.8	2.5636	−0.1267	−1.5716	0.0640	0.7253	0.1598	2.1370	−0.1194	0.2588	0.3919
0.9	2.5597	−0.1328	−1.5771	0.0651	0.7293	0.1659	2.1420	−0.1205	0.2588	0.3920
1	2.5570	−0.1384	−1.5787	0.0658	0.7321	0.1716	2.1431	−0.1212	0.2588	0.3921

Table 6 WMDPDEs of the mean time to the occurrence of tumors (in months), $\widehat{E}[T]$

Strain	Gender	Conc	$\widehat{E}[T]$					
			$\beta = 0$	$\beta = 0.2$	$\beta = 0.4$	$\beta = 0.6$	$\beta = 0.8$	$\beta = 1$
0	0	60	17.461	17.316	17.395	17.433	17.449	17.455
0	1	60	30.103	30.188	30.598	30.712	30.713	30.694
1	0	60	18.438	18.039	18.031	18.029	18.036	18.044
1	1	60	31.787	31.447	31.719	31.762	31.747	31.728
0	0	120	14.676	14.565	14.577	14.594	14.605	14.610
0	1	120	25.301	25.392	25.643	25.710	25.707	25.691
1	0	120	15.496	15.173	15.111	15.092	15.096	15.103
1	1	120	26.715	26.451	26.582	26.588	26.572	26.557
0	0	200	12.348	12.265	12.231	12.231	12.238	12.243
0	1	200	21.288	21.381	21.514	21.547	21.541	21.529
1	0	200	13.039	12.776	12.678	12.649	12.650	12.656
1	1	200	22.478	22.273	22.302	22.283	22.266	22.254
0	0	400	8.991	8.942	8.858	8.840	8.843	8.848
0	1	400	15.500	15.589	15.583	15.574	15.566	15.558
1	0	400	9.493	9.315	9.183	9.142	9.141	9.146
1	1	400	16.366	16.240	16.153	16.106	16.090	16.082

dihydrochloride in the shape parameter of the gamma distribution, while b_1, b_2, and b_3 denote similarly for the scale parameter, respectively. The WMDPDEs of model parameters as well as of the mean time to occurrence of tumors for each group, for different values of $\beta \in \{0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1\}$, are computed, and these are presented in Tables 5 and 6, respectively, where strain= 0 for F1 strain of offspring, and gender= 0 for females.

There is a significant difference between genders, with males having a higher expected lifetime. Also, tumors are induced by an increase in the dosage of benzidine dihydrochloride. Empirical mean absolute error (MAB) and root of mean square error (RMSE), measured by comparing predicted probabilities to the observed ones, are also presented in Table 5. In both cases, maximum likelihood estimator presents the maximum error. Applying the ad hoc data-driven procedure presented in Sect. 5.2, we obtain $\beta = 0.2$ as the optimal tuning parameter. As an illustration of the proposed Wald-type test for this dataset, we consider the problem of testing $H_0 : a_2 = 0$. The p value obtained based on the classical test ($\beta = 0$) is $p_0 \approx 0.08$, while the p value obtained based on the Wald-type test with $\beta = 0.2$ is $p_{0.2} \approx 0.0004$, and so the test decision at significance level $\alpha = 0.05$ clearly changes with our choice.

7 Concluding remarks

In this paper, we have developed the WMDPDEs for one-shot device testing data under the gamma lifetime model. Through a simulation study, these estimators, as well as the Wald-type tests derived from them, are shown to be good alternatives to the classical MLE and tests based on them in terms of robustness. These estimators are finally applied to a tumor toxicological dataset, for the purpose of illustration.

Another problem that will be of practical interest here is in the application of minimum density power divergence estimators to one-shot device testing data when there are competing risks for failure. Work in this direction is currently under progress, and we hope to report these findings in a future paper.

Acknowledgements We would like to thank the editor, the associate editor, and the referees for their helpful comments and suggestions. This research is partially supported by Grant MTM2015-67057-P and Grant FPU16/03104 from Ministerio de Economia y Competitividad and Ministerio de Educacion, Cultura y Deporte (Spain).

Compliance with ethical standards

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest

Proofs of the main results

Proof of Theorem 1

From (8), we have

$$
\begin{aligned}
d_{KL}(\widehat{\boldsymbol{p}},\boldsymbol{p}(\boldsymbol{\theta})) &= \sum_{i=1}^{I}\sum_{j=1}^{2}\frac{n_{ij}}{K}\log\left(\frac{\frac{n_{ij}}{K_i}}{F_j(IT_i;\boldsymbol{x}_i,\boldsymbol{\theta})}\right)\\
&= \sum_{i=1}^{I}\sum_{j=1}^{2}\frac{n_{ij}}{K}\log\left(\frac{n_i}{K_i}\right)-\sum_{i=1}^{I}\sum_{j=1}^{2}\frac{n_{ij}}{K}\log\left(F_j(IT_i;\boldsymbol{x}_i,\boldsymbol{\theta})\right)\\
&= c-\frac{1}{K}\sum_{i=1}^{I}\left\{n_i\log\left(F(IT_i;\boldsymbol{x}_i,\boldsymbol{\theta})\right)+(K_i-n_i)\log\left(R(IT_i;\boldsymbol{x}_i,\boldsymbol{\theta})\right)\right\}\\
&= c-\frac{1}{K}\log\left(\prod_{i=1}^{I}F(IT_i;\boldsymbol{x}_i,\boldsymbol{\theta})^{n_i}R^{K_i-n_i}(IT_i;\boldsymbol{x}_i,\boldsymbol{\theta})\right)\\
&= c-\frac{1}{K}\log\mathcal{L}(\boldsymbol{\theta}),
\end{aligned}
$$

where $c = \sum_{i=1}^{I} \sum_{j=1}^{2} \frac{n_{ij}}{K} \log\left(\frac{n_i}{K_i}\right)$ does not depend on the vector of parameters $\boldsymbol{a}$.

