Robust estimation based on one-shot device test data under log-normal lifetimes

N. Balakrishnan and E. Castilla

aDepartment of Mathematics and Statistics, McMaster University, Hamilton, Ontario, Canada; bDepartamento de Matematica Aplicada, Rey Juan Carlos University, Madrid, Spain

ABSTRACT
In this paper, we present robust estimators for one-shot device test data under log-normal lifetimes. Based on these estimators, confidence intervals and Wald-type tests are also developed. Their robustness feature is illustrated through a simulation study as well as two numerical examples.

ARTICLE HISTORY
Received 18 July 2022
Accepted 2 July 2023

KEYWORDS
Divergences; reliability; robustness; one-shot devices

1. Introduction
A destructive one-shot device is a product, system or an unit that can be used only once. After use, it gets destroyed or must be rebuilt. Some examples are munitions, automobile air bags, heat detectors and antigen tests. The study of one-shot device data has been developed considerably recently, mainly motivated by the works of Ling [1] and So [2]. For a complete review of all recent developments concerning the analysis of one-shot device data, one may refer to the book by Balakrishnan et al. [3]. Note that these data can also be analyzed in survival analysis in the context of ‘current status data’. Some examples of current status data include animal carcinogenicity experiments [4] or HIV transmission [5].

Due to significant developments in manufacturing technology, such one-shot devices usually possess long lifetimes, with a low failure rate. That makes the initial data collection and processing stages more complex. To save time and cost, accelerated life tests (ALTs) are commonly employed as they induce early failures by testing items under high stresses, such as temperature, air-pressure or voltage. A very common type of ALT is the constant-stress ALT (CSALT), which assumes that each device is subject to only pre-specified stress levels; see, for example, Balakrishnan and Ling [6] and Balakrishnan et al. [7]. Other forms of ALTs for one-shot devices have also been discussed in the literature; see Ling [8], Ling and Hu [9] and Zhu et al. [10], among others.

From here on, we assume CSALTs are performed on $I$ groups of one-shot devices, each of which is subject to $J$ types of stress levels, and that in the $i$th testing condition $K_i$ devices are tested at a pre-specified inspection time $\tau_i$, for $i = 1, \ldots, I$. Then, in the $i$th test group, the number of failures, $n_i$, is collected. The experimental condition and the data so obtained are as summarized in Table 1. Note that one-shot device test data are an extreme case of
Table 1. Data collected on one-shot devices at multiple stress levels and collected at different inspection times.

| Condition | Inspection Time | Devices | Failures | Covariates |
|-----------|-----------------|---------|----------|------------|
|           |                 |         |          | Stress 1   |
| 1         | $\tau_1$        | $K_1$   | $n_1$    | $x_{11}$   |
| 2         | $\tau_2$        | $K_2$   | $n_2$    | $x_{21}$   |
| ...       | ...             | ...     | ...      | ...        |
| I         | $\tau_I$        | $K_I$   | $n_I$    | $x_{I1}$   |

censoring since we do not observe any failure time, only the status of the device (failure or not) gets observed at the inspection time.

Most of the existing works on one-shot devices are based on parametric models. Under the parametric model, the lifetimes are described by a statistical distribution which relates the stress levels to the model parameter vector $\theta \in \Theta$. Inference for exponential, gamma and Weibull distributions has been successfully developed for one-shot device test data in this context; see Balakrishnan and Ling [6,11–13] and Balakrishnan et al. [14,15], among others. However, log-normal lifetime distribution has not been studied in detail in this set-up. While the hazard function (which measures the instantaneous rate of failure) for exponential distribution is always a constant, and that of Weibull and gamma distributions are either increasing or decreasing, the log-normal distribution has increasing-decreasing behaviour of hazard which is encountered often in practice as units usually experience early failure and then stabilize over time in terms of performance. In this respect, Balakrishnan and Castilla [7] developed an EM algorithm for the likelihood estimation of the model parameter vector based on one-shot device test data under log-normal distribution. Upadhyay and Sharma [16] developed Bayesian techniques in the context of current status data, also based on the likelihood function and assuming log-normal lifetimes, and applied them to a real dataset obtained from a survival experiment conducted at the National Center for Toxicological Research (NCTR). However, the maximum likelihood estimator (MLE) is known for its lack of robustness. For this reason, we develop here robust divergence-based inference for one-shot device test data under log-normal distribution, following an approach similar to that of Balakrishnan et al. [17].

The rest of this paper is organized as follows. Section 2 introduces the likelihood approach for one-shot device testing under the log-normal distribution. Section 3 presents the weighted minimum DPD estimators as an alternative to the MLEs. Their asymptotic distribution is presented and a family of Wald-type test statistics is then developed. The robustness features of the new families of estimators and associated tests are evaluated by means of the boundedness of their Influence Function (IF) in Section 4. Section 5 presents the results of an elaborate simulation study. In Section 6, two data sets are analyzed to illustrate all the inferential results developed here. Finally, some concluding remarks are made in Section 7.

2. Model description and likelihood inference

We shall assume that the lifetimes of the units, under the $i$th testing condition, follow log-normal distribution with probability density function and cumulative distribution
function as

\[ f(t; x_i, \theta) = \frac{1}{\sqrt{2\pi}\sigma_i} \exp \left\{ -\frac{(\log(t) - \mu_i)^2}{2\sigma_i^2} \right\}, \quad t > 0, \tag{1} \]

and

\[ F(t; x_i, \theta) = \Phi \left( \frac{\log(t) - \mu_i}{\sigma_i} \right), \quad t > 0, \tag{2} \]

respectively. Here, \( \Phi(\cdot) \) denotes the cumulative distribution function of the standard normal distribution, and \( \mu_i \in \mathbb{R} \) and \( \sigma_i \in \mathbb{R}^+ \) are, respectively, the scale and shape parameters, which we assume are related to the stress levels in the form

\[ \mu_i = \sum_{j=0}^{J} a_j x_{ij}, \quad \sigma_i = \exp \left\{ \sum_{j=0}^{J} b_j x_{ij} \right\}, \]

with \( x_{i0} = 1 \), for all \( i \). Here, \( \theta = (a_0, \ldots, a_J, b_0, \ldots, b_J)^T \in \Theta = \mathbb{R}^{(2J+1)} \) is the model parameter vector. In this case, the mean lifetime and the hazard function are known to be [18]

\[ E(x_i, \theta) = \exp \left\{ \mu_i + \sigma_i^2/2 \right\}, \]

\[ h(t; x_i, \theta) = \frac{1}{\sqrt{2\pi}\sigma_i} \exp \left\{ -\frac{(\log(t) - \mu_i)^2}{2\sigma_i^2} \right\} \frac{1}{1 - \Phi \left( \frac{\log(t) - \mu_i}{\sigma_i} \right)}, \tag{3} \]

respectively.

Instead of working with log-normal lifetimes, it is more convenient to work with log-lifetimes, \( \omega_{ik} = \log(t_{ik}) \), since this belongs to a location-scale family of distributions; see Meeker [19] and Yuan et al. [20]. In fact, the log-lifetimes follow a normal distribution with density and cumulative distribution functions as

\[ f_{\omega}(\omega; x_i, \theta) = \phi \left( \frac{\omega - \mu_i}{\sigma_i} \right) = \frac{1}{\sqrt{2\pi}\sigma_i} \exp \left\{ -\frac{(\omega - \mu_i)^2}{2\sigma_i^2} \right\}, \quad -\infty < \omega < \infty, \tag{4} \]

\[ F_{\omega}(\omega; x_i, \theta) = \Phi \left( \frac{\omega - \mu_i}{\sigma_i} \right), \quad -\infty < \omega < \infty, \]

respectively.

The likelihood function of the complete data is given by

\[ L_c(\theta) = \prod_{i=1}^{I} \prod_{k=1}^{K_i} f_{\omega}(\omega_{ik}; \theta), \tag{5} \]

where \( \omega_{ik} \) is the true log-lifetime of the \( k \)th device in the \( i \)th condition. The corresponding MLE of \( \theta \), \( \hat{\theta}_{\text{MLE}} \), will be obtained by maximization of (5) or, equivalently, its logarithm:

\[ \ell_c(\theta) = \sum_{i=1}^{I} \sum_{k=1}^{K_i} \log(f_{\omega}(\omega_{ik}; \theta)). \tag{6} \]

However, one-shot device testing is an extreme case of interval censoring, in which the true lifetimes are unknown. Therefore, we cannot compute the exact value of (6). Balakrishnan
and Castilla [7], developed an EM algorithm for the estimation of the parameter vector. In the first step of this iterative algorithm, the true lifetime of each device is imputed with its expected value. In the second step, the expected value of the complete log-likelihood is maximized. In this paper, we will alternatively use the likelihood based on the observed data. For convenience, let us now denote $z = \{W_i, K_i, n_i, i = 1, \ldots, I\}$ with $W_i = \log(\tau_i)$ for the observed data. Assuming independent observations, the likelihood function based on the observed data is given by

$$L(\theta; z) \propto \prod_{i=1}^{I} F_{\omega}^{n_i}(W_i; x_i, \theta) R_{\omega}^{K_i-n_i}(W_i; x_i, \theta),$$  

where $R_{\omega}(\omega; x_i, \theta) = 1 - F_{\omega}(\omega; x_i, \theta)$ is the reliability function at test condition $i$. The corresponding MLE of $\theta$, $\hat{\theta}_{MLE}$, will be obtained by maximization of (8) or, equivalently, its logarithm, as

$$\hat{\theta} = \arg \max_{\theta \in \Theta} \log L(\theta; z).$$  

Differentiating (7) with respect to $\theta$ and equating to zero, we obtain the estimating equations for the MLE as

$$\sum_{i=1}^{I} \delta_i (K_i F_{\omega}(W_i; x_i, \theta) - n_i) \left( F_{\omega}^{-1}(W_i; x_i, \theta) + R_{\omega}^{-1}(W_i; x_i, \theta) \right) x_i = 0_{2(J+1)},$$

where $0_{2(J+1)}$ is the null column vector of dimension $2(J+1)$,

$$\delta_i x_i = \frac{\partial F_{\omega}(W_i; x_i, \theta)}{\partial \theta} = \left( \frac{1}{\sigma_i} \phi \left( \frac{W_i - \mu_i}{\sigma_i} \right) \frac{(W_i - \mu_i)}{\sigma_i} \phi \left( \frac{W_i - \mu_i}{\sigma_i} \right) \right)^T x_i,$$

and $\phi(\cdot)$ is the density function of the standard normal distribution.

