Recent Advances in Robust Design for Accelerated Failure Time Models with Type I Censoring

María J. Rivas-López 1, Raúl Martín-Martín 2,* and Irene García-Camacha Gutiérrez 2

1 Department of Statistics, University of Salamanca, Plaza de los Caidos, 37008 Salamanca, Spain; chusrl@usal.es
2 Department of Mathematics, University of Castilla-La Mancha, Avda. Carlos III s/n, 45071 Toledo, Spain; Irene.GarciaCamacha@uclm.es
* Correspondence: Raul.MMartin@uclm.es; Tel.: +34-925-268-800 (ext. 5375)

Abstract: Many fields including clinical and manufacturing areas usually perform life-testing experiments and accelerated failure time models (AFT) play an essential role in these investigations. In these models the covariate causes an accelerant effect on the course of the event through the term named acceleration factor (AF). Despite the influence of this factor on the model, recent studies state that the form of AF is weakly or partially known in most real applications. In these cases, the classical optimal design theory may produce low efficient designs since they are highly model dependent. This work explores planning and techniques that can provide the best robust designs for AFT models with type I censoring when the form of the AF is misspecified, which is an issue little explored in the literature. Main idea is focused on considering the AF to vary over a neighbourhood of perturbation functions and assuming the mean square error matrix as the basis for measuring the design quality. A key result of this research was obtaining the asymptotic MSE matrix for type I censoring under the assumption of known variance regardless the selected failure time distribution. In order to illustrate the applicability of previous result to a study case, analytical characterizations and numerical approaches were developed to construct optimal robust designs under different contaminating scenarios for a failure time following a log-logistic distribution.

Keywords: acceleration factor; type-I censoring; asymptotic mean squared error; log-logistic distribution; robust design; optimal experimental design; survival models

1. Introduction

Failure time data, also called time-to-event data or survival data, arise in many areas, including medical follow-up studies and reliability research. In these investigations often the event times are not always exactly observed because some units under observation may not experience the event of interest; that is, the data are censored and partial information from them is available. Different circumstances can produce different types of censoring. A right-censored failure time implies that the event is observed in a time greater than the last known survival time, either due to the end of the study or lose to follow-up. On the contrary, left censoring occurs when the actual survival time of an event is less than that observed. This kind of censoring occurs far less commonly than right censoring [1]. Another class of censored data is interval censoring, which is produced when the event is imprecisely experienced within an interval of time. The emphasis of this work will be on the analysis of data from an experiment which stops at a predetermined time, so that the survival time of any subjects remaining at this point are right-censored. The latter is also known as type I censoring.

Although the Cox’s Proportional Hazards (PH) model is commonly applied to analyze these data, it can generate severe problems when proportional hazards assumption is violated. Accelerated Failure Time (AFT) models have been considered as an alternative to PH models [2]. The AFT model has long been used in strength testing of materials,
in which the units are exposed to higher than usual levels of stress to observe earlier failures. The type of distribution of the time-to-event variable has to be specified, and the objective is to estimate the distribution of the failure time under usual conditions through the observed distribution of the failure time under accelerated conditions. Thus, AFT models provide a more appropriate modeling framework in many applications. A comprehensive and detailed exposition on accelerated failure testing can be found in [3].

Trial design is a primarily relevant stage in survival studies since practical limitations may involve time-consuming and expensive follow-ups [4]. In particular, planning decisions in AFT experiments are challenging due to the practical limitations of the study cases characterized by data censoring, non-normal distributions, non-linearity of the response and uncertainty about the AF form, to name a few. Thus, new design strategies are needed to overcome previous difficulties. Design issues in AFT experiments using the basis of Optimal Experimental Design (OED) theory were investigated for very rigid examples of particular distributions of the failure time (see [5–7] for Poisson, Weibull and Log-Normal distribution, respectively). Later, Müller [8] gave some theoretical results to construct local \( D \)-optimal designs considering the exponential distribution and based on equivalence theorem ideas. Rivas-López et al. [9] extended this theory obtaining a general expression of the information matrix for type I censoring regardless the assumed distribution. Recent real applications proposed sophisticated design plans fitted to the special features of their experimental applications [10,11]. Differently from OED techniques, other authors consider more rudimentary design strategies based on classical design [12], progressive censoring [13] or adaptive trials [14].

Robustness in the context of experimental design may be conceived from different approaches. Focusing on AFT models, robust designs were constructed to seek protection against misspecification of: (1) model parameter values, (2) the assumed distribution, and (3) the underlying stress–life relationship. The latter type of robustness is crucial for AFT models because the use level is often below the range of stress levels at which tests are conducted [15]. The interested reader is referred to [15] for an extensive review aimed at these three streams of research. Previous approaches have been addressed in the literature following two strategies mainly: (a) seeking a well design performance under different candidate models, and (b) assuming that the model form is only approximate and unknown departures, but controlled in some way, from the considered model may occur during the experiments. In the first vein, it should be noted the works due to Ginebra and Sen (1998) [16] and Pascual and Montepiedra (2002) [17]. The former computed minimax AFT designs when the parameter values are misspecified, whereas the latter discussed design issues comprising uncertainty on lifetime distribution. The main criticism of this strategy is that emphasizes on the design efficiency rather than on directly addressing the bias as a result of not choosing the “true” model [15]. On the other hand, there is a vast theory following (b), initiated by Box and Draper in 1959 [18] and latter extended in depth by Wiens [19–21] for other classes of models. In a survival context, this idea has been previously explored considering prediction and extrapolation problems for censored data [22] and the PH model [23]. As far as the authors know, the only existing work for AFT models under this perspective is due to Pascual (2006) [15]. In this work the author investigated the bias produced when a linear model is used to estimate life quantiles in AFT but the “true” one is quadratic.