Proof of Theorem 6

Since

$$\frac{\partial}{\partial \boldsymbol{\theta}} \sum_{i=1}^{I} \frac{K_i}{K} d_\beta^*(\widehat{\boldsymbol{p}}_i, \boldsymbol{\pi}_i(\boldsymbol{\theta})) = \sum_{i=1}^{I} \frac{K_i}{K} \frac{\partial}{\partial \boldsymbol{\theta}} d_\beta^*(\widehat{\boldsymbol{p}}_i, \boldsymbol{\pi}_i(\boldsymbol{\theta}))$$

and

$$\begin{aligned}
&\frac{\partial}{\partial \boldsymbol{\theta}} d_\beta^*(\widehat{\boldsymbol{p}}_i, \boldsymbol{\pi}_i(\boldsymbol{\theta})) \\
&\quad = \left(\frac{\partial}{\partial \boldsymbol{\theta}} \pi_{i1}^{\beta+1}(\boldsymbol{\theta}) + \frac{\partial}{\partial \boldsymbol{\theta}} \pi_{i2}^{\beta+1}(\boldsymbol{\theta})\right) - \frac{\beta+1}{\beta}\left(\widehat{p}_{i1} \frac{\partial}{\partial \boldsymbol{\theta}} \pi_{i1}^{\beta}(\boldsymbol{\theta}) + \widehat{p}_{i2} \frac{\partial}{\partial \boldsymbol{\theta}} \pi_{i2}^{\beta}(\boldsymbol{\theta})\right) \\
&\quad = (\beta+1)\left(\pi_{i1}^{\beta}(\boldsymbol{\theta}) - \pi_{i2}^{\beta}(\boldsymbol{\theta}) - \widehat{p}_{i1} \pi_{i1}^{\beta-1}(\boldsymbol{\theta}) + \widehat{p}_{i2} \pi_{i2}^{\beta-1}(\boldsymbol{\theta})\right) \frac{\partial}{\partial \boldsymbol{\theta}} \pi_{i1}(\boldsymbol{\theta}) \\
&\quad = (\beta+1)\Big(\big(\pi_{i1}(\boldsymbol{\theta}) - \widehat{p}_{i1}\big) \pi_{i1}^{\beta-1}(\boldsymbol{\theta}) - \big(\pi_{i2}(\boldsymbol{\theta}) - \widehat{p}_{i2}\big) \pi_{i2}^{\beta-1}(\boldsymbol{\theta})\Big) \frac{\partial}{\partial \boldsymbol{\theta}} \pi_{i1}(\boldsymbol{\theta}) \\
&\quad = (\beta+1)\Big(\big(\pi_{i1}(\boldsymbol{\theta}) - \widehat{p}_{i1}\big) \pi_{i1}^{\beta-1}(\boldsymbol{\theta}) + \big(\pi_{i1}(\boldsymbol{\theta}) - \widehat{p}_{i1}\big) \pi_{i2}^{\beta-1}(\boldsymbol{\theta})\Big) \frac{\partial}{\partial \boldsymbol{\theta}} \pi_{i1}(\boldsymbol{\theta}) \\
&\quad = (\beta+1)\big(\pi_{i1}(\boldsymbol{\theta}) - \widehat{p}_{i1}\big)\left(\pi_{i1}^{\beta-1}(\boldsymbol{\theta}) + \pi_{i2}^{\beta-1}(\boldsymbol{\theta})\right) \frac{\partial}{\partial \boldsymbol{\theta}} \pi_{i1}(\boldsymbol{\theta}) \\
&\quad = (\beta+1)\big(\pi_{i1}(\boldsymbol{\theta}) - \widehat{p}_{i1}\big)\left(\pi_{i1}^{\beta-1}(\boldsymbol{\theta}) + \pi_{i2}^{\beta-1}(\boldsymbol{\theta})\right) f\big(IT_i; \boldsymbol{x}_i, \boldsymbol{\theta}\big) \frac{\partial f\big(IT_i; \boldsymbol{x}_i, \boldsymbol{\theta}\big)}{\partial \boldsymbol{\theta}},
\end{aligned}$$

we obtain

$$\begin{aligned}
\frac{\partial}{\partial \boldsymbol{\theta}} \sum_{i=1}^{I} \frac{K_i}{K} d_\beta^*(\widehat{\boldsymbol{p}}_i, \boldsymbol{\pi}_i(\boldsymbol{\theta})) = \frac{\beta+1}{K} \sum_{i=1}^{I} \big(K_i \pi_{i1}(\boldsymbol{\theta}) - n_i\big)\left(\pi_{i1}^{\beta-1}(\boldsymbol{\theta}) + \pi_{i2}^{\beta-1}(\boldsymbol{\theta})\right) \\
f\big(IT_i; \boldsymbol{x}_i, \boldsymbol{\theta}\big) \frac{\partial f\big(IT_i; \boldsymbol{x}_i, \boldsymbol{\theta}\big)}{\partial \boldsymbol{\theta}}.
\end{aligned}$$

In a similar way, we can get the derivative with respect to $\boldsymbol{b}$, and then, the required results follow.