### 2.1. Asymptotic distribution of the MLE and confidence intervals

The observed Fisher Information matrix for model parameters, $\mathcal{I}_{obs}$, is given by the second-order derivatives of the observed log-likelihood function in (8), given by

$$\mathcal{I}_{obs}(\theta) = \sum_{i=1}^{I} K_i \left( F^{-1}(IT_i; x_i, \theta) + R^{-1}(IT_i; x_i, \theta) \right) \left( \frac{\partial F(IT_i; x_i, \theta)}{\partial \theta} \right) \left( \frac{\partial F(IT_i; x_i, \theta)}{\partial \theta} \right)^T.$$  

The asymptotic properties of MLE for regular distributions were stated by Chimitova and Balakrishnan [21]. The asymptotic variance-covariance matrix of the MLEs of the model parameters can then be obtained by inverting the above observed Fisher information matrix as

$$V(\theta) = \mathcal{I}^{-1}_{obs}(\theta).$$

In particular, the variance-covariance matrix is the Rao-Cramer lower bound, implying the efficiency of the MLEs if the model is correctly specified. A detailed explanation of this result can be found in Balakrishnan and Ling [13], or in Appendix B of the book by Balakrishnan et al. [3].
The asymptotic variance of the MLE of lifetime characteristics, such as reliability and mean lifetime of units at normal operating conditions \((\omega_0, x_0)\), can be computed readily by employing delta method, as

\[
V_R(\theta) = P_R^T(\hat{\theta})V(\theta)P_R(\hat{\theta}),
\]

\[
V_E(\theta) = P_E^T(\hat{\theta})V(\theta)P_E(\hat{\theta}),
\]

where

\[
P_R(\hat{\theta}) = \frac{\partial R(\omega_0; x_0, \theta)}{\partial \theta} \bigg|_{\theta = \hat{\theta}} \quad \text{and} \quad P_E(\hat{\theta}) = \frac{\partial E(x_0, \theta)}{\partial \theta} \bigg|_{\theta = \hat{\theta}}.
\]

In practice, as we do not know the real value of \(\theta\), and so we will replace it by the estimated \(\hat{\theta}\).

The 100\((1 - \alpha)\)% asymptotic confidence interval for any parameter of interest \(\upsilon\) is given by

\[
(\hat{\upsilon} - z_{\alpha/2}se(\hat{\upsilon}), \hat{\upsilon} + z_{\alpha/2}se(\hat{\upsilon}))
\]

where \(z_{\alpha/2}\) is the upper \(\alpha/2\) standard normal quantile and \(se(\hat{\upsilon})\) is the standard error of \(\hat{\upsilon}\) obtained from the asymptotic variance-covariance matrix. However, the asymptotic confidence interval is based on the asymptotic properties of the MLE, and so it may be satisfactory only for large sample sizes. Further, we would need to truncate the bounds of the confidence intervals for the mean lifetime and reliability of the devices, since the mean lifetime has to be positive and the reliability lies between 0 and 1. For this reason, Balakrishnan and Castilla [7] studied the behaviour of the hyperbolic arcsecant (arsech) and logit transformations of confidence intervals of reliability under the log-normal distribution.

In the arsech-approach, the 100\((1 - \alpha)\)% asymptotic confidence interval for reliability takes on the form

\[
\left( \exp\left(\frac{2}{\exp(-U_f) + \exp(U_f)}\right), \exp\left(\frac{2}{\exp(-L_f) + \exp(L_f)}\right) \right),
\]

with

\[
U_f = \hat{f} + z_{\alpha/2}se(\hat{f}), \quad L_f = \hat{f} - z_{\alpha/2}se(\hat{f}),
\]

and

\[
\hat{f} = \log\left(\frac{1 + \sqrt{1 - \hat{R}^2}}{\hat{R}}\right), \quad se(\hat{f}) = \frac{se(\hat{R})}{\hat{R}\sqrt{1 - \hat{R}^2}},
\]

where \(\hat{R}\) denotes, for simplicity, the estimated reliability. In the logit-approach, the 100\((1 - \alpha)\)% asymptotic confidence interval for reliability is of the form

\[
\left( \frac{\hat{R}}{\hat{R} + (1 - \hat{R})S}, \frac{\hat{R}}{\hat{R} + (1 - \hat{R})/S} \right),
\]

where \(S = \exp(z_{\alpha/2}se(\hat{R})/\hat{R}(1 - \hat{R}))\). For more details concerning the confidence intervals in (10) and (11), see Balakrishnan and Ling [11]. These authors also studied the performance of the
log-approach for the construction of a 100(1 − α)% confidence interval for the expected lifetime, \( \hat{T} \), of the form

\[
\left( \hat{T} \exp \left\{ -\frac{z_{\alpha/2} \text{se}(\hat{T})}{\hat{T}} \right\}, \hat{T} \exp \left\{ \frac{z_{\alpha/2} \text{se}(\hat{T})}{\hat{T}} \right\} \right).
\]  

(12)

2.2. Wald test

The Wald test is a well-known multivariate test of hypothesis that allows testing a set of parameters of the considered model simultaneously.

Let us consider the function \( m: \mathbb{R}^{2(J+1)} \rightarrow \mathbb{R}^r \), where \( r \leq 2(J + 1) \). Then, \( m(\theta) = \mathbf{0}_r \) represents a composite null hypothesis. We assume that the \( 2(J + 1) \times r \) matrix

\[
M(\theta) = \frac{\partial m^T(\theta)}{\partial \theta}
\]

exists and is continuous in \( \theta \), with rank \( M(\theta) = r \). Then, for testing

\[
H_0 : \theta \in \Theta_0 \quad \text{against} \quad H_1 : \theta \notin \Theta_0,
\]

(13)

with \( \Theta_0 = \{ \theta \in \Theta : m(\theta) = \mathbf{0}_r \} \), we can consider the following Wald test statistic:

\[
W_K(\hat{\theta}) = K m^T(\hat{\theta}) \left( M^T(\hat{\theta}) \Sigma(\hat{\theta}) M(\hat{\theta}) \right)^{-1} m(\hat{\theta}),
\]

(14)

where, for further convenience, \( \Sigma(\hat{\theta}) = V(\hat{\theta}) / K \). Here, \( K = \sum_{i=1}^J K_i \) is the total number of devices.

The asymptotic null distribution of the Wald test statistic, given in (14), is a chi-squared (\( \chi^2 \)) distribution with \( r \) degrees of freedom, i.e.

\[
W_K(\hat{\theta}) \xrightarrow{K \to \infty} \chi^2_r.
\]

Based on this, we will reject the null hypothesis in (13) if

\[
W_K(\hat{\theta}) > \chi^2_{r,\alpha},
\]

where \( \chi^2_{r,\alpha} \) is the upper \( \alpha \) percentage point of \( \chi^2_r \) distribution.

Example 2.1: In the context of one-shot device testing under multiple stress levels, we may be interested in checking whether there is a significant relationship between the \( j \)th stress factor and the lifetime of the device. This can be tested through Wald test in (14), where \( m(\theta) = (\theta_j, \theta_{(J+1)+j}) \), and

\[
M^T(\theta) = \begin{pmatrix}
0, & \ldots, & (J+1) & \ldots, & (J+1) & \ldots, & (J+1+j) & \ldots, & (2J+2)
\end{pmatrix}
\]

(0, \ldots, 0, \ldots, 1, \ldots, 0, \ldots, 0).
2.3. Model misspecification

The main problem when considering a parametric approach is that we are assuming the distribution associated to our data beforehand. In reliability studies, practitioners usually rely on underlying distribution to develop the corresponding inference. However, model misspecification can be a serious issue, causing significant errors in the estimation. Many authors have dealt with the problem of model misspecification in the presence of censored data. For example, Dey and Kundu [22] investigated the effect of model misspecification in the case of Type-II censoring. In the context of one-shot device test data, Chimitova and Balakrishnan [21] proposed four different statistics for the purpose of testing the goodness-of-fit. On the other hand, the Akaike Information Criterion (AIC) was used by Ling and Balakrishnan [23] to analyse the model misspecification of Weibull and gamma models, and in Balakrishnan et al. [24] to evaluate the effect of model misspecification in the optimal design of experiments. If the model depends on a parameter \( \theta \) of dimension \( R \) and has log-likelihood \( \ell(\theta) \), the AIC associated to the model is given by

\[
AIC = 2R - 2\ell(\theta).
\]

Given a set of competing models, the preferred model is the one with smallest value of AIC.