Motivated by the gap in literature on the construction of robust designs for AFT models for a wider and not-predetermined class of possible departures from the considered model when AF is partially known, the present work arises. On the other hand, Nelson [24] holds that the Acceleration Factor (AF) is only partially known in most of practical cases, so that a better analysis would include the uncertainty in it. Under this framework, this work provides theory and methods for constructing robust designs when the AF is misspecified and the variance of the time-to-event variable is known. The approach adopted in this research lies in allowing the AF to vary over a neighborhood of possible departures of the considered form of the AF. The class of functions proposed by Huber [25] will be considered
in this research for covering a sufficiently wide range of plausible alternate models and being the most commonly used in bias-robust literature. Construction strategies developed in this work were set for the example presented in [26], in which the log-logistic AFT model exhibited a better performance regarding goodness of fit compared with other life-testing models.

The next section contains the main features and the notation used for AFT models. The incorporation of the contaminating function in the AF and its implications over the design problem is also included in this section. The key result of this work is presented in Section 3: an analytical expression of the asymptotic mean squared error matrix for type I censoring which does not depend on the assumed distribution. In order to illustrate the applicability of previous result to a study case, optimal robust designs are constructed under different contaminating scenarios for the log-logistic distribution in Section 4. Finally, a brief discussion about the obtained results as well as the open lines of research are presented in Section 5.

2. Accelerated Failure Time Models

The AFT models look for the parameters of a probability distribution for a time-to-event variable, $T$, that is affected by some covariates $X = (X_1, X_2, \ldots, X_k)$ giving the conditions under which the experiment runs. Let $T_u$ be the random variable $T$ observed in the usual levels of the covariates (for simplicity $X_i = 0, \forall i = 1, 2, \ldots, k$) following a distribution given by a baseline survival function, $S_u$. An AFT model is usually defined through this survival function,

$$S(t | x, \theta) = S_u(t / AF(x, \theta)),$$

where $AF(x, \theta)$ is an acceleration factor, which speeds up the effect of the survival time in the survival function.

The AF is typically the exponential of a linear relationship,

$$AF(x, \theta) = \exp(\theta'x).$$

Alternatively, the AFT model can be expressed through its log-linear form with respect to time as (see [27] for details)

$$\log(T) = \mu + \log(AF(x, \theta)) + \sigma \epsilon = \mu + \theta'x + \sigma \epsilon$$

where $\mu \equiv E(\log T_u)$ is the intercept, $\sigma \equiv \text{Var}(\log T_u)$ is the scale parameter and $\epsilon$ is a random variable with a particular distribution with mean 0 and variance 1. It is assumed that the random variable $T$, for any values of the covariates $x$, follows the same type of distribution than $T_u$.

The survival function of $T$ expressed in terms of the survival function of $\epsilon$ is

$$S(t | x; \mu, \theta) = S_\epsilon \left( \frac{\log(t) - \mu - \theta'x}{\sigma} \right) = S_\epsilon(z_t),$$

where it is defined

$$z_t = \frac{\log(t) - \mu - \theta'x}{\sigma}.$$

From the above, the corresponding relationship between the Probability Density Functions (PDF) of $T$ and $\epsilon$ is

$$f(t | x; \mu, \theta) = \frac{1}{\sigma |t|} \cdot f_\epsilon \left( \frac{\log(t) - \mu - \theta'x}{\sigma} \right) = \frac{1}{\sigma |t|} \cdot f_\epsilon(z_t).$$

Members of the AFT model class include the Exponential, Weibull, Log-Logistic, Log-Normal and Gamma, among other distributions.
For interpretation purposes let us consider this class of models assuming an only dichotomous covariate \( X \) with two levels (e.g., labels of groups \( x = 1 \) vs. \( x = 0 \)), then the expected ratio of the survival times for the two levels is \( \exp(\theta) \). Therefore, AFT models can be interpreted in terms of the rate of progression of a disease: a ratio above (under) 1 means that the time is accelerated (slowed down) by a constant factor.

### 2.1. Adding Uncertainty in the Acceleration Factor

If the experimenter’s choice on the acceleration factor is only approximate, 

\[
\eta(x, \theta) \approx \exp(\theta x),
\]

and the “true” model is 

\[
\eta(x, \theta) = \exp \left( \theta x + \frac{\psi(x)}{\sqrt{n}} \right)
\]

for some unknown but ‘small’ function \( \psi \) belonging to certain class of ‘contaminating’ functions \( \Psi \), then the “true” log-linear model is 

\[
\log(T) = \mu + \theta' x + \frac{\psi(x)}{\sqrt{n}} + \sigma \epsilon = \mu + \theta' x + \sigma \cdot \left( \epsilon + \frac{\psi(x)}{\sigma \sqrt{n}} \right) = \mu + \theta' x + \sigma \tilde{\epsilon}, \tag{2}
\]

with \( \tilde{\epsilon} \sim \left( \mu_\tilde{\epsilon} = \frac{\psi(x)}{\sigma \sqrt{n}}, \sigma_\tilde{\epsilon} = 1 \right) \) and \( n \) denotes the sample size. The term \( \sqrt{n} \) is included joint with the contaminating function for an adequate asymptotic treatment of the bias as it will be seen later. Then, the PDF for the “true” model is 

\[
f(t|x; \mu, \theta) = \frac{1}{\sigma t} \cdot f_\tilde{\epsilon} \left( \frac{\log(t) - \mu - \theta' x - \frac{\psi(x)}{\sigma \sqrt{n}}}{\sigma} \right) \cdot f_t(z_t)
\]

since \( f_t(z_t) = f_t(z_t - \frac{\psi(x)}{\sigma \sqrt{n}}) \).