Proof of Theorem 7

Let us denote

$$\boldsymbol{u}_{ij}(\boldsymbol{\theta}) = \left(\frac{\partial \log \pi_{ij}(\boldsymbol{\theta})}{\partial \boldsymbol{a}}, \frac{\partial \log \pi_{ij}(\boldsymbol{\theta})}{\partial \boldsymbol{b}}\right)^T = \left(\frac{1}{\pi_{ij}(\boldsymbol{\theta})}\frac{\partial \pi_{ij}(\boldsymbol{\theta})}{\partial \boldsymbol{a}}, \frac{1}{\pi_{ij}(\boldsymbol{\theta})}\frac{\partial \pi_{ij}(\boldsymbol{\theta})}{\partial \boldsymbol{b}}\right)^T$$
$$= \left(\frac{(-1)^{j+1}}{\pi_{ij}(\boldsymbol{\theta})} l_i \boldsymbol{x}_i, \frac{(-1)^{j+1}}{\pi_{ij}(\boldsymbol{\theta})} s_i \boldsymbol{x}_i\right)^T,$$

with

$$l_i = \alpha_i \left\{ -\Psi\left(\alpha_i\right)\pi_{i1}(\boldsymbol{\theta}) + \log\left(\frac{IT_i}{\lambda_i}\right)\pi_{i1}(\boldsymbol{\theta}) - \frac{\left(\frac{IT_i}{\lambda_i}\right)^{\alpha_i}}{\alpha_i^2 \Gamma(\alpha_i)} {}_2F_2\left(\alpha_i, \alpha_i; 1+\alpha_i, 1+\alpha_i; -\frac{IT_i}{\lambda_i}\right)\right\} \tag{27}$$

and

$$s_i = -f\left(IT_i; \boldsymbol{x}_i, \boldsymbol{\theta}\right) IT_i; \tag{28}$$

see Balakrishnan and Ling (2014) for more details.

Upon using Theorem 3.1 of Ghosh and Basu (2013), we have

$$\sqrt{K}\left(\widehat{\boldsymbol{\theta}}_\beta - \boldsymbol{\theta}_0\right) \xrightarrow[K\to\infty]{\mathcal{L}} \mathcal{N}\left(\boldsymbol{0}_{2J}, \boldsymbol{J}_\beta^{-1}(\boldsymbol{\theta}_0)\boldsymbol{K}_\beta(\boldsymbol{\theta}_0)\boldsymbol{J}_\beta^{-1}(\boldsymbol{\theta}_0)\right),$$

where

$$\boldsymbol{J}_\beta(\boldsymbol{\theta}) = \sum_{i=1}^{I}\sum_{j=1}^{2}\frac{K_i}{K}\boldsymbol{u}_{ij}(\boldsymbol{\theta})\boldsymbol{u}_{ij}^T(\boldsymbol{\theta})\pi_{ij}^{\beta+1}(\boldsymbol{\theta}),$$
$$\boldsymbol{K}_\beta(\boldsymbol{\theta}) = \left(\sum_{i=1}^{I}\sum_{j=1}^{2}\frac{K_i}{K}\boldsymbol{u}_{ij}(\boldsymbol{\theta})\boldsymbol{u}_{ij}^T(\boldsymbol{\theta})\pi_{ij}^{2\beta+1}(\boldsymbol{\theta}) - \sum_{i=1}^{I}\frac{K_i}{K}\boldsymbol{\xi}_{i,\beta}(\boldsymbol{\theta})\boldsymbol{\xi}_{i,\beta}^T(\boldsymbol{\theta})\right),$$

with

$$\boldsymbol{\xi}_{i,\beta}(\boldsymbol{\theta}) = \sum_{j=1}^{2}\boldsymbol{u}_{ij}(\boldsymbol{\theta})\pi_{ij}^{\beta+1}(\boldsymbol{\theta})$$
$$= \left(l_i\boldsymbol{x}_i, s_i\boldsymbol{x}_i\right)^T \sum_{j=1}^{2}(-1)^{j+1}\pi_{ij}^{\beta}(\boldsymbol{\theta}).$$

Now, for $\boldsymbol{u}_{ij}(\boldsymbol{\theta})\boldsymbol{u}_{ij}^T(\boldsymbol{\theta})$, we have

$$\boldsymbol{u}_{ij}(\boldsymbol{\theta})\boldsymbol{u}_{ij}^T(\boldsymbol{\theta}) = \frac{1}{\pi_{ij}^2(\boldsymbol{\theta})}\begin{pmatrix} l_i^2\boldsymbol{x}_i^T\boldsymbol{x}_i & l_is_i\boldsymbol{x}_i^T\boldsymbol{x}_i \\ l_is_i\boldsymbol{x}_i^T\boldsymbol{x}_i & s_i^2\boldsymbol{x}_i^T\boldsymbol{x}_i \end{pmatrix} = \frac{1}{\pi_{ij}^2(\boldsymbol{\theta})}\boldsymbol{M}_i,$$

with

$$\boldsymbol{M}_i = \begin{pmatrix} l_i^2\boldsymbol{x}_i^T\boldsymbol{x}_i & l_is_i\boldsymbol{x}_i^T\boldsymbol{x}_i \\ l_is_i\boldsymbol{x}_i^T\boldsymbol{x}_i & s_i^2\boldsymbol{x}_i^T\boldsymbol{x}_i \end{pmatrix}. \tag{29}$$