3. Divergence-based inference

We first introduce some notation in order to define the MLE on the basis of Kullback-Leibler divergence. For each testing condition \( i, i = 1, \ldots, I \), the empirical and theoretical probability vectors are given, respectively, by

\[
\hat{p}_i = (\hat{p}_{i1}, \hat{p}_{i2})^T = \left( \frac{n_i}{K_i}, 1 - \frac{n_i}{K_i} \right)^T,
\]

\[
\pi_i(\theta) = (\pi_{i1}(\theta), \pi_{i2}(\theta))^T = (F_{\omega}(W_i; x_i, \theta), R_{\omega}(W_i; x_i, \theta))^T.
\]

The empirical and theoretical vectors provide, for each testing condition, the probability of failure and success based on the data collected and on the theoretical parametric model considered, respectively.

Let us consider the weighted Kullback-Leibler divergence measure of all the units given by

\[
\frac{I}{K} \sum_{i=1}^{I} K_i d_{KL}(\hat{p}_i, \pi_i(\theta)) = \frac{1}{K} \sum_{i=1}^{I} K_i \left[ \hat{p}_{i1} \log \left( \frac{\hat{p}_{i1}}{\pi_{i1}(\theta)} \right) + \hat{p}_{i2} \log \left( \frac{\hat{p}_{i2}}{\pi_{i2}(\theta)} \right) \right],
\]

where \( K = \sum_{i=1}^{I} K_i \) is the total number of devices.

The following result establishes the relationship between the weighted Kullback-Leibler divergence in (18) and the MLE.

**Result 3.1:** The MLE can be obtained as the minimization of the weighted Kullback-Leibler divergence measure between probability vectors in (16) and (17); that is,

\[
\hat{\theta} = \arg \min_{\theta \in \Theta} \sum_{i=1}^{I} \frac{K_i}{K} d_{KL}(\hat{p}_i, \pi_i(\theta)).
\]
Remark 3.2: Result 3.1 provides us the basis for the development of alternative estimators for one-shot device testing. The idea is to choose a different divergence measure (between empirical and theoretical probability vectors) to be minimized. In particular, the DPD is well-known for its robustness properties.

Given the probability vectors \( \hat{p}_i \) and \( \pi_i(\theta) \) defined in (16) and (17), respectively, the density power divergence (DPD) between the two probability vectors (see Basu et al. [25]), with tuning parameter \( \beta \geq 0 \), is given by

\[
d_{\beta}(\hat{p}_i, \pi_i(\theta)) = \left( \pi_{i1}^{\beta+1}(\theta) + \pi_{i2}^{\beta+1}(\theta) \right) - \frac{\beta + 1}{\beta} \left( \hat{p}_{i1}\pi_{i1}^\beta(\theta) + \hat{p}_{i2}\pi_{i2}^\beta(\theta) \right) + \frac{1}{\beta} \left( \hat{p}_{i1}^{\beta+1} + \hat{p}_{i2}^{\beta+1} \right), \quad \text{if } \beta > 0,
\]

(20)

As the term \( \frac{1}{\beta} (\hat{p}_{i1}^{\beta+1} + \hat{p}_{i2}^{\beta+1}) \) does not have any role in the minimization with respect to \( \theta \) in (20), we can consider the equivalent measure for \( \beta > 0 \) as

\[
d^*_\beta(\hat{p}_i, \pi_i(\theta)) = \lim_{\beta \to 0^+} d_{\beta}(\hat{p}_i, \pi_i(\theta)) = d_{KL}(\hat{p}_i, \pi_i(\theta)).
\]

(21)

Based on Result 3.1, we can now define the weighted minimum DPD estimators for the one-shot device model with multiple stress levels.

Definition 3.3: Let us consider the framework in Table 1, and define the weighted minimum DPD estimator for \( \theta \) as

\[
\hat{\theta}_\beta = \arg \min_{\theta \in \Theta} \sum_{i=1}^{I} \frac{K_i}{K} d^*_\beta(\hat{p}_i, \pi_i(\theta)), \quad \text{for } \beta > 0,
\]

where \( d^*_\beta(\hat{p}_i, \pi_i(\theta)) \) is as given in (22), and \( \hat{p}_i \) and \( \pi_i(\theta) \) are as given in (16) and (17), respectively. For \( \beta = 0 \), we get the MLE, \( \hat{\theta} \).

Differentiating \( d^*_\beta(\hat{p}_i, \pi_i(\theta)) \) with respect to \( \theta \) and equating to zero, we have for \( \beta \geq 0 \), the estimating system of equations to be

\[
\sum_{i=1}^{I} \delta_i \left( K_iF_{\omega}(W_i; \theta) - n_i \right) \left[ F_{\omega}^{\beta-1}(W_i; \theta) + R_{\omega}^{\beta-1}(W_i; \theta) \right] x_i = 0_{2(I+1)},
\]

with \( \delta_i \) as in (9).

A detailed review of divergence-based robust inferential methods for one-shot device testing under different lifetime distributions can be found in Balakrishnan et al. [17]. We now derive the asymptotic distribution of the proposed weighted minimum DPD estimators, which will facilitate the development of a new family of Wald-type tests for hypothesis testing.
3.1. Asymptotic distribution and confidence intervals

In the following result, we present the asymptotic distribution of the proposed weighted minimum DPD estimators.

**Result 3.4:** Let \( \theta^0 \) be the true value of the parameter \( \theta \). Then, the asymptotic distribution of the weighted minimum DPD estimator \( \hat{\theta}_\beta \) is given by

\[
\sqrt{K} \left( \hat{\theta}_\beta - \theta^0 \right) \xrightarrow{K \to \infty} N \left( 0_{2(J+1)}, \Sigma_\beta(\theta^0) \right),
\]

where

\[
\Sigma_\beta(\theta) = KV_\beta(\theta) = J^{-1}(\theta)K_\beta(\theta)J^{-1}(\theta),
\]

(23)

with

\[
J_\beta(\theta) = \sum_i K_i \frac{\Lambda_i}{K} \left( F_{\omega}^{-1}(W_i; x_i, \theta) + R_{\omega}^{-1}(W_i; x_i, \theta) \right) x_i x_i^T,
\]

\[
K_\beta(\theta) = \sum_i K_i \Lambda_i F_{\omega}(W_i; x_i, \theta) R_{\omega}(W_i; x_i, \theta) \left( F_{\omega}^{-1}(W_i; x_i, \theta) + R_{\omega}^{-1}(W_i; x_i, \theta) \right)^2 x_i x_i^T
\]

and \( \Lambda_i = \delta_i \delta_i^T \).

The proof of this result follows from Theorem 3.1 of Ghosh and Basu [26]; see Appendix A for more details. Using the asymptotic distribution of the weighted minimum DPD estimator given in Result 3.4, we can construct confidence intervals similar to those detailed in Section 2.1 for the model parameters.

3.2. Wald-type tests

Let us consider again the problem of testing the null hypothesis in (13). We can consider the following Wald-type test statistic as a generalization of the classical Wald test defined in (14):

\[
W_K(\hat{\theta}_\beta) = Km^T(\hat{\theta}_\beta) \left( M^T(\hat{\theta}_\beta) \Sigma(\hat{\theta}_\beta)M(\hat{\theta}_\beta) \right)^{-1} m(\hat{\theta}_\beta),
\]

(24)

where \( \Sigma(\theta) \) is as given in (23). The classical Wald test given in (14) is then deduced for \( \beta = 0 \).

**Result 3.5:** The asymptotic null distribution of the proposed Wald-type test statistic, given in (24), is a chi-squared \( \chi^2 \) distribution with \( r \) degrees of freedom; that is,

\[
W_K(\hat{\theta}_\beta) \xrightarrow{K \to \infty} \chi^2_r.
\]
Result 3.5 shows that we can reject the null hypothesis in (13) if

\[ W_{K} (\hat{\theta}) > \chi_{r, \alpha}^2, \]  

where \( \chi_{r, \alpha}^2 \) is the upper \( \alpha \) percentage point of \( \chi^2_r \) distribution; see Appendix A for details on the proof of Result 3.5.

4. Study of the influence function

The Influence Function (IF) is an important concept in statistical robustness theory. Introduced by Hampel et al. [27], the (first-order) IF of an estimator, as a function of \( t \), measures the standardized asymptotic bias (in its first-order approximation) caused by the infinitesimal contamination at the point \( t \). The larger the value of IF is, the less robust the estimator is.

Let us denote by \( G_i \) the true distribution function of a Bernoulli random variable with an unknown probability of success, for the \( i \)th group of \( K_i \) observations, having mass function \( g_i \). Similarly, we denote by \( F_{i, \theta} \) the distribution function of a Bernoulli random variable having a probability of success as \( \pi_{i1}(\theta) \), with probability mass function \( f_i(\cdot, \theta) \) \((i = 1, \ldots, I)\), which is related to the considered model. In vector notation, we consider \( G = (G_1 \otimes 1_{K_1}^T, \ldots, G_I \otimes 1_{K_I}^T)^T \) and \( F_{\theta} = (F_{1, \theta} \otimes 1_{K_1}^T, \ldots, F_{I, \theta} \otimes 1_{K_I}^T)^T \).