Under this framework, the ‘target’ parameter \( \beta_0 = (\mu_0, \theta_0) \)’ is defined as 

\[
\beta_0 = \arg \min_\beta \int_{\chi} \left( \E[\log(T)|x] - \log(T)|_\beta \right)^2 dx,
\]

where \( \beta = (\mu, \theta) \)' and \( \chi \) denotes the covariate design region. The above minimum condition requires to verify the following orthogonality condition: 

\[
\int_{\chi} h(x) \psi(x) dx = 0,
\]

where \( h(x) = (d\log(T)/d\beta|_{\beta=\beta_0})' \beta \).

Two types of contaminating functions will be considered in this work:

\[
\Psi_1 = \left\{ \psi : \max_x |\psi(x)| \leq \tau, \int_{\chi} h(x) \psi(x) dx = 0, \psi(0) = 0 \right\},
\]

and

\[
\Psi_2 = \left\{ \psi : \int_{\chi} \psi^2(x) dx \leq \eta^2 / n, \int_{\chi} h(x) \psi(x) dx = 0, \psi(0) = 0 \right\},
\]

where \( \tau \) and \( \eta \) are constants characterizing the neighborhood radio. The initial condition \( \psi(0) = 0 \) is needed for satisfying that \( \Delta F(0) = 1 \). Although the class \( \Psi_1 \) is frequently found in several works [28–30], it has the disadvantage of being too ‘thin’ in the sense of limited robustness against realistic departures from the assumed model. Conversely, Huber [25] proposes a wider class of plausible alternative responses \( \Psi_2 \), which requires to consider continuous designs. Even though some example is stated in this work considering the class \( \Psi_1 \), our interest is in \( \Psi_2 \) for being the most commonly used class in bias-robust literature.
2.2. A Minimax Strategy to Construct Robust Designs

If model (1) is fitted but the “true” one is (2), the least squares estimator of the model parameters \( \hat{\beta} = (\hat{\mu}, \hat{\theta}) \) or the predicted response \( \hat{Y}(x) \) is biased, so that the Mean Squared Error (MSE) matrix of the estimates rather than variance becomes a suitable measure of design quality.

An objective function must be defined from the MSE matrix seeking to optimize some aspect of the sampling planning. The classical alphabetic optimality criteria are extended in robust designs replacing the variance-covariance matrix of the estimators by the MSE matrix. These optimality criteria are named loss functions in the bias-robust approach context. Two loss functions were selected in this research for covering both estimation purposes: model parameters and response. Thus, the \( D \)- and \( I \)-losses are defined as

\[
\mathcal{L}_D(\psi, \xi) = \det \text{MSE}(\hat{\beta})^{1/p} = \frac{\sigma^2}{n} \left( \frac{1 + b'(\psi, \xi) [\frac{\sigma^2}{n} \text{M}(\xi)]^{-1} b(\psi, \xi)}{\det \text{M}(\xi)} \right)^{1/p}
\]

and

\[
\mathcal{L}_I(\psi, \xi) = \int_X \text{MSE}(\hat{Y}(x)) dx = \frac{\sigma^2}{n} \text{Tr}(AM^{-1}(\xi)) + b'(\psi, \xi)H(\xi)b(\psi, \xi) + \int_X \psi^2(x) dx
\]

respectively, where \( p \) is the number of model parameters,

\[
A = \int_X h(x)h(x)'dx, \\
M(\xi) = \int_X h(x)h(x)'d\xi(x), \\
H(\xi) = M(\xi)A^{-1}M(\xi), \\
b(\psi, \xi) = \int_X h(x)\psi(x)d\xi(x).
\]

Our attention in this work is focused on considering both errors due to bias and sampling in the design problem, so that a minimax approach will be adopted. Therefore, it is defined a robust design or optimal robust design as one that minimizes the maximum of some scalar-valuated function of the MSE matrix of the estimates

\[
\xi^* = \arg \min_{\xi} \max_{\psi} \mathcal{L}(\psi, \xi).
\]

3. Asymptotic Mean Squared Error Matrix for Type I Censoring and Known Variance

Following [9] let us consider that all the individuals in the study are recruited at time 0 and that there is a follow-up until time \( d \). The end of study is the only censoring considered here. We also assume \( \sigma \equiv \text{var}(\log T) \) to be known.

Let \( \delta \equiv 1_{\{T < d\}} \) be an indicator variable that is equal to 1 if the observation corresponds to an event and 0 if it corresponds to a right-censored observation. The log-likelihood function for model (1) and just one observation at value \( x \) is

\[
\text{LL}(\mu, \theta; t, x) = \delta \log(f(t|x; \mu, \theta)) + (1 - \delta) \log(S(t|x; \theta)) \\
= -\delta \log(\sigma t) + \delta \log(f_s(z_i)) + (1 - \delta) \log(S_s(z_i))
\]

where \( z_i = \frac{\log(t_i) - \mu - \theta'x_i}{\sigma} \).