It then follows that

$$\begin{aligned}\boldsymbol{J}_\beta(\boldsymbol{\theta}) &= \sum_{i=1}^{I}\frac{K_i}{K}\boldsymbol{M}_i\sum_{j=1}^{2}\pi_{ij}^{\beta-1}(\boldsymbol{\theta}) \\ &= \sum_{i=1}^{I}\frac{K_i}{K}\boldsymbol{M}_i\left(\pi_{i1}^{\beta-1}(\boldsymbol{\theta}) + \pi_{i2}^{\beta-1}(\boldsymbol{\theta})\right).\end{aligned}$$

In a similar manner,

$$\boldsymbol{\xi}_{i,\beta}(\boldsymbol{\theta})\boldsymbol{\xi}_{i,\beta}^T(\boldsymbol{\theta}) = \boldsymbol{M}_i\left(\sum_{j=1}^{2}(-1)^{j+1}\pi_{ij}^{\beta}(\boldsymbol{\theta})\right)^2$$

and

$$\boldsymbol{K}_\beta(\boldsymbol{\theta}) = \sum_{i=1}^{I}\frac{K_i}{K}\boldsymbol{M}_i\left(\sum_{j=1}^{2}\pi_{ij}^{2\beta-1}(\boldsymbol{\theta}) - \left(\sum_{j=1}^{2}(-1)^{j+1}\pi_{ij}^{\beta}(\boldsymbol{\theta})\right)^2\right).$$

Since

$$\sum_{j=1}^{2}\pi_{ij}^{2\beta-1}(\boldsymbol{\theta}) - \left(\sum_{j=1}^{2}(-1)^{j+1}\pi_{ij}^{\beta}(\boldsymbol{\theta})\right)^2 = \pi_{i1}(\boldsymbol{\theta})\pi_{i2}(\boldsymbol{\theta})\left(\pi_{i1}^{\beta-1}(\boldsymbol{\theta}) + \pi_{i2}^{\beta-1}(\boldsymbol{\theta})\right)^2,$$

we have

$$\boldsymbol{K}_\beta(\boldsymbol{\theta}) = \sum_{i=1}^{I}\frac{K_i}{K}\boldsymbol{M}_i\pi_{i1}(\boldsymbol{\theta})\pi_{i2}(\boldsymbol{\theta})\left(\pi_{i1}^{\beta-1}(\boldsymbol{\theta}) + \pi_{i2}^{\beta-1}(\boldsymbol{\theta})\right)^2.$$

Proof of Theorem 9

Let $\boldsymbol{\theta}_0 \in \Theta$ be the true value of parameter $\boldsymbol{\theta}$. It is clear that

$$\begin{aligned} \boldsymbol{m}\left(\widehat{\boldsymbol{\theta}}_{\beta}\right) &= \boldsymbol{m}(\boldsymbol{\theta}_0) + \boldsymbol{M}^T\left(\widehat{\boldsymbol{\theta}}_{\beta}\right)\left(\widehat{\boldsymbol{\theta}}_{\beta} - \boldsymbol{\theta}_0\right) + o_p\left(\left\|\widehat{\boldsymbol{\theta}}_{\beta} - \boldsymbol{\theta}_0\right\|\right) \\ &= \boldsymbol{M}^T\left(\widehat{\boldsymbol{\theta}}_{\beta}\right)\left(\widehat{\boldsymbol{\theta}}_{\beta} - \boldsymbol{\theta}_0\right) + o_p(K^{-1/2}). \end{aligned}$$

But, under H_0,

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

Therefore, under H_0,

$$\sqrt{K}\boldsymbol{m}\left(\widehat{\boldsymbol{\theta}}_{\beta}\right) \xrightarrow[K\to\infty]{\mathcal{L}} \mathcal{N}\left(\mathbf{0}_r, \boldsymbol{M}^T(\boldsymbol{\theta}_0)\boldsymbol{\Sigma}_{\beta}\left(\widehat{\boldsymbol{\theta}}_{\beta}\right)\boldsymbol{M}(\boldsymbol{\theta}_0)\right)$$

and taking into account that $rank(\boldsymbol{M}(\boldsymbol{\theta}_0)) = r$, we get

$$K\boldsymbol{m}\left(\widehat{\boldsymbol{\theta}}_{\beta}\right)^T(\boldsymbol{M}^T(\boldsymbol{\theta}_0)\boldsymbol{\Sigma}_{\beta}(\boldsymbol{\theta}_0)\boldsymbol{M}(\boldsymbol{\theta}_0))^{-1}\boldsymbol{m}\left(\widehat{\boldsymbol{\theta}}_{\beta}\right) \xrightarrow[K\to\infty]{\mathcal{L}} \chi^2_r.$$