For any estimator defined in terms of a statistical functional \( U(G) \) in the set-up of data from the true distribution function \( G \), its IF is defined as

\[ IF(t, U, G) = \lim_{\varepsilon \downarrow 0} \frac{U(G_{\varepsilon, t}) - U(G)}{\varepsilon} = \frac{\partial U(G_{\varepsilon, t})}{\partial \varepsilon} \bigg|_{\varepsilon = 0^+}, \]

where \( G_{\varepsilon, t} = (1 - \varepsilon)G + \varepsilon \Delta_t \), with \( \varepsilon \) being the contamination proportion and \( \Delta_t \) being the distribution function of the degenerate random variable at the contamination point

\[ t = (t_{11}, \ldots, t_{1K_1}, \ldots, t_{I1}, \ldots, t_{IK_I})^T \in \mathbb{R}^{IK}. \]

We first need to define the statistical functional \( U_\beta(G) \) corresponding to the weighted minimum DPD estimator as the minimizer of the weighted sum of DPDs between the true and model densities. This is defined as the minimizer of

\[ \sum_{i=1}^{I} \frac{K_i}{K} \left\{ \sum_{y \in \{0, 1\}} \left[ f_i^{\beta + 1}(y, \theta) - \frac{\beta + 1}{\beta} f_i^{\beta}(y, \theta) g_i(y) \right] \right\}, \]  

where \( g_i(y) \) is the probability mass function associated to \( G_i \) and

\[ f_i(y, \theta) = y \pi_{i1}(\theta) + (1 - y) \pi_{i2}(\theta), \quad y \in \{0, 1\}. \]

If we choose \( g_i(y) = f_i(y, \theta) \), the expression in (26) is minimized at \( \theta = \theta^0 \), implying the Fisher consistency of the minimum DPD estimator functional \( U_\beta(G) \) in the considered model.

We can derive the IF of the minimum DPD estimators at \( F_{\theta^0} \) with respect to the \( k \)th observation of the \( i_0 \)th group and with respect to all the observations as in the following results.
Result 4.1: Let us consider the one-shot device testing under the log-normal distribution. Then, the IF with respect to the \(k\)th observation of the \(i_0\)th group is given by

\[
IF(t_{i_0,k}, U_\beta, F_{\theta^0}) = J_\beta^{-1}(\theta^0) \frac{K_{i_0}}{K} \delta_{i_0} x_{i_0} \\
\times \left( F_{\omega}^{\beta-1}(W_{i_0}; \theta^0) + R_{\omega}^{\beta-1}(W_{i_0}; \theta^0) \right) \left( F_{\omega}(W_{i_0}; \theta^0) - \Delta_{(1)_{i_0,k}}(t_{i_0,k}) \right),
\]

where \(\Delta_{(1)_{i_0,k}}(t_{i_0,k})\) is the degenerating function at point \((t_{i_0,k})\).

Result 4.2: Let us consider the one-shot device testing under the log-normal distribution. Then, the IF with respect to all the observations is given by

\[
IF(t, U_\beta, F_{\theta^0}) = J_\beta^{-1}(\theta^0) \sum_{i=1}^{I} \frac{K_i}{K} \delta_i x_i \\
\times \left( F_{\omega}^{\beta-1}(W_i; \theta^0) + R_{\omega}^{\beta-1}(W_i; \theta^0) \right) \left( F_{\omega}(W_i; \theta^0) - \Delta_{(1)_{i_0,k}} \right),
\]

where \(\Delta_{(1)_{i_0,k}} = \sum_{k=1}^{K_i} \Delta_{(1)_{i_0,k}}\).

Proofs of Results 4.1 and 4.2 require some heavy manipulations. We abstain from presenting them here, but one may refer to Balakrishnan et al. [15] for a similar derivation for the case of Weibull distribution.

Let \(\theta^T = (a^T, b^T)\), with \(a = (a_0, \ldots, a_J)^T\) and \(b = (b_0, \ldots, b_J)^T\). Then, the factors of the influence function of \(\theta\) given in (27) and (28), associated with \(a\) and \(b\), respectively, are

\[
h_1(\omega, x, \theta, \beta) = \frac{1}{\sigma} \phi \left( \frac{\omega - \mu}{\sigma} \right) \left[ \Phi \left( \frac{\omega - \mu}{\sigma} \right)^{\beta-1} + \left( 1 - \Phi \left( \frac{\omega - \mu}{\sigma} \right) \right)^{\beta-1} \right] x, \tag{29}
\]

\[
h_2(\omega, x, \theta, \beta) = \frac{\omega - \mu}{\sigma} \phi \left( \frac{\omega - \mu}{\sigma} \right) \left[ \Phi \left( \frac{\omega - \mu}{\sigma} \right)^{\beta-1} + \left( 1 - \Phi \left( \frac{\omega - \mu}{\sigma} \right) \right)^{\beta-1} \right] x. \tag{30}
\]

Based on these, we may comment on conditions for boundedness of the IFs presented here, either with respect to an observation or with respect to all the observations, that they are bounded on \(t_{i_0,k}\) or \(t\); but, if \(\beta = 0\), the norm of the influence functions can be very large on \((x, \omega)\), in comparison with \(\beta > 0\), as can be seen in Figure 1 for the case of only one factor of stress.

Specifically, to see the effect of \(\omega\) in (29) and (30), we fix \(\mu = 1, \sigma = 1\) and \(x = 1\). When \(\omega\) increases, the classical MLE becomes clearly non-robust. This certainly necessitates the use of DPD-based estimators. On the other hand, to see the effect of \(x\) in (29) and (30), we take \(\omega = 1\) with \(\theta^T = (0, -1, 0, 1)\) or \(\theta^T = (0, -1, 0, -1)\). Note that we are assuming \(x\) to be positive, incrementing the stress as \(x \to \infty\). Thus, we set \(a_1 < 0\) to ensure an increment of the stress implies a reduction of the expected lifetime in (3). Weighted minimum DPDs with \(\beta > 0\) do present a more robust behaviour than the classical MLE.
Figure 1. Effect of the change of $\omega$ and $x$ in $h_1(\omega, x, \theta, \beta)$ (left) and $h_2(\omega, x, \theta, \beta)$ (right) for different values of the tuning parameter $\beta$. 
4.1. Influence function of wald-type tests

The functional associated with Wald-type statistics for testing the composite null hypothesis in (13) is, ignoring the multiplier $K$, given by

$$W_K(U_\beta) = (M^T U_\beta - m)^T \left[ M^T V(U_\beta) M \right]^{-1} (M^T U_\beta - m),$$

and the IF with respect to the $k$th observation of the $i_0$th group of observations, is then given by

$$IF(t_{i_0,k}, W_K, F_{\theta^0}) = \frac{\partial W_K(U_\beta(F_{\theta^0}^{e}))}{\partial \varepsilon}\bigg|_{\varepsilon=0^+} = 0.$$

It, therefore, becomes necessary to consider the second-order IF, as presented in the following results.

**Result 4.3:** The second-order IF of the functional associated with Wald-type test statistics, with respect to the $k$th observation of the $i_0$th group of observations, is given by

$$IF_2(t_{i_0,k}, W_K, F_{\theta^0}) = \frac{\partial^2 W_K(U_\beta(F_{\theta^0}^{e}))}{\partial \varepsilon^2}\bigg|_{\varepsilon=0^+} = 2 IF(t_{i_0,k}, U_\beta, F_{\theta^0})(M^T \theta^0 - m)^T \left[ M^T V(\theta^0) M \right]^{-1} (M^T \theta^0 - m)IF(t_{i_0,k}, U_\beta, F_{\theta^0}),$$

where $IF(t_{i_0,k}, U_\beta, F_{\theta^0})$ is as given in (27).

Similarly, for all the observations, we have the following result.

**Result 4.4:** The second-order IF of the functional associated with Wald-type test statistics, with respect to all the observations, is given by

$$IF_2(t, W_K, F_{\theta^0}) = \frac{\partial^2 W_K(U_\beta(F_{\theta^0}^{e}))}{\partial \varepsilon^2}\bigg|_{\varepsilon=0^+} = 2 IF(t, U_\beta, F_{\theta^0})(M^T \theta^0 - m)^T \left[ M^T V(\theta^0) M \right]^{-1} (M^T \theta^0 - m)IF(t, U_\beta, F_{\theta^0}),$$

where $IF(t, U_\beta, F_{\theta^0})$ is as given in (28).

Note that the second-order influence functions of the proposed Wald-type tests are quadratic functions of the corresponding IFs of the weighted minimum DPD estimator for any type of contamination. Therefore, the boundedness of the IFs of the weighted minimum DPD estimators at $\beta > 0$ also indicates the boundedness of the IFs of the Wald-type test functional, implying their robustness against contamination.

5. Monte carlo simulation study

In this section, an elaborate simulation study is carried out to evaluate the performance of the proposed weighted minimum DPD estimators and the Wald-type tests under the assumption of log-normal lifetimes. We consider different scenarios with different sample sizes and under different degrees of contamination. The results are recorded and averaged over $S = 1000$ simulation runs, using the R statistical software.
5.1. Weighted minimum DPD estimators

We evaluate the robustness of the estimators by means of standardized mean absolute error (SMAE) of some specific parameters of interest for different values of the tuning parameter $\beta \in \{0, 0.2, 0.4, 0.6\}$. After $S$ simulation runs, the SMAE of a parameter of interest $\nu$ is computed as

$$SMAE(\nu) = \frac{1}{S} \sum_{i=1}^{S} \left| \frac{\hat{\nu}_i - \nu^0}{\nu^0} \right|,$$

where $\nu^0$ is the true value of $\nu$ and $\hat{\nu}_i$ is the estimated value of $\nu$ obtained from the $i$th run.

To obtain more information about the inherent distribution of the data, we evaluate devices under different testing conditions (inspection times and stress levels). Note that, in practice, the inspection frequency, number of inspections at each stress level, and allocation of the test devices can be determined from optimal design of the CSALT. See Balakrishnan et al. [24] for more details.