Let \( \xi \) be an approximate design

\[
\xi = \left\{ \begin{array}{c} x_1 \omega_1 \\
\vdots \omega_2 \\
\vdots \\
x_m \omega_m \\
\end{array} \right\}, \quad 0 < \omega_i \leq 1, \quad \sum_{i=1}^m \omega_i = 1
\]
where \( x_1, x_2, \ldots, x_m \in \chi \) \((m \leq n)\) are the distinct experimental values for the covariates and \( \omega_i \) is the proportion of observations taken at \( x_i \).

In order to find the maximum likelihood estimator for the vector of model parameters assuming that the “true” model is (2), it is necessary to differentiate twice the log-likelihood, to compute its expectation and to analyze its limiting behavior. Thus the asymptotic information matrix of \( \hat{\beta}_0 \) is

\[
M(\xi) = M(\xi, \hat{\beta}_0) = \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} I^{(2)}(\beta_0; t, x_i) = \sum_{i=1}^{m} \omega_i I^{(2)}_{11}(\beta_0; t, x_i) \left( \begin{array}{c} 1 \\ x_i \\ x_i' \end{array} \right),
\]

with

\[
I^{(2)}_{11}(\beta_0; t, x) = E_{(2)} \left[ -\frac{\partial^2 LL(\mu, \theta; t, x)}{\partial \mu^2} \right] =
\]

\[
= \frac{1}{\sigma^2} \left( \frac{f_t'(z)}{S_t(z)} + \int_{-\infty}^{z_d} \frac{f_t'(z)^2}{f_t(z)} dz \right) + \\
\psi'(x) \frac{1}{\sigma^3 \sqrt{n}} \left[ \frac{f_t'(z)^2}{f_t(z)} - \frac{f_t'(z)^2}{f_t(z)} \right] (-\infty) - \int_{-\infty}^{z_d} \frac{f_t'(z)^3}{f_t(z)^2} dz + \\
- \psi''(x) \frac{1}{\sigma^3 \sqrt{n}} \left( \frac{f_t'(z)^2}{S_t(z)} \right) \left( \frac{f_t(z)}{S_t(z)} \right) \left( \frac{f_t'(z)}{f_t(z)} \right) + o \left( \frac{\psi(x)}{\sigma^2 \sqrt{n}} \right),
\]

where the two-term Taylor expansion of \( f_t(z) \) was used. \( f_t(z) \) must be of class \( C^2 \) to satisfy the conditions given in Lemma 2.1 of [9]. Details for obtaining the previous formulae are shown in the Appendix A.

On the other hand, the asymptotic expectation of the score function at \( \beta_0 \) is

\[
\hat{b}(\xi, \psi) = \hat{b}(\xi, \psi, \beta_0) = \frac{1}{\sqrt{n}} \lim_{n \to \infty} \frac{1}{n} E_{(2)} \left[ \sum_{i=1}^{n} \frac{\partial LL(\beta; t, x_i)}{\partial \beta} \right] =
\]

\[
= \frac{1}{\sigma^2 \sqrt{n}} \sum_{i=1}^{m} \omega_i \left( \frac{f_t'(z)^2}{S_t(z)} \right) \left( \frac{f_t(z)}{S_t(z)} \right) \left( \frac{f_t'(z)}{f_t(z)} \right) + o \left( \frac{\psi(x)}{\sigma^2 \sqrt{n}} \right).
\]

Again, further explanation of above calculations is given in the Appendix A. Furthermore, the asymptotic variance–covariance matrix of the score function evaluated at \( \beta_0 \) is

\[
C(\xi, \beta_0) = \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{m} \text{Cov} \left[ \frac{\partial LL(\beta; t, x_i)}{\partial \beta} \right] = \sum_{i=1}^{m} \omega_i I^{(2)}_{11}(\beta_0; t, x_i) \left( \begin{array}{c} 1 \\ x_i \\ x_i' \end{array} \right).
\]

It should be noted that the asymptotic matrices \( M(\xi) \) and \( C(\xi) \) are identical and that the maximum likelihood estimate \( \hat{\beta} \) is a root of the score function \( s(\beta) \). Therefore, using Taylor expansion of this function around \( \beta_0 \), the following is obtained

\[
s(\beta) = s(\beta_0) + s'(\beta_0)(\beta - \beta_0) + \ldots \Rightarrow 0 = s(\hat{\beta}) \approx s(\beta_0) + s'(\beta_0)(\hat{\beta} - \beta_0).
\]

Then \( (\hat{\beta} - \beta_0) \approx M^{-1}(\xi, \beta_0)s(\beta_0) \) and, since \( \sqrt{n} s(\beta_0) \sim AN(b(\xi, \psi), C(\xi) = M(\xi)) \), the asymptotic distribution of \( \sqrt{n} (\hat{\beta} - \beta_0) \) is \( AN(M^{-1}(\xi)b(\xi, \psi), M^{-1}(\xi)) \).
Finally, gathering previous results, it is obtained that the asymptotic MSE matrix is
\[
\text{MSE}(\xi, \psi) = (M^{-1}(\xi)b(\xi, \psi)(M^{-1}(\xi)b'(\xi, \psi)) + M^{-1}(\xi)
\]
\[
= M^{-1}(\xi)(b(\xi, \psi)b'(\xi, \psi) + M(\xi))M^{-1}(\xi).
\]

4. Log-Logistic Study Case

To illustrate previous results, let us assume that the variable \(T\) time-to-event to follow
a log-logistic distribution with parameter \(\sigma\) known. That is \(T|x \sim \text{Log} - \text{Log}\left(e^{\mu(x)}, \frac{1}{\sigma}\right)\)
and variable log \(\log(T|x) \sim \text{Log}[\mu, \sigma]\).