But, $\left(\boldsymbol{M}\left(\widehat{\boldsymbol{\theta}}_{\beta}\right)^T\boldsymbol{\Sigma}_{\beta}\left(\widehat{\boldsymbol{\theta}}_{\beta}\right)\boldsymbol{M}\left(\widehat{\boldsymbol{\theta}}_{\beta}\right)\right)^{-1}$ is a consistent estimator of $\left(\boldsymbol{M}(\boldsymbol{\theta}_0)^T\boldsymbol{\Sigma}_{\beta}(\boldsymbol{\theta}_0)\boldsymbol{M}(\boldsymbol{\theta}_0)\right)^{-1}$. Therefore,

$$W_K\left(\widehat{\boldsymbol{\theta}}_{\beta}\right) \xrightarrow[K\to\infty]{\mathcal{L}} \chi^2_r.$$

Proof of Theorem 10

Under the assumption that

$$\widehat{\boldsymbol{\theta}}_{\beta} \xrightarrow[K\to\infty]{P} \boldsymbol{\theta}^*,$$

the asymptotic distribution of $\ell_{\beta}\left(\widehat{\boldsymbol{\theta}}_1, \widehat{\boldsymbol{\theta}}_2\right)$ coincides with the asymptotic distribution of $\ell_{\beta}\left(\widehat{\boldsymbol{\theta}}_1, \boldsymbol{\theta}^*\right)$. A first-order Taylor expansion of $\ell_{\beta}\left(\widehat{\boldsymbol{\theta}}_{\beta}, \boldsymbol{\theta}\right)$ at $\widehat{\boldsymbol{\theta}}_{\beta}$, around $\boldsymbol{\theta}^*$, gives

$$\left(\ell_{\beta}\left(\widehat{\boldsymbol{\theta}}_{\beta}, \boldsymbol{\theta}^*\right) - \ell_{\beta}(\boldsymbol{\theta}^*, \boldsymbol{\theta}^*)\right) = \left.\frac{\partial \ell_{\beta}(\boldsymbol{\theta}, \boldsymbol{\theta}_*)}{\partial \boldsymbol{\theta}^T}\right|_{\boldsymbol{\theta}=\boldsymbol{\theta}_*}\left(\widehat{\boldsymbol{\theta}}_{\beta} - \boldsymbol{\theta}^*\right) + o_p(K^{-1/2}).$$

Now, the result follows since

$$\sqrt{K}\left(\widehat{\boldsymbol{\theta}}_{\beta} - \boldsymbol{\theta}^*\right) \xrightarrow[K\to\infty]{\mathcal{L}} \mathcal{N}(\mathbf{0}_{J+1}, \boldsymbol{\Sigma}_{\beta}(\boldsymbol{\theta}^*)).$$

Proof of Theorem 12

A Taylor series expansion of $\boldsymbol{m}(\widehat{\boldsymbol{\theta}}_\beta)$ around $\boldsymbol{\theta}_n$ yields

$$\boldsymbol{m}(\widehat{\boldsymbol{\theta}}_\beta) = \boldsymbol{m}(\boldsymbol{\theta}_n) + \boldsymbol{M}^T(\boldsymbol{\theta}_n)(\widehat{\boldsymbol{\theta}}_\beta - \boldsymbol{\theta}_n) + o\left(\left\|\widehat{\boldsymbol{\theta}}_\beta - \boldsymbol{\theta}_n\right\|\right).$$

From (23), we have

$$\boldsymbol{m}(\widehat{\boldsymbol{\theta}}_\beta) = K^{-1/2}\boldsymbol{M}^T(\boldsymbol{\theta}_0)\boldsymbol{d} + \boldsymbol{M}^T(\boldsymbol{\theta}_n)(\widehat{\boldsymbol{\theta}}_\beta - \boldsymbol{\theta}_n) + o\left(\left\|\widehat{\boldsymbol{\theta}}_\beta - \boldsymbol{\theta}_n\right\|\right) + o\left(\|\boldsymbol{\theta}_n - \boldsymbol{\theta}_0\|\right).$$

As

$$\sqrt{K}(\widehat{\boldsymbol{\theta}}_\beta - \boldsymbol{\theta}_n) \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}(\mathbf{0}_{J+1}, \boldsymbol{\Sigma}_\beta(\boldsymbol{\theta}_0))$$

and $\sqrt{K}\left(o\left(\left\|\widehat{\boldsymbol{\theta}}_\beta - \boldsymbol{\theta}_n\right\|\right) + o\left(\|\boldsymbol{\theta}_n - \boldsymbol{\theta}_0\|\right)\right) = o_p(1)$, we have

$$\sqrt{n}\boldsymbol{m}(\widehat{\boldsymbol{\theta}}_\beta) \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}(\boldsymbol{M}^T(\boldsymbol{\theta}_0)\,\boldsymbol{d}, \boldsymbol{M}^T(\boldsymbol{\theta}_0)\boldsymbol{\Sigma}_\beta(\boldsymbol{\theta}_0)\boldsymbol{M}(\boldsymbol{\theta}_0)).$$