Particularly, in this study, the lifetimes of devices are simulated from the log-normal distribution, under three different stress conditions with one stress factor at three levels, taken to be $\{x_1, x_2, x_3\} = \{30, 40, 50\}$. A balanced data with equal sample size for each group was considered. $K_i$ was taken to range from small to large sample sizes, and the model parameters were set to be $\theta = (a_0, -0.1, -0.6, 0.02)^T$, while $a_0$ was chosen to be 5.8, 6.0 and 6.2 corresponding to devices with low, moderate, and high reliability, respectively. Then, all devices under each stress condition were tested at four different inspection times $\tau = \{\tau_1, \tau_2, \tau_3, \tau_4\}$. As it is important to prevent many zero-observations in test groups, the inspection times were set as $\tau = \{5, 10, 15, 20\}$ for the case of low reliability, $\tau = \{8, 16, 24, 36\}$ for the case of moderate reliability, and $\tau = \{12, 24, 36, 48\}$ for the case of high reliability. Contaminated data were generated by setting $b_0$ as zero. SMAEs of the estimators are presented in Tables 2, 3 and 4 for low, moderate and high reliability, respectively. As contamination increases, the precision of estimation decreases.

Minimum DPD estimators with $\beta > 0$ once again demonstrate that they are more robust than the MLEs.

Finally, the relation of $\beta$ to the sample size is illustrated in the top of Figure 3 for the moderate reliability scenario. As can be seen, there is not a significant difference on how the tuning parameter performs in relation to the sample size.

5.2. Confidence intervals

The coverage probability (CP) and average width (AW) are used for evaluating the 90% confidence intervals of reliability and expected lifetime under normal operating conditions for different sample sizes and the moderate-reliability scenario defined above.
Table 2. SMAE of estimates of some parameters of interest at various levels of reliability and different sample sizes. Low reliability case.

| Size   | True value | Pure data | Contaminated data |
|--------|------------|-----------|-------------------|
|        |            | $\beta = 0$ | $\beta = 0.2$ | $\beta = 0.4$ | $\beta = 0.6$ | $\beta = 0$ | $\beta = 0.2$ | $\beta = 0.4$ | $\beta = 0.6$ |
| $K_i = 50$ |            |            |            |            |            |            |            |            |            |
| $a_0$  | 5.8000     | 0.11043   | 0.10998   | 0.11121   | 0.11244   | 0.19983   | 0.18969   | 0.18078   | 0.17416   |
| $a_1$  | -0.1000    | 0.19862   | 0.19792   | 0.20004   | 0.20264   | 0.37594   | 0.35573   | 0.33816   | 0.32510   |
| $b_0$  | -0.6000    | 0.79721   | 0.79567   | 0.80597   | 0.82651   | 1.37877   | 1.31916   | 1.26482   | 1.22463   |
| $b_1$  | 0.0200     | 0.62687   | 0.62877   | 0.64152   | 0.66236   | 1.33147   | 1.26703   | 1.20589   | 1.15780   |
| $\theta$ |           | 0.43323   | 0.43308   | 0.43968   | 0.45099   | 0.82150   | 0.78290   | 0.74741   | 0.72042   |
| $R(60, 15)$ | 0.6093   | 0.26439   | 0.26402   | 0.26730   | 0.26945   | 0.42266   | 0.40963   | 0.39723   | 0.38748   |
| $E(15)$ | 96.9704    | 0.79662   | 0.91662   | 1.21543   | 2.18463   | 1.13724   | 1.05631   | 0.99568   | 0.95358   |

Table 3. SMAE of estimates of some parameters of interest at various levels of reliability and different sample sizes. Moderate reliability case.

| Size   | True value | Pure data | Contaminated data |
|--------|------------|-----------|-------------------|
|        |            | $\beta = 0$ | $\beta = 0.2$ | $\beta = 0.4$ | $\beta = 0.6$ | $\beta = 0$ | $\beta = 0.2$ | $\beta = 0.4$ | $\beta = 0.6$ |
| $K_i = 50$ |            |            |            |            |            |            |            |            |            |
| $a_0$  | 5.8000     | 0.08459   | 0.08545   | 0.08668   | 0.08685   | 0.17929   | 0.16786   | 0.15790   | 0.15054   |
| $a_1$  | -0.1000    | 0.15300   | 0.15438   | 0.15632   | 0.15649   | 0.33903   | 0.31687   | 0.29735   | 0.28283   |
| $b_0$  | -0.6000    | 0.60755   | 0.60786   | 0.61688   | 0.62660   | 1.31003   | 1.24521   | 1.18285   | 1.13126   |
| $b_1$  | 0.0200     | 0.48175   | 0.48409   | 0.49447   | 0.50574   | 1.29470   | 1.22338   | 1.15544   | 1.09756   |
| $\theta$ |           | 0.33172   | 0.33296   | 0.33859   | 0.34392   | 0.78076   | 0.73833   | 0.69838   | 0.66555   |
| $R(60, 15)$ | 0.6093   | 0.21098   | 0.21319   | 0.21610   | 0.21604   | 0.42420   | 0.40421   | 0.38484   | 0.37303   |
| $E(15)$ | 96.9704    | 0.38567   | 0.38170   | 0.38343   | 0.38624   | 0.71432   | 0.66372   | 0.62268   | 0.59530   |

Table 4. SMAE of estimates of some parameters of interest at various levels of reliability and different sample sizes. High reliability case.
Table 4. SMAE of estimates of some parameters of interest at various levels of reliability and different sample sizes. High reliability case.

| Size   | Pure data | Contaminated data |
|--------|-----------|-------------------|
|        | True value | $\beta = 0$ | $\beta = 0.2$ | $\beta = 0.4$ | $\beta = 0.6$ | $\beta = 0$ | $\beta = 0.2$ | $\beta = 0.4$ | $\beta = 0.6$ |
| $K_i = 30$ | $a_0$ | 6.2000 | 0.08996 | 0.09088 | 0.09138 | 0.09197 | 0.13200 | 0.12833 | 0.12478 | 0.12192 |
|         | $a_1$ | −0.1000 | 0.16250 | 0.16404 | 0.16522 | 0.16642 | 0.24958 | 0.24179 | 0.23460 | 0.22861 |
|         | $b_0$ | −0.6000 | 0.77303 | 0.77453 | 0.78027 | 0.78982 | 1.29934 | 1.26680 | 1.24278 | 1.22847 |
|         | $b_1$ | 0.0200 | 0.58364 | 0.58674 | 0.59442 | 0.60514 | 1.19338 | 1.15206 | 1.11768 | 1.09119 |
|         | $\theta$ | − | 0.40228 | 0.40405 | 0.40782 | 0.41334 | 0.71858 | 0.69725 | 0.67996 | 0.66755 |
| $R(15; 60)$ | 0.7932 | 0.14379 | 0.14452 | 0.14551 | 0.14666 | 0.17712 | 0.17399 | 0.17135 | 0.16947 |
| $E(15)$ | 144.6628 | 0.86043 | 0.86502 | 0.87026 | 0.87939 | 105.452 | 0.94573 | 0.88077 | 0.84203 |
| $K_i = 50$ | $a_0$ | 6.2000 | 0.07261 | 0.07315 | 0.07337 | 0.07333 | 0.10841 | 0.10474 | 0.10181 | 0.09915 |
|         | $a_1$ | −0.1000 | 0.12960 | 0.13039 | 0.13069 | 0.13090 | 0.21525 | 0.17574 | 0.14379 | 0.11189 |
|         | $b_0$ | −0.6000 | 0.63278 | 0.63262 | 0.63835 | 0.64322 | 1.21525 | 1.17574 | 1.14379 | 1.11819 |
|         | $b_1$ | 0.0200 | 0.47713 | 0.47706 | 0.48087 | 0.48600 | 1.14733 | 1.10086 | 1.06085 | 1.02765 |
|         | $\theta$ | − | 0.32803 | 0.32831 | 0.33082 | 0.33336 | 0.66933 | 0.64493 | 0.62457 | 0.60780 |
| $R(15; 60)$ | 0.7932 | 0.12330 | 0.12381 | 0.12731 | 0.12351 | 0.17075 | 0.16613 | 0.16228 | 0.15947 |
| $E(15)$ | 144.6628 | 0.44060 | 0.44289 | 0.44501 | 0.44693 | 0.48049 | 0.46618 | 0.45525 | 0.44650 |
| $K_i = 100$ | $a_0$ | 6.2000 | 0.06201 | 0.06305 | 0.06327 | 0.06349 | 0.10028 | 0.09600 | 0.09247 | 0.08929 |
|         | $a_1$ | −0.1000 | 0.11093 | 0.11248 | 0.11301 | 0.11341 | 0.19164 | 0.18242 | 0.17491 | 0.16837 |
|         | $b_0$ | −0.6000 | 0.54893 | 0.54975 | 0.55234 | 0.55988 | 1.17898 | 1.14104 | 1.10818 | 1.08244 |
|         | $b_1$ | 0.0200 | 0.41805 | 0.41983 | 0.42345 | 0.42988 | 1.13684 | 1.08996 | 1.04886 | 1.01434 |
|         | $\theta$ | − | 0.28498 | 0.28628 | 0.28802 | 0.29167 | 0.65193 | 0.62736 | 0.60610 | 0.58861 |
| $R(15; 60)$ | 0.7932 | 0.10550 | 0.10682 | 0.10730 | 0.10785 | 0.16873 | 0.16354 | 0.15927 | 0.15546 |
| $E(15)$ | 144.6628 | 0.34205 | 0.34669 | 0.34836 | 0.35068 | 0.40811 | 0.39425 | 0.38297 | 0.37325 |

Figure 2. SMAE of the parameter vector $\theta$ for different levels of contamination.