In the non-contaminated case \(T_u \sim \text{Log} - \text{Log}\left(e^{\mu_u}, \frac{1}{\sigma_u}\right)\)
and \(\text{log}(T_u) \sim \text{Log}[\mu_u, \sigma]\), so
that \(E[\text{log}(T_u)] = \mu_u\) and \(\text{Var}[\text{log}(T_u)] = \sigma^2\pi^2 = \sigma_u\). Then, the corresponding log-linear form is
\[
\text{log}(T|x) = \mu_u + \theta'x + \sigma_u\varepsilon' = \mu_u + \theta'x + \frac{\sigma\pi}{\sqrt{3}}\varepsilon' = \mu_u + \theta'x + \sigma\varepsilon
\]
where \(\varepsilon' \sim \text{Log}[0, \sqrt{3}/\pi]\) and \(\varepsilon = \pi\varepsilon'/\sqrt{3} \sim \text{Log}[0, 1]\). Consequently, \(\varepsilon\) follows a standard
logistic distribution with PDF
\[
f_{\varepsilon}(z) = \frac{e^z}{(1 + e^z)^2}, \quad -\infty < z < \infty
\]
and survival function
\[
S_z(z) = \frac{1}{1 + e^z}, \quad -\infty < z < \infty,
\]
where \(\sigma\) is again assumed known.

Finally, to obtain the MSE matrix (5) for this study case, it is necessary to compute
\[
M(\xi) = M(\xi, \beta_0) = \sum_{i=1}^{m} \omega_i \mathcal{T}_1^{(2)}(\beta_0, t, x_i)(1 \begin{pmatrix} x_i \end{pmatrix} x_i x_i^T)
\]
and
\[
b(\xi, \psi) = \sum_{i=1}^{m} \omega_i \psi(x_i)\mathcal{T}_1^{(2)}(\beta_0, t, x_i)(1 \begin{pmatrix} x_i \end{pmatrix} x_i x_i^T),
\]
where
\[
\mathcal{T}_1^{(2)} = \frac{1}{3\sigma^2} \left(1 - S(z_d)^3\right) + o\left(\frac{\psi(x)}{\sigma^4}\right)
\]
by virtue of results given in Section 3 for Functions (6) and (7). Details for obtaining \(\mathcal{T}_1^{(2)}\)
are given in Appendix B.

4.1. Robust Designs on \(\Psi_1\)

Let us consider a design region \(\chi = \{0, 1\}\) corresponding, for example, to two experi-
mental treatments, A and B. Since there are two parameters to be optimally estimated in
the model, both points 0 and 1 must be support points of the D-optimal robust design to be
obtained. The following theorem provides an analytical solution of the minimax problem
when the class of functions \(\Psi_1\) is assumed.

**Theorem 1.** Let \(\psi \in \Psi_1\). The D-optimal robust design on \(\chi = \{0, 1\}\) allocates a proportion of \(\omega^*\)
observations at point \(x = 0\), where
\[
\omega^* = \frac{A(\theta, \mu, \sigma, d) \cdot (1 + \tau^2) - B(\theta, \mu, \sigma) - \sqrt{(A^2(\theta, \mu, \sigma, d) \cdot (1 + \tau^2) - \tau^2B(\theta, \mu, \sigma))}}{
\tau^2 \cdot (A(\theta, \mu, \sigma, d) - B(\theta, \mu, \sigma))}
\]
with \( A(\theta, \mu, \sigma, d) = (e^{\theta + \mu} + d^{1/\nu})^3 \), \( B(\theta, \mu, \sigma) = e^{\theta + \mu} \) and \( \omega^* = 1/2 \) if \( \tau = 0 \).

A proof of the above result is given in the Appendix C. Figure 1 shows the weight variation on \( x = 0 \) as the neighborhood radio increases. From the graph it follows that the classical \( D \)-optimal design is obtained, which is equally weighted on all support points, when the contamination is almost negligible. Nevertheless, the higher the neighborhood radio is, the more emphasis is placed on point \( x = 0 \).

**Figure 1.** Minimax \( D \)- and \( I \)-optimal weights at point \( x = 0 \) for \( \psi \in \Psi_1 \).

In the case of the \( I \)-optimality, it was not possible to find an analytical expression of the optimal weights. The solutions of the function to be maximized are the roots of a nine-order polynomial on \( \omega \). For \( \theta, \mu, \sigma, d, \) and \( \tau \) fixed, the numerical calculation of \( \omega^* \) is straightforward. The same example as for \( D \)-optimization is also shown in the Figure 1 for \( I \)-optimality.

### 4.2. Robust Designs on \( \Psi_2 \)

The fact of assuming that \( T \) follows a log-logistic distribution implies that

\[
M(\xi) = \int_\chi \frac{1}{3\sigma^2}(1 - S_3(z_d))^3 \left( \frac{1}{x} \right) d\xi(x),
\]

where \( h(x) = \sqrt{\alpha(x)g(x)} \) being \( a(x) = \frac{1}{3\sigma^2}(1 - S_3(z_d))^3 \) and \( g(x) = (1 x)' \). Matrices intervening in (3) and (4) are now

\[
A = \int_\chi a(x)g(x)g'(x)dx,
\]

\[
M(\xi) = \int_\chi a(x)g(x)g'(x)d\xi(x),
\]

\[
H(\xi) = M(\xi)A^{-1}M(\xi),
\]

\[
b(\xi, \psi) = \int_\chi a(x)\psi(x)g'(x)d\xi(x).
\]

Let us define also

\[
K(\xi) = \int_\chi a^2(x)g(x)g'(x)d\xi(x),
\]

\[
G(\xi) = K(\xi) - M(\xi)A^{-1}M(\xi),
\]

which will be useful later. In order to avoid singularity problems in matrices \( A \) and \( M(\xi) \), let us assume that for each vector \( v \neq 0 \), the set \( \{ x \in \chi, v'g(x) = 0 \} \) has Lebesgue measure zero [22].