We can observe from the relationship $\boldsymbol{d}^* = \boldsymbol{M}(\boldsymbol{\theta}_0)^T\boldsymbol{d}$, if $\boldsymbol{m}(\boldsymbol{\theta}_n) = n^{-1/2}\boldsymbol{d}^*$, that

$$\sqrt{n}\boldsymbol{m}(\widehat{\boldsymbol{\theta}}_\beta) \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}(\boldsymbol{d}^*, \boldsymbol{M}^T(\boldsymbol{\theta}_0)\boldsymbol{\Sigma}_\beta(\boldsymbol{\theta}_0)\boldsymbol{M}(\boldsymbol{\theta}_0)).$$

So, the quadratic form here is

$$W_K\left(\widehat{\boldsymbol{\theta}}_\beta\right) = \boldsymbol{Z}^T\boldsymbol{Z}$$

with

$$\boldsymbol{Z} = \sqrt{n}\boldsymbol{m}(\widehat{\boldsymbol{\theta}}_\beta)\left(\boldsymbol{M}^T(\boldsymbol{\theta}_0)\boldsymbol{\Sigma}_\beta(\boldsymbol{\theta}_0)\boldsymbol{M}(\boldsymbol{\theta}_0)\right)^{-1/2}$$

and

$$\boldsymbol{Z} \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}\left(\left(\boldsymbol{M}^T(\boldsymbol{\theta}_0)\boldsymbol{\Sigma}_\beta(\boldsymbol{\theta}_0)\boldsymbol{M}(\boldsymbol{\theta}_0)\right)^{-1/2}\boldsymbol{M}(\boldsymbol{\theta}_0)^T\boldsymbol{d}, \boldsymbol{I}_r\right),$$

where $\boldsymbol{I}_r$ is the identity matrix or order r. Hence, the required the result follows immediately, and the non-centrality parameter is

$$\boldsymbol{d}^T\boldsymbol{M}(\boldsymbol{\theta}_0)\left(\boldsymbol{M}^T(\boldsymbol{\theta}_0)\boldsymbol{\Sigma}_\beta(\boldsymbol{\theta}_0)\boldsymbol{M}(\boldsymbol{\theta}_0)\right)^{-1}\boldsymbol{M}(\boldsymbol{\theta}_0)^T\boldsymbol{d} = \boldsymbol{d}^{*T}\left(\boldsymbol{M}^T(\boldsymbol{\theta}_0)\boldsymbol{\Sigma}_\beta(\boldsymbol{\theta}_0)\boldsymbol{M}(\boldsymbol{\theta}_0)\right)^{-1}\boldsymbol{d}^*.$$

Proof of Theorem 13

The influence function of $W_K(\widehat{\boldsymbol{\theta}}_\beta)$, with respect to the j_0th observation of the i_0th group of observations, is defined as

$$\mathcal{IF}\left(i_0, j_0, x, T_\beta, \boldsymbol{F}_{\underline{K},\boldsymbol{\theta}}\right) = \left.\frac{\partial W_\beta\left(\boldsymbol{F}_{\underline{K},\boldsymbol{\theta},x,\varepsilon}\right)}{\partial \varepsilon}\right|_{\varepsilon=0^+}$$

where

$$\begin{aligned}
&\frac{\partial W_K\left(\boldsymbol{F}_{\underline{K},\boldsymbol{\theta},x,\varepsilon}\right)}{\partial \varepsilon}\\
&\quad= 2\boldsymbol{m}\left(\boldsymbol{U}_\beta\left(\boldsymbol{F}_{\underline{K},\boldsymbol{\theta},x,\varepsilon}\right)\right)\left(\boldsymbol{M}^T\left(\boldsymbol{U}_\beta\left(\boldsymbol{F}_{\underline{K},\boldsymbol{\theta},x,\varepsilon}\right)\right)\boldsymbol{\Sigma}\left(\boldsymbol{U}_\beta\left(\boldsymbol{F}_{\underline{K},\boldsymbol{\theta},x,\varepsilon}\right)\right)\boldsymbol{M}\left(\boldsymbol{U}_\beta\left(\boldsymbol{F}_{\underline{K},\boldsymbol{\theta},x,\varepsilon}\right)\right)\right)^{-1}\\
&\qquad\times \boldsymbol{M}^T\left(\boldsymbol{U}_\beta\left(\boldsymbol{F}_{\underline{K},\boldsymbol{\theta},x,\varepsilon}\right)\right)\frac{\partial \boldsymbol{U}_\beta\left(\boldsymbol{F}_{\underline{K},\boldsymbol{\theta},x,\varepsilon}\right)}{\partial \varepsilon}\\
&\qquad+ \boldsymbol{m}^T\left(\boldsymbol{U}_\beta\left(\boldsymbol{F}_{\underline{K},\boldsymbol{\theta},x,\varepsilon}\right)\right)\frac{\partial}{\partial \varepsilon}\left(\boldsymbol{M}^T\left(\boldsymbol{U}_\beta\left(\boldsymbol{F}_{\underline{K},\boldsymbol{\theta},x,\varepsilon}\right)\right)\boldsymbol{\Sigma}\left(\boldsymbol{U}_\beta\left(\boldsymbol{F}_{\underline{K},\boldsymbol{\theta},x,\varepsilon}\right)\right)\boldsymbol{M}\left(\boldsymbol{U}_\beta\left(\boldsymbol{F}_{\underline{K},\boldsymbol{\theta},x,\varepsilon}\right)\right)\right)^{-1}\\
&\qquad\times \boldsymbol{m}\left(\boldsymbol{U}_\beta\left(\boldsymbol{F}_{\underline{K},\boldsymbol{\theta},x,\varepsilon}\right)\right)
\end{aligned}$$