The results presented in Tables 5 and 6 reveal that the logit and arsech approaches outperform the asymptotic confidence interval for reliability, for both pure and contaminated data. For the expected lifetime, there is not a significant difference between the asymptotic confidence interval and the log approach. But, the latter is slightly more robust, but is accompanied with an increase in the AW.
Figure 3. SMAE of the parameter vector (above), empirical levels (middle) and powers (below) for pure (left) and contaminated (right) data.
Table 5. Coverage probabilities for reliabilities and expected lifetime at different sample sizes.

| Size   | True value | Method  | 0.2 | 0.4 | 0.6   | 0.2 | 0.4 | 0.6   |
|--------|------------|---------|-----|-----|-------|-----|-----|-------|
| $K_i = 50$ | R(15;60) 0.7080 asy | 0.787 | 0.794 | 0.790 | 0.785 | 0.539 | 0.566 | 0.582 |
|        | R(15;60) 0.7080 logit | 0.891 | 0.896 | 0.892 | 0.888 | 0.977 | 0.974 | 0.953 |
|        | R(15;60) 0.7080 arsech | 0.862 | 0.858 | 0.848 | 0.844 | 0.630 | 0.656 | 0.678 |
|        | E(15) 118.4399 asy | 0.881 | 0.881 | 0.884 | 0.880 | 0.817 | 0.829 | 0.840 |
|        | E(15) 118.4399 logit | 0.828 | 0.823 | 0.825 | 0.828 | 0.843 | 0.836 | 0.834 |

| $K_i = 80$ | R(15;60) 0.7080 asy | 0.817 | 0.820 | 0.820 | 0.816 | 0.497 | 0.527 | 0.574 |
|        | R(15;60) 0.7080 logit | 0.901 | 0.903 | 0.902 | 0.903 | 0.999 | 0.991 | 0.987 |
|        | R(15;60) 0.7080 arsech | 0.873 | 0.866 | 0.866 | 0.868 | 0.624 | 0.658 | 0.681 |
|        | E(15) 118.4399 asy | 0.889 | 0.886 | 0.884 | 0.887 | 0.821 | 0.831 | 0.833 |
|        | E(15) 118.4399 logit | 0.857 | 0.855 | 0.855 | 0.853 | 0.878 | 0.878 | 0.872 |

| $K_i = 100$ | R(15;60) 0.7080 asy | 0.836 | 0.836 | 0.833 | 0.828 | 0.478 | 0.518 | 0.545 |
|        | R(15;60) 0.7080 logit | 0.911 | 0.916 | 0.913 | 0.910 | 0.999 | 0.996 | 0.996 |
|        | R(15;60) 0.7080 arsech | 0.878 | 0.875 | 0.873 | 0.869 | 0.574 | 0.606 | 0.633 |
|        | E(15) 118.4399 asy | 0.895 | 0.895 | 0.893 | 0.892 | 0.814 | 0.825 | 0.837 |
|        | E(15) 118.4399 logit | 0.861 | 0.861 | 0.861 | 0.859 | 0.890 | 0.889 | 0.881 |

Table 6. Average widths of confidence intervals for reliabilities and expected lifetime at different sample sizes.

| Size   | True value | Method  | 0.2 | 0.4 | 0.6   | 0.2 | 0.4 | 0.6   |
|--------|------------|---------|-----|-----|-------|-----|-----|-------|
| $K_i = 50$ | R(15;60) 0.7080 asy | 0.362 | 0.362 | 0.364 | 0.365 | 0.325 | 0.334 | 0.341 |
|        | R(15;60) 0.7080 logit | 0.378 | 0.380 | 0.383 | 0.386 | 0.545 | 0.524 | 0.511 |
|        | R(15;60) 0.7080 arsech | 0.310 | 0.310 | 0.310 | 0.310 | 0.233 | 0.242 | 0.249 |
|        | E(15) 118.4399 asy | 148.30 | 148.33 | 149.15 | 150.60 | 161.66 | 159.13 | 157.19 |
|        | E(15) 118.4399 logit | 162.09 | 162.06 | 163.01 | 164.71 | 181.40 | 177.53 | 174.64 |

| $K_i = 80$ | R(15;60) 0.7080 asy | 0.304 | 0.305 | 0.307 | 0.308 | 0.270 | 0.280 | 0.288 |
|        | R(15;60) 0.7080 logit | 0.310 | 0.312 | 0.314 | 0.316 | 0.452 | 0.444 | 0.438 |
|        | R(15;60) 0.7080 arsech | 0.275 | 0.276 | 0.277 | 0.278 | 0.216 | 0.226 | 0.234 |
|        | E(15) 118.4399 asy | 107.45 | 107.54 | 107.90 | 108.56 | 120.73 | 118.82 | 117.21 |
|        | E(15) 118.4399 logit | 113.60 | 113.70 | 114.13 | 114.88 | 129.99 | 127.50 | 125.47 |

| $K_i = 100$ | R(15;60) 0.7080 asy | 0.280 | 0.281 | 0.282 | 0.284 | 0.246 | 0.256 | 0.265 |
|        | R(15;60) 0.7080 logit | 0.282 | 0.283 | 0.285 | 0.287 | 0.398 | 0.392 | 0.387 |
|        | R(15;60) 0.7080 arsech | 0.258 | 0.259 | 0.260 | 0.261 | 0.205 | 0.214 | 0.223 |
|        | E(15) 118.4399 asy | 90.93 | 90.87 | 91.264 | 91.68 | 105.06 | 103.29 | 101.79 |
|        | E(15) 118.4399 logit | 95.06 | 95.01 | 95.43 | 95.89 | 111.62 | 109.50 | 107.67 |

5.3. Wald-type tests

To evaluate the performance of the proposed Wald-type tests, we consider the scenario of moderate reliability mentioned in the last section. We consider the testing problem

$$H_0 : a_0 = 6.0 \text{ against } H_1 : a_0 \neq 6.0.$$
Figure 4. Reliability and hazard functions, expected lifetime and variance of one-shot device testing data generated under Log($6, -0.1, -0.6, 0.02$) and Wei($6.5, -0.11, 0.2, 0$) under different inspection times and stress levels.

We first evaluate the empirical levels, measured as the proportion of test statistics exceeding the corresponding chi-square critical value for a nominal significance level of $\alpha = 0.05$. The empirical powers are computed in a similar way, with $a_0 = 5$. The corresponding results are shown in Figure 3 for both pure and contaminated data (left and right, respectively). To contaminate the data here, we set $\tilde{a}_0$ as 5.6. For pure data, the whole family of Wald-type tests has their levels very close to the nominal level. However, when the data get contaminated, Wald-type tests with $\beta > 0$ have much more stable robustness properties. With respect to the power, the classical Wald test outperforms other tests for pure data, but is less robust when contamination is present in the data.

5.4. Model misspecification

Let us consider now that our data are generated from the log-normal distribution with parameters $\theta = (6, -0.1, -0.6, 0.02)^T$ or the Weibull distribution with parameters $\theta =$
Table 7. AIC values under the model misspecification scenarios.

| True Distribution      | $K_i=50$ | $K_i=60$ | $K_i=70$ | $K_i=80$ | $K_i=90$ | $K_i=100$ | $K_i=110$ | $K_i=120$ |
|------------------------|----------|----------|----------|----------|----------|----------|----------|----------|
| Weibull distribution   | 0.658    | 0.648    | 0.651    | 0.679    | 0.712    | 0.707    | 0.744    | 0.727    |
| Lognormal distribution | 0.975    | 0.976    | 0.984    | 0.987    | 0.990    | 0.994    | 0.997    | 0.990    |

As can be seen in Figure 4, both models present similar reliability and hazard functions, and similar expected lifetime and variance under some particular normal testing conditions. Table 7 presents the probability of correct selection based on the AIC given in (15). AIC works very well when the data follow the log-normal distribution. It presents more difficulties when the data are generated from Weibull distribution, but improves for large sample sizes.

6. Illustrative examples

In this section, we present two real-life examples to illustrate all the methods of inference developed in the preceding sections:

**Electro-explosive device data:** In this one-shot device testing experiment described in Balakrishnan and Ling [12] (Table 1), the number of failures of 90 electro-explosive devices tested at temperatures $x_i \in \{35, 45, 55\}$ and inspection times $\tau_i \in \{10, 20, 30\}$ were collected. In this example, we have only one stress factor, with $I = 9$ and $K_i = 10$, for $i = 1, \ldots, I$.

**Electric current data:** These data, presented in Balakrishnan and Ling [12] (Table 2), present the number of failures on 120 devices placed under 12 different conditions in the experiment: four accelerated conditions with higher-than-normal temperature $x_{i1} \in \{55, 85\}$ and electric current $x_{i2} \in \{70, 100\}$, and three different inspection times $\tau_i \in \{2, 5, 8\}$. In this case, we have $J = 2$, $I = 12$ and $K_i = 10$, for $i = 1, \ldots, I$.

We apply the AIC presented in (15) to compare Weibull and log-normal distributions for fitting these data. For the electro-explosive device data, AIC of log-normal model is substantially lower than the AIC of the Weibull model. In particular, based on the complete likelihood, AIC(log) = 114.92 and AIC(Weibull) = 123.37. On the other hand, for the electric current data example, AIC of the Weibull model is slightly lower than AIC of the log-normal model. In particular, AIC(log) = 127.51 and AIC(Weibull) = 127.48. However, the difference is negligible to have a clear conclusion.