Note that

(a) \( G(\xi) \) defined as above is positive semidefinite since

\[
G = \int_\chi [(a(x)I - M(\xi)A^{-1})g(x)][(a(x)I - M(\xi)A^{-1})g(x)]'d\xi(x).
\]
(b) \( a(x) \) is a nonnegative function.

**Theorem 2.** The maximums of \( L_D(\xi, \psi) \) and \( L_I(\xi, \psi) \) in \( \Psi_2 \) are

\[
\sup_{\psi \in \Psi_2} L_D(\xi, \psi) = (1 - \nu) \left( 1 + \frac{\nu}{\det M(\xi)} \lambda_{\text{max}}[M^{-1}(\xi)G(\xi)] \right)^{1/p},
\]

\[
\sup_{\psi \in \Psi_2} L_I(\xi, \psi) = (1 - \nu) \text{Tr}(AM^{-1}(\xi)) + \nu \lambda_{\text{max}}[K(\xi)H^{-1}(\xi)],
\]

respectively, both attained at

\[\psi(x) = \eta \sqrt{n} g'(x) \{a(x)I - A^{-1}M(\xi)\} c_{\lambda_{\text{max}}}\]

where \( c_{\lambda_{\text{max}}} \) is the normalized eigenvector corresponding to the maximum eigenvalue, \( \lambda_{\text{max}} \), of the matrices \( M^{-1}(\xi)G(\xi) \) and \( K(\xi)H^{-1}(\xi) \), respectively. The constant

\[\nu = \frac{\eta^2}{\sigma^2 + \eta^2} \in [0, 1]\]

is defined representing the relative importance, to the experimenter, of errors due to bias rather than to variance.

An outline of the proof is given in Appendix D. Then, the initial minimax problem for both \( D \)- and \( I \)-optimality criteria has become one of finding a design that minimizes \((14)\) and \((15)\), respectively. To this end, and since practitioners require exact designs to implement their experiments, we propose the following strategy:

1. To discretize the design region \( \chi \).
2. To consider \( n \)-point exact designs \( \xi = \{x_1, \ldots, x_n\} \), \( x_i \in \chi \) not necessarily different and being \( \omega_i \)'s integer multiples of \( 1/n \).
3. To define the discrete version of matrices \((8)-(13)\) replacing \( \int \) by \( \sum \) properly.
4. To use an heuristic algorithm to numerically minimize \((14)\) and \((15)\) according to Theorem 2.

In order to illustrate the previous results, let us consider the real application proposed by Qi (2009) [26]. Here the author considered a randomized placebo-controlled trial to prevent tuberculosis in Ugandan adults infected with HIV. Log-logistic distribution exhibited the best fit in this study. Let \( x \in \chi = [0, 1] \) be the covariate (dose) to be designed and let \( d = 3, \sigma = 0.71, \mu = -2.97, \) and \( \theta = 0.1 \) be the selected nominal values as in [26]. An heuristic algorithm is required for the loss minimization according to step 4.

In this work, a Genetic Algorithm (GA) similar to that presented in García-Camacha et al. (2020) [31] will be used for this purpose.

Tables 1–4 collect the \( D \)- and \( I \)-optimal robust designs with 2 and 10 points, respectively. It is important to note that the \( D \)-optimal robust design achieved when the variance is almost ignored (\( \nu = 0.91 \)) is not unique. On the other hand, the achieved \( I \)-optimal robust design remains unchanged regardless the certainty that the experimenter has about model adequacy. A similar phenomenon is documented in the context of mixture designs [32] as both terms involved in \((4)\) are proportionals. This result is a very desirable design property for seeking robustness.

The classical alphabetic optimality criteria, based on variance alone, have the appealing property of being convex functionals of the design. This allows for the establishment of equivalence theorems, by which a design is optimal if it satisfies first-order conditions. When bias is incorporated into the loss, this convexity is lost and so the first-order conditions, necessary for optimality, cannot be proved to be sufficient [33]. Therefore, the only way to validate the obtained designs in this work is to compare with those obtained in...
the literature (if available). In this regard, the classical 2-point $D$-optimal design obtained in this work is the same as that obtained in [9] (case $\nu = 0$). Tables 5 and 6 show the computational cost, measured in terms of number of the algorithm iterations, and the obtained loss function values for $D$- and $I$-optimal robust designs.