and

$$\begin{aligned}
&\left.\frac{\partial W_K\left(\boldsymbol{F}_{\underline{K},\boldsymbol{\theta},x,\varepsilon}\right)}{\partial \varepsilon}\right|_{\varepsilon=0^+}\\
&\quad= 2\boldsymbol{m}\left(\boldsymbol{U}_\beta\left(\boldsymbol{F}_{\underline{K},\boldsymbol{\theta},x}\right)\right)\left(\boldsymbol{M}^T\left(\boldsymbol{U}_\beta\left(\boldsymbol{F}_{\underline{K},\boldsymbol{\theta},x}\right)\right)\boldsymbol{\Sigma}\left(\boldsymbol{U}_\beta\left(\boldsymbol{F}_{\underline{K},\boldsymbol{\theta},x}\right)\right)\boldsymbol{M}\left(\boldsymbol{U}_\beta\left(\boldsymbol{F}_{\underline{K},\boldsymbol{\theta},x}\right)\right)\right)^{-1}\\
&\qquad\times \boldsymbol{M}^T\left(\boldsymbol{U}_\beta\left(\boldsymbol{F}_{\underline{K},\boldsymbol{\theta},x,\varepsilon}\right)\right)\mathcal{IF}\left(\boldsymbol{F}_{\underline{K},\boldsymbol{\theta},x,\varepsilon}\right)\\
&\qquad+ \boldsymbol{m}^T\left(\boldsymbol{U}_\beta\left(\boldsymbol{F}_{\underline{K},\boldsymbol{\theta},x}\right)\right)\left.\frac{\partial}{\partial \varepsilon}\left(\boldsymbol{M}^T\left(\boldsymbol{U}_\beta\left(\boldsymbol{F}_{\underline{K},\boldsymbol{\theta},x,\varepsilon}\right)\right)\boldsymbol{\Sigma}\left(\boldsymbol{U}_\beta\left(\boldsymbol{F}_{\underline{K},\boldsymbol{\theta},x}\right)\right)\boldsymbol{M}\left(\boldsymbol{U}_\beta\left(\boldsymbol{F}_{\underline{K},\boldsymbol{\theta},x}\right)\right)\right)^{-1}\right|_{\varepsilon=0^+}\\
&\qquad\times \boldsymbol{m}\left(\boldsymbol{U}_\beta\left(\boldsymbol{F}_{\underline{K},\boldsymbol{\theta},x}\right)\right)
\end{aligned}$$

For $\boldsymbol{\theta} = \boldsymbol{\theta}_0, \boldsymbol{U}_\beta\left(\boldsymbol{F}_{\underline{K},\boldsymbol{\theta}_0,x}\right) = \boldsymbol{\theta}_0$ and $\boldsymbol{m}^T(\boldsymbol{\theta}_0) = 0$. Therefore,

$$\mathcal{IF}\left(i_0, j_0, x, T_\beta, \boldsymbol{F}_{\underline{K},\boldsymbol{\theta}_0}\right) = 0,$$

in a similar way, we have

$$\mathcal{IF}\left(i_0, j_0, \underline{x}, T_\beta, \boldsymbol{F}_{\underline{K},\boldsymbol{\theta}_0}\right) = 0$$

In order to get the second-order influence function, with respect to the jth observation of the ith group of observation, we must get

$$\mathcal{IF}_2\left(i_0, j_0, x, T_\beta, \boldsymbol{F}_{\underline{K},\boldsymbol{\theta}_0}\right) = \frac{\partial^2}{\partial \varepsilon} W_K(\boldsymbol{F}_{\underline{K}}, \boldsymbol{\theta}_0, x, \varepsilon)\Big|_{\varepsilon=0^+}$$

We can express

$$\begin{aligned}\frac{\partial^2 W_K(\boldsymbol{F}_{\underline{K}}, \boldsymbol{\theta}, x, \varepsilon)}{\partial \varepsilon} &= \boldsymbol{l}(K, \boldsymbol{\theta}, x, \varepsilon) + 2\frac{\partial \boldsymbol{U}_\beta(\boldsymbol{F}_{\underline{K}}, \boldsymbol{\theta}, x, \varepsilon)}{\partial \varepsilon}\left(\boldsymbol{U}_\beta\left(\boldsymbol{F}_{\underline{K},\boldsymbol{\theta},x,\varepsilon}\right)\right) \\ &\times \left(\boldsymbol{M}^T\left(\boldsymbol{U}_\beta\left(\boldsymbol{F}_{\underline{K},\boldsymbol{\theta},x,\varepsilon}\right)\right)\boldsymbol{\Sigma}\left(\boldsymbol{U}_\beta\left(\boldsymbol{F}_{\underline{K},\boldsymbol{\theta},x,\varepsilon}\right)\right)\right. \\ &\quad \left.\boldsymbol{M}\left(\boldsymbol{U}_\beta\left(\boldsymbol{F}_{\underline{K},\boldsymbol{\theta},x,\varepsilon}\right)\right)\right)^{-1}\boldsymbol{M}^T\left(\boldsymbol{U}_\beta\left(\boldsymbol{F}_{\underline{K},\boldsymbol{\theta},x,\varepsilon}\right)\right) \\ &\times \frac{\partial \boldsymbol{U}_\beta\left(\boldsymbol{F}_{\underline{K},\boldsymbol{\theta},x,\varepsilon}\right)}{\partial \varepsilon}\end{aligned}$$