Weighted minimum DPD estimators under log-normal distribution were computed for different values of the tuning parameter $\beta$ and the estimated values of some specific parameters of interest are presented in Table 8 (Electro-explosive data) and Table 9 (Electric current data). For the CI of reliability, we used the logit-approach, while for the CI of mean lifetime, we used the log-approach. For both examples, the predicted probabilities are compared to the observed ones (left of Figure 5).

In the preceding discussion, we noted that the robustness feature of the proposed weighted minimum DPD estimators usually increases with increasing $\beta$; but, the efficiency decreases slightly. It seems, therefore, that a moderately large value of $\beta$ would provide the best trade-off for possibly contaminated data. However, a data-driven choice of $\beta$ may be more helpful in practice. In the context of one-shot device testing, Castilla and Chocano
### Table 8. Inferential results for electro-explosive device data.

| Parameter | Estimate | Confidence Interval | Estimate | Confidence Interval |
|-----------|----------|---------------------|----------|---------------------|
| $a_0$     | 4.78801  | (3.21825, 6.35778)  | 4.78801  | (3.21861, 6.35742)  |
| $a_1$     | -0.04305 | (-0.07661, 0.00949) | -0.04308 | (-0.07665, -0.00952)|
| $b_0$     | 0.80430  | (-1.93904, 3.54764) | 0.8043   | (-1.94143, 3.55002) |
| $b_1$     | -0.01833 | (-0.07575, 0.03910) | -0.01833 | (-0.07583, 0.03917) |
| $R(25;10)$| 0.84059  | (0.34212, 0.98164)  | 0.84049  | (0.34145, 0.98167)  |
| $E(25)$   | 111.16472| (5.27000, 2344.89266)| 111.02344| (5.27234, 2337.89866)|

| Parameter | Estimate | Confidence Interval | Estimate | Confidence Interval |
|-----------|----------|---------------------|----------|---------------------|
| $a_0$     | 4.78858  | (3.25352, 6.32364)  | 4.78693  | (3.27006, 6.30380)  |
| $a_1$     | -0.04325 | (-0.07645, -0.01006)| -0.04329 | (-0.07628, -0.01031)|
| $b_0$     | 0.60442  | (-2.08826, 3.29709) | 0.50556  | (-2.16925, 3.18036) |
| $b_1$     | -0.01420 | (-0.07111, 0.03910) | -0.0122  | (-0.06901, 0.04461) |
| $R(25;10)$| 0.86317  | (0.32925, 3.29709)  | 0.84767  | (0.47042, 0.97257)  |
| $E(25)$   | 92.80086 | (7.57122, 1137.46539)| 85.75138 | (8.77878, 837.62209)|

### Table 9. Inferential results for electric current data.

| Parameter | Estimate | Confidence Interval | Estimate | Confidence Interval |
|-----------|----------|---------------------|----------|---------------------|
| $a_0$     | 6.91992  | (4.61124, 9.22859)  | 7.04428  | (4.59463, 9.49392)  |
| $a_1$     | -0.03979 | (-0.05979, -0.01980)| -0.04062 | (-0.06159, -0.01965)|
| $a_2$     | -0.03734 | (-0.05785, -0.01683)| -0.03820 | (-0.05989, -0.01651)|
| $b_0$     | 0.01003  | (-0.01212, 0.03219) | 0.00914  | (-0.01358, 0.03187) |
| $b_1$     | 0.01354  | (-0.00837, 0.03546) | 0.01314  | (-0.00936, 0.03563) |
| $b_2$     | 0.01354  | (-0.00837, 0.03546) | 0.01314  | (-0.00936, 0.03563) |
| $R(25,35,60)$| 0.94600 | (0.00124, 100.000)  | 0.95153  | (0.00153, 100.000)  |
| $R(35,50,60)$| 0.46376 | (0.01032, 0.98625)  | 0.52216  | (0.01568, 0.98684)  |
| $E(25,35)$| 106.83129| (27.56069, 414.10155)| 116.33824| (27.38673, 494.20227)|

[28] studied different methods for the the choice of the ‘optimal’ tuning parameter. The Warwick and Jones procedure [29] and its iterative form [30] propose a minimization of the asymptotic mean square error (MSE) through the computation of the estimated asymptotic variance-covariance matrix of the model parameters in a grid of tuning parameters. However, as seen in Castilla and Chocano [28], minimizing a loss function which relates empirical and theoretical probabilities may provide an easy-to-compute criteria. Particularly, the minimization of the maximum of the absolute errors (MaxAE) is seen to work well. We compute the MaxAE and RMSE for our examples, and we observe how for the Electric current data, the estimation with higher tuning parameters clearly reduces the MaxAE.
7. Conclusion and discussion

The aim of the present research was to develop robust statistical inference for one-shot device testing model under log-normal distribution. We have presented a family of weighted minimum DPD estimators, as a generalization of the classical MLE. Based on this family, we are also able to develop confidence intervals and a family of Wald-type tests, which generalizes the classical Wald test. The robustness of the proposed statistics has been evaluated through the study of Influence Function and an extensive simulation study. Finally, the methods developed here have been applied to two real-life examples for illustrative purpose. The evidence from this study suggests the use of these methods as a more robust alternative to those based on the MLE, with an unavoidable loss of efficiency in the presence of contamination on the data.

This research has raised some questions that need further investigation. The model misspecification has been focused to distinguish between Weibull and log-normal
distributions. A more detailed study may be done by including other important lifetime distribution candidates.

On the other hand, although the study was limited to a binary response (failure or success of the device when being tested), it may be of interest to consider the ‘competing risks’ scenario, in which we assume the products under study can experience any one of various, say $R$, possible causes of failure. In this case, the likelihood function of the complete data is given by

$$
\mathcal{L}_c(\theta) = \prod_{i=1}^{I} \prod_{k=1}^{K_i} \prod_{r=1}^{R} f_{i0}(\omega^{(\delta_{ik})}_{rik}; \theta),
$$

where $\delta_{ik}$ is an indicator of the survival/failure cause of the $k$th device in the $i$th condition.

As with the binary case, we do not know the true log-lifetimes $\omega_{rik}$. So an EM algorithm may be developed for the estimation of the parameter vector in (32), with the difficulty of computation of expectations of log-lifetimes conditioned on different causes of failure. This is an interesting problem for further consideration. On the other hand, the observed likelihood function is

$$
\mathcal{L}(\theta) \propto \prod_{i=1}^{I} \pi_{i0}(\theta)^{n_{i0}} \left\{ \prod_{r=1}^{R} \pi_{ir}(\theta)^{n_{ir}} \right\},
$$

where $\pi_{i0}$ and $n_{i0}$ are the probability of survival and number of survivals in the $i$th condition, respectively, and $\pi_{i0}$ and $n_{ir}$ are the probability of failure and the number of failures due to cause $r$, respectively. Divergence-based robust estimation methods could also be developed using an approach similar to the one developed in this paper. We are currently working on these problems and hope to report the findings in a future paper.

**Acknowledgments**

We would like to thank the referees for their helpful comments and suggestions on an earlier version of this manuscript. Their comments have improved the paper.

**Disclosure statement**

No potential conflict of interest was reported by the author(s).

**ORCID**

N. Balakrishnan  
http://orcid.org/0000-0001-5842-8892

E. Castilla  
http://orcid.org/0000-0002-9626-6449

**References**

[1] Ling MH. Inference for one-shot device testing data [PhD thesis]. Hamilton, Ontario, Canada: McMaster University; 2012. http://hdl.handle.net/11375/12377.

[2] So HY. Some inferential results for one-shot device testing data analysis [PhD thesis]. Hamilton, Ontario, Canada: McMaster University; 2016. http://hdl.handle.net/11375/19438.
Ling MH, Balakrishnan N. Model mis-specification analyses of Weibull and gamma models based on one-shot device test data. IEEE Trans Reliab. 2017;66(3):641–650. doi: 10.1109/TR.2017.2703111

Balakrishnan N, Castilla E, Ling MH. Optimal designs of constant-stress accelerated life-tests for one-shot devices with model misspecification analysis. Qual Reliab Eng Int. 2021;38(2):989–1012. doi: 10.1002/qre.v38.2

Basu A, Harris IR, Hjort NL, et al. Robust and efficient estimation by minimising a density power divergence. Biometrika. 1998;85(3):549–559. doi: 10.1093/biomet/85.3.549

Balakrishnan N, Castilla E, Ling MH. Optimal designs of constant-stress accelerated life-tests for one-shot devices with model misspecification analysis. Qual Reliab Eng Int. 2021;38(2):989–1012. doi: 10.1002/qre.v38.2

Basu A, Harris IR, Hjort NL, et al. Robust and efficient estimation by minimising a density power divergence. Biometrika. 1998;85(3):549–559. doi: 10.1093/biomet/85.3.549

Balakrishnan N, Castilla E, Ling MH. Optimal designs of constant-stress accelerated life-tests for one-shot devices with model misspecification analysis. Qual Reliab Eng Int. 2021;38(2):989–1012. doi: 10.1002/qre.v38.2

Basu A, Harris IR, Hjort NL, et al. Robust and efficient estimation by minimising a density power divergence. Biometrika. 1998;85(3):549–559. doi: 10.1093/biomet/85.3.549

Balakrishnan N, Castilla E, Ling MH. Optimal designs of constant-stress accelerated life-tests for one-shot devices with model misspecification analysis. Qual Reliab Eng Int. 2021;38(2):989–1012. doi: 10.1002/qre.v38.2

Basu A, Harris IR, Hjort NL, et al. Robust and efficient estimation by minimising a density power divergence. Biometrika. 1998;85(3):549–559. doi: 10.1093/biomet/85.3.549