In order to show the benefits of the obtained robust designs ($\nu > 0$) versus designs calculated using the classical optimal design theory ($\nu = 0$), efficiencies were calculated as $\frac{\text{loss}_v(\xi^*_D)}{\text{loss}_v(\xi^*_{\nu=0})} \times 100$. These efficiencies range from 0 to 1, so that values close to 0 indicate that classical design is not robust at all, while a value of 1 means the opposite case. Note that for study cases in which the robust design coincides with the classical one, calculating the efficiency is meaningless (which occurs for the $I$-optimal robust designs and for $\nu \leq 0.7$ in the case of $D$-optimality and two-point designs). When the knowledge of model is almost non-existent or many deviations of it occur during experiments ($\nu = 0.91$), two-point classical $D$-optimal design is 3.42% efficient compared with the proposed ones in Table 1. Efficiencies of ten-point classical $D$-optimal design compared with the obtained robust designs are shown in Table 5.

| Table 1. 2-point $D$-optimal robust designs computed for the log-logistic model. |
| --- |
| $v$ | $D$-Optimal Robust Designs |
| 0.01 | $\xi^*_D = \{ 0 \ 0.5 \ 1 \}$ |
| 0.3 | $\xi^*_D = \{ 0.34 \ 0.99 \}$ |
| 0.91 | $\xi^*_D = \{ 0.01 \ 0.66 \}$ |

| Table 2. 10-point $D$-optimal robust designs computed for the log-logistic model. |
| --- |
| $v$ | $D$-Optimal Robust Designs |
| 0.01 | $\xi^*_D = \{ 0 \ 0.01 \ 0.02 \ 0.03 \ 0.04 \ 0.96 \ 0.97 \ 0.98 \ 0.99 \ 1 \}$ |
| 0.3 | $\xi^*_D = \{ 0 \ 0.08 \ 0.54 \ 0.60 \ 0.73 \ 0.76 \ 0.86 \ 0.93 \ 0.96 \ 1 \}$ |
| 0.7 | $\xi^*_D = \{ 0.39 \ 0.43 \ 0.60 \ 0.62 \ 0.67 \ 0.70 \ 0.79 \ 0.83 \ 0.94 \ 0.96 \}$ |

| Table 3. 2-point $I$-optimal robust designs computed for the log-logistic model. |
| --- |
| $v$ | $I$-Optimal Robust Designs |
| 0.01 | $\xi^*_I = \{ 0 \ 0.5 \ 1 \}$ |
| 0.3 | $\xi^*_I = \{ 0.1 \ 0.1 \ 0.1 \ 0.1 \ 0.1 \ 0.1 \ 0.1 \ 0.1 \ 0.1 \}$ |
| 0.7 | $\xi^*_I = \{ 0.01 \ 0.03 \ 0.04 \ 0.96 \ 0.97 \ 0.98 \ 0.99 \ 1 \}$ |

| Table 4. 10-point $I$-optimal robust designs computed for the log-logistic model. |
| --- |
| $v$ | $I$-Optimal Robust Designs |
| 0.01 | $\xi^*_I = \{ 0 \ 0.01 \ 0.02 \ 0.03 \ 0.04 \ 0.96 \ 0.97 \ 0.98 \ 0.99 \ 1 \}$ |
| 0.3 | $\xi^*_I = \{ 0.1 \ 0.1 \ 0.1 \ 0.1 \ 0.1 \ 0.1 \ 0.1 \ 0.1 \ 0.1 \}$ |
| 0.7 | $\xi^*_I = \{ 0.01 \ 0.03 \ 0.04 \ 0.96 \ 0.97 \ 0.98 \ 0.99 \ 1 \}$ |
Table 5. Number of iterations needed for achieving the $D$-optimal robust designs computed in Tables 1 and 2, their corresponding loss function values and efficiencies.

| $\nu$ | No. It (2p) | $D$-Loss (2p) | Efficiency | No. It (10p) | $D$-Loss (10p) | Efficiency |
|-------|-------------|---------------|------------|-------------|---------------|------------|
| 0.01  | 1037        | 2.2989        | -          | 1744        | 3.04          | -          |
| 0.3   | 1008        | 1.961         | -          | 1239        | 1.903         | 99.05      |
| 0.5   | 1436        | 1.237         | -          | 1270        | 1.002         | 96.37      |
| 0.7   | 1080        | 0.4338        | -          | 1125        | 0.001823      | 0.31       |
| 0.91  | 1061        | 0.01426       | 3.42       | 1617        | 0.00103       | 0.22       |

Table 6. Number of iterations needed for achieving the $I$-optimal robust designs computed in Tables 3 and 4, and their corresponding loss function values.

| $\nu$ | No. It (2p) | $I$-Loss (2p) | No. It (10p) | $I$-Loss (10p) |
|-------|-------------|---------------|-------------|---------------|
| 0.01  | 1063        | 2.011         | 1916        | 2.05          |
| 0.3   | 1120        | 1.569         | 1543        | 1.479         |
| 0.5   | 1015        | 1.263         | 1184        | 1.085         |
| 0.7   | 1103        | 0.9579        | 1284        | 0.6909        |
| 0.91  | 1006        | 0.6374        | 1328        | 0.2773        |

5. Discussion

Accelerated failure time models have been widely used in industrial processes, but nowadays they are becoming popular in many other fields, including medical research. This fact is not only due to their relaxed conditions compared with proportional hazard models, but also due to the easy interpretation of the effects for applicability purposes [9]. The main feature characterizing this type of models is that the covariates have a multiplicative effect which accelerates/decelerates the survival. Nevertheless, practitioners rarely know the exact form of the acceleration factor [24], which is something necessary to obtain optimal designs, strongly dependent of the model.

In this work we provide robust designs against departures from the assumed acceleration factor. To this end, a space of contaminating functions was considered, and the mean square error matrix was established as the basis for measuring the design quality. A key result of this research was obtaining the asymptotic expression of this matrix for type I censoring and known variance. This expression is general in the sense that it does not depend on the selected failure time distribution.