With $\boldsymbol{l}(K, \boldsymbol{\theta}, x, \varepsilon)$, we denote all the terms which contain the expression $\boldsymbol{m}\left(\boldsymbol{U}_\beta\left(\boldsymbol{F}_{\underline{K},\boldsymbol{\theta},x,\varepsilon}\right)\right)$, because for $\boldsymbol{\theta} = \boldsymbol{\theta}_0$ and $\varepsilon = 0$, we have $\boldsymbol{m}\left(\boldsymbol{U}_\beta\left(\boldsymbol{F}_{\underline{K},\boldsymbol{\theta}_0,x}\right)\right) = 0$. Therefore, we have

$$\begin{aligned}\mathcal{IF}_2\left(i_0, j_0, x, T_\beta, \boldsymbol{F}_{\underline{K},\boldsymbol{\theta}_0}\right) &= \frac{\partial^2}{\partial \varepsilon} W_K(\boldsymbol{F}_{\underline{K}}, \boldsymbol{\theta}_0, x, \varepsilon)\Big|_{\varepsilon=0^+} \\ &= 2\mathcal{IF}^T\left(i_0, j_0, x, T_\beta, \boldsymbol{F}_{\underline{K},\boldsymbol{\theta}_0}\right)\boldsymbol{M}(\boldsymbol{\theta}_0)\left(\boldsymbol{M}^T(\boldsymbol{\theta}_0)\boldsymbol{\Sigma}(\boldsymbol{\theta}_0)\boldsymbol{M}(\boldsymbol{\theta}_0)\right)^{-1} \\ &\quad \times \boldsymbol{M}^T(\boldsymbol{\theta}_0)\mathcal{IF}\left(i_0, j_0, x, T_\beta, \boldsymbol{F}_{\underline{K},\boldsymbol{\theta}_0}\right).\end{aligned}$$

In a similar way, we obtain the expression of $\mathcal{IF}_2\left(\underline{x}, T_\beta, \boldsymbol{F}_{\underline{K},\boldsymbol{\theta}_0}\right)$.

References

Balakrishnan N, Ling MH (2012) Multiple-stress model for one-shot device testing data under exponential distribution. IEEE Trans Reliab 61(3):809–821

Balakrishnan N, Ling MH (2013) Expectation maximization algorithm for one shot device accelerated life testing with Weibull lifetimes, and variable parameters over stress. IEEE Trans Reliab 62(2):537–551

Balakrishnan N, Ling MH (2014) Gamma lifetimes and one-shot device testing analysis. Reliab Eng Syst Saf 126:54–64

Balakrishnan N, Castilla E, Martín N, Pardo L (2019) Robust estimators and test-statistics for one-shot device testing under the exponential distribution. IEEE Trans Inf Theory 65(5):3080–3096. https://doi.org/10.1109/tit.2019.2903244

Basu A, Ghosh A, Mandal A, Martin N, Pardo L (2017) A Wald-type test statistic for testing linear hypothesis in logistic regression models based on minimum density power divergence estimator. Electonic J Stat 11:2741–2772

Fan TH, Balakrishnan N, Chang CC (2009) The Bayesian approach for highly reliable electro-explosive devices using one-shot device testing. J Stat Comput Simul 79(9):1143–1154

Ghosh A, Basu A (2013) Robust estimation for independent non-homogeneous observations using density power divergence with applications to linear regression. Electron J Stat 7:2420–2456

Heilbron DC (1994) Zero-altered and other regression models for count data with added zeros. Biom J 36:531–547

Husak GJ, Michaelsen J, Funk C (2007) Use of the gamma distribution to represent monthly rainfall in Africa for drought monitoring applications. Int J Climatol 27:935–944

Kodell RL, Nelson CJ (1980) An illness-death model for the study of the carcinogenic process using survival/sacrifice data. Biometrics 36:267–277

Kullback S (1959) Statistics and information theory. Wiley, New York

Kullback S, Leibler RA (1951) On information and sufficiency. Ann Math Stat 22(1):79–86

Kwon K, Frangopol DM (2010) Bridge fatigue reliability assessment using probability density functions of equivalent stress range based on field monitoring data. Int J Fatigue 32:1221–1223

Lambert D (1992) Zero-inflated Poisson regression, with an application to defects in manufacturing. Technometrics 34(1):1–14

Meeter CA, Meeker WQ (1994) Optimum accelerated life tests with a nonconstant scale parameter. Technometrics 36:71–83

Meeker WQ, Escobar LA, Lu CJ (1998) Accelerated degradation tests: modeling and analysis. Technometrics 40:89–99

Pardo L (2006) Statistical inference based on divergence measures. Chapman & Hall/ CRC Press, Boca Raton

Seabom JB (1991) Hypergeometric functions and their applications. Springer, New York

Tseng ST, Balakrishnan N, Tsai CC (2009) Optimal step-stress accelerated degradation test plan for gamma degradation processes. IEEE Trans Reliab 58:611–618

Warwick J, Jones MC (2005) Choosing a robustness tuning parameter. J Stat Comput Simul 75:581–588

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.