Balakrishnan N, Castilla E, Ling MH. Optimal designs of constant-stress accelerated life-tests for one-shot devices with model misspecification analysis. Qual Reliab Eng Int. 2021;38(2):989–1012. doi: 10.1002/qre.v38.2

Basu A, Harris IR, Hjort NL, et al. Robust and efficient estimation by minimising a density power divergence. Biometrika. 1998;85(3):549–559. doi: 10.1093/biomet/85.3.549

Balakrishnan N, Castilla E, Ling MH. Optimal designs of constant-stress accelerated life-tests for one-shot devices with model misspecification analysis. Qual Reliab Eng Int. 2021;38(2):989–1012. doi: 10.1002/qre.v38.2

Basu A, Harris IR, Hjort NL, et al. Robust and efficient estimation by minimising a density power divergence. Biometrika. 1998;85(3):549–559. doi: 10.1093/biomet/85.3.549

Balakrishnan N, Castilla E, Ling MH. Optimal designs of constant-stress accelerated life-tests for one-shot devices with model misspecification analysis. Qual Reliab Eng Int. 2021;38(2):989–1012. doi: 10.1002/qre.v38.2

Basu A, Harris IR, Hjort NL, et al. Robust and efficient estimation by minimising a density power divergence. Biometrika. 1998;85(3):549–559. doi: 10.1093/biomet/85.3.549

Balakrishnan N, Castilla E, Ling MH. Optimal designs of constant-stress accelerated life-tests for one-shot devices with model misspecification analysis. Qual Reliab Eng Int. 2021;38(2):989–1012. doi: 10.1002/qre.v38.2

Basu A, Harris IR, Hjort NL, et al. Robust and efficient estimation by minimising a density power divergence. Biometrika. 1998;85(3):549–559. doi: 10.1093/biomet/85.3.549

Balakrishnan N, Castilla E, Ling MH. Optimal designs of constant-stress accelerated life-tests for one-shot devices with model misspecification analysis. Qual Reliab Eng Int. 2021;38(2):989–1012. doi: 10.1002/qre.v38.2

Basu A, Harris IR, Hjort NL, et al. Robust and efficient estimation by minimising a density power divergence. Biometrika. 1998;85(3):549–559. doi: 10.1093/biomet/85.3.549

Balakrishnan N, Castilla E, Ling MH. Optimal designs of constant-stress accelerated life-tests for one-shot devices with model misspecification analysis. Qual Reliab Eng Int. 2021;38(2):989–1012. doi: 10.1002/qre.v38.2

Basu A, Harris IR, Hjort NL, et al. Robust and efficient estimation by minimising a density power divergence. Biometrika. 1998;85(3):549–559. doi: 10.1093/biomet/85.3.549

Appendix. Proof of Results

A.1 Proof of result 3.1

Proof: We have

\[
\sum_{i=1}^{l} \frac{K_i}{K} d_{KL}(\hat{\theta}_i, \pi_i(\theta)) = \sum_{i=1}^{l} \frac{n_i}{K} \log \left( \frac{n_i}{K} \right) F_{\omega}(W_i; x_i, \theta) + \frac{K_i - n_i}{K} \log \left( \frac{K_i - n_i}{K_i} \right) R_{\omega}(W_i; x_i, \theta)
\]

\[
= c - \frac{1}{K} \sum_{i=1}^{l} \left\{ n_i \log \left( F_{\omega}(W_i; x_i, \theta) \right) + (K_i - n_i) \log \left( R_{\omega}(W_i; x_i, \theta) \right) \right\}
\]

\[
= c - \frac{1}{K} \log \left( \prod_{i=1}^{l} F_{\omega}^{n_i}(W_i; x_i, \theta) R_{\omega}^{K_i-n_i}(W_i; x_i, \theta) \right)
\]

\[
= c - \frac{1}{K} \log L(\theta; z),
\]

where \(c = \sum_{i=1}^{l} \frac{n_i}{K} \log \left( \frac{n_i}{K} \right) + \frac{K_i - n_i}{K_i} \log \left( \frac{K_i - n_i}{K_i} \right)\) does not depend on the parameter \(\theta\). Therefore, the maximization of the likelihood is equivalent to the minimization of the weighted Kullback-Leibler divergence measure between probability vectors in (16) and (17).

A.2 Proof of result 3.4

Proof: Let us denote

\[
u_{ij}(\theta) = \frac{\partial \log \pi_{ij}(\theta)}{\partial \theta} = \frac{1}{\pi_{ij}(\theta)} \left( \frac{\partial \pi_{ij}(\theta)}{\partial \theta} \right) = \frac{1}{\pi_{ij}(\theta)} \delta_i x_i,
\]

with \(\delta_i\) as given in (9). Upon using Theorem 3.1 of Ghosh and Basu [26], we have

\[
\sqrt{K} \left( \hat{\beta} - \theta^0 \right) \xrightarrow{K \to \infty} N \left( 0_{2^{(l+1)}}, J^{-1}(\theta^0) K^\beta(\theta^0) J^{-1}(\theta^0) \right)
\]
where
\[
J_\beta(\theta) = \sum_{i=1}^{I} \sum_{j=1}^{2} \frac{K_i}{K} u_{ij}(\theta) u_{ij}^T(\theta) \pi_{ij}^{\beta+1}(\theta),
\]
\[
K_\beta(\theta) = \left( \sum_{i=1}^{I} \sum_{j=1}^{2} \frac{K_i}{K} u_{ij}(\theta) u_{ij}^T(\theta) \pi_{ij}^{2\beta+1}(\theta) - \sum_{i=1}^{I} \frac{K_i}{K} \xi_{i,\beta}(\theta) \xi_{i,\beta}^T(\theta) \right),
\]
with
\[
\xi_{i,\beta}(\theta) = \sum_{j=1}^{2} u_{ij}(\theta) \pi_{ij}^{\beta+1}(\theta) = \delta_i x_i \sum_{j=1}^{2} (-1)^{i+1} \pi_{ij}^{\beta}(\theta).
\]
Now, we have \(u_{ij}(\theta) u_{ij}^T(\theta) = \frac{1}{\pi_{ij}(\theta)} \Delta_i x_i x_i^T\), with \(\Delta_i = \delta_i \delta_i^T\). It then follows that
\[
J_\beta(\theta) = \sum_{i=1}^{I} \frac{K_i}{K} \Delta_i \left( \sum_{j=1}^{2} \pi_{ij}^{\beta-1}(\theta) - \left( \sum_{j=1}^{2} (-1)^{i+1} \pi_{ij}^{\beta}(\theta) \right) \right) x_i x_i^T.
\]
In a similar manner, \(\xi_{i,\beta}(\theta) \xi_{i,\beta}^T(\theta) = \Delta_i \left( \sum_{j=1}^{2} (-1)^{i+1} \pi_{ij}^{\beta}(\theta) \right)^2\) and
\[
K_\beta(\theta) = \sum_{i=1}^{I} \frac{K_i}{K} \Delta_i \left( \sum_{j=1}^{2} \pi_{ij}^{2\beta-1}(\theta) - \left( \sum_{j=1}^{2} (-1)^{i+1} \pi_{ij}^{\beta}(\theta) \right)^2 \right) x_i x_i^T.
\]
Because
\[
\sum_{j=1}^{2} \pi_{ij}^{2\beta-1}(\theta) - \left( \sum_{j=1}^{2} (-1)^{i+1} \pi_{ij}^{\beta}(\theta) \right)^2 = \pi_{i1}(\theta) \pi_{i2}(\theta) \left( \pi_{i1}^{\beta-1}(\theta) + \pi_{i2}^{\beta-1}(\theta) \right)^2,
\]
we have
\[
K_\beta(\theta) = \sum_{i=1}^{I} \frac{K_i}{K} \Delta_i \pi_{i1}(\theta) \pi_{i2}(\theta) \left( \pi_{i1}^{\beta-1}(\theta) + \pi_{i2}^{\beta-1}(\theta) \right)^2 x_i x_i^T.
\]

### A.3 Proof of result 3.5

**Proof:** Let \(\theta^0 \in \Theta_0\) be the true value of parameter \(\theta\). It is then clear that
\[
m(\hat{\theta}_\beta) = m(\theta^0) + M^T(\hat{\theta}_\beta)(\hat{\theta}_\beta - \theta^0) + o_p\|\theta^0 - \theta^0\| \rightarrow M^T(\hat{\theta}_\beta)(\hat{\theta}_\beta - \theta^0) + o_p(K^{-1/2}).
\]
But, \(\sqrt{K}(\hat{\theta}_\beta - \theta^0) \xrightarrow{L} N(0_{j+1}, \Sigma_{\beta}(\hat{\theta}_\beta))\). Therefore, we have
\[
\sqrt{K} m(\hat{\theta}_\beta) \xrightarrow{L} N\left(0_r, M^T(\theta^0) \Sigma_{\beta}(\theta^0) M(\theta^0)\right)
\]
and taking into account that \(\text{rank}(M(\theta^0)) = r\), we readily obtain
\[
K m(\hat{\theta}_\beta) \left( M^T(\theta^0) \Sigma_{\beta}(\theta^0) M(\theta^0) \right)^{-1} m(\hat{\theta}_\beta) \xrightarrow{L} \chi^2_r.
\]
As \((M^T(\hat{\theta}_\beta) \Sigma_{\beta}(\hat{\theta}_\beta) M(\hat{\theta}_\beta))^{-1}\) is a consistent estimator of \((M^T(\theta^0) \Sigma_{\beta}(\theta^0) M(\theta^0))^{-1}\), we get
\[
W_K(\hat{\theta}_\beta) \xrightarrow{L} \chi^2_r.
\]