Several study cases were analysed considering the log-logistic distribution to illustrate the previous result. First example comprises a bounded contaminating space and a placebo-treatment type design region. Theoretical support is provided for constructing the $D$-optimal robust design, whereas the $I$-optimal problem can be only solved numerically. Two types of contaminating functions are considered in this work. The maximization problem over these classes of functions is theoretically solved and a construction strategy similar to [31] for computing the $D$- and $I$-robust designs is used when the contaminating functions are a $L_2$-type neighbourhood.

Methodology used for calculating the obtained designs was confirmed when there was available information in the literature. Note that there is a lack of equivalence theorems since convexity might be violated under bias-robust approach [33]. $D$-optimal robust designs tend to be equally weighted on the extremes when the model form is either known or almost known, whereas the allocation is more random when bias emphasizes. This latter is a quite common result in robust design [34]. It seems logical to think that an adequate strategy to seek protection against the misspecification total or quasi total of the model is not favoring any design region. On the other hand, the achieved $I$-optimal robust design remains unchanged regardless the certainty that the experimenter has about model adequacy. This is a very desirable design property for seeking robustness.

As we mentioned before, peculiarities of real experiments lead to different censoring. Some authors warn of the risks of addressing data as right-censored when interval censoring...
occurs [35,36]. This may lead to biased estimates or underestimates of the “true” error variance or the survival itself. So censoring mechanisms can be quite complicated and thus special methods are necessary for their appropriate treatment. Further research must be aimed at this important issue. New alternatives for interval censored data under this bias-robust approach proposed in this work are being explored.

Author Contributions: Conceptualization, M.J.R.-L. and R.M.-M.; methodology, M.J.R.-L., R.M.-M. and I.G.-C.G.; software, I.G.-C.G., R.M.-M. and M.J.R.-L.; writing—original draft preparation, I.G.-C.G.; writing—review and editing, M.J.R.-L., R.M.-M. and I.G.-C.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Ministerio de Economía, Industria y Competitividad—Agencia Estatal de Investigación FEDER (MTM2016-80539-C2-1-R and MTM2016-80539-C2-2-R), Ministerio de Ciencia e Innovación—Agencia Estatal de Investigación (PID2020-113443RB-C21), Junta de Comunidades de Castilla-La Mancha through Fondo Europeo de Desarrollo Regional (SBPLY/17/180501/000380) and Junta de Castilla y León (SA080P17 and SA105P20).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Abbreviations
The following abbreviations are used in this manuscript:

AFT Accelerated Failure Time
AF Acceleration Factor
OED Optimal Experimental Design
MSE Mean Square Error
PH Proportional Hazard
PDF Probability Density Function
FIM Fisher Information Matrix

Appendix A
Fisher’s information matrix for a single observation is defined as

\[ M(\xi) = \int_{\chi} \mathcal{I}^{(2)}(\mu, \theta; t, x) d\xi(x), \]

where

\[ \mathcal{I}^{(2)}(\mu, \theta; t, x) = E_{(2)} \left[ -\frac{\partial^2 LL(\mu, \theta; t, x)}{\partial \beta \partial \beta'} \right] = \mathcal{I}_{11}^{(2)}(\mu, \theta; t, x) \left( \begin{array}{c} 1 \\ \mu_x x \end{array} \right), \]

by virtue of results obtained in [9]. Using two-term Taylor expansion,

\[ f(\tilde{z}_i) = f(z_i - \frac{\psi(x)}{\sigma\sqrt{n}}) \approx f(z_i) - \frac{\psi(x)}{\sigma\sqrt{n}} f'(z_i) + o\left( \frac{\psi(x)}{\sigma\sqrt{n}} \right), \]

the following expression can be obtained

\[ E_{(2)}[H(z)] = \int_{-\infty}^{\infty} H(z) f(z) dz = \int_{-\infty}^{\infty} H(z) f(z) \left( z - \frac{\psi(x)}{\sigma\sqrt{n}} \right) dz \]

\[ = E_{(1)}[H(z)] - \frac{\psi(x)}{\sigma\sqrt{n}} \int_{-\infty}^{\infty} H(z) f'(z) dz + o\left( \frac{\psi(x)}{\sigma\sqrt{n}} \right), \]
where the subscript (1) corresponds to non-contaminated case computed in [9]. Thus,

\[ I^{(2)}_{11} = E_{(2)} \left[ -\frac{\partial^2 LL(\mu, \theta; t, x)}{\partial \mu^2} \right] \]

\[ = E_{(1)} \left[ -\frac{\partial^2 LL(\mu, \theta; t, x)}{\partial \mu^2} \right] + \frac{\psi(x)}{\sigma \sqrt{n}} \int_{-\infty}^{\infty} \frac{\partial^2 LL(\mu, \theta; t, x)}{\partial \mu^2} f'_\epsilon(z) dz + o \left( \frac{\psi(x)}{\sigma \sqrt{n}} \right) \]

\[ = I^{(1)}_{11} + \frac{\psi(x)}{\sigma^3 \sqrt{n}} \left[ \int_{-\infty}^{\epsilon^2} \frac{f'_\epsilon(z)^2}{f_\epsilon(z)} dz \right] - \frac{\psi(x)}{\sigma^3 \sqrt{n}} \left[ \left[ f'_\epsilon(z) \right]_{(-\infty)} f_\epsilon(z) \right] \]

\[ \quad - \frac{\psi(x)}{\sigma^3 \sqrt{n}} \left( \frac{f'_\epsilon(z) S_\epsilon(z) + f_\epsilon(z) S_\epsilon(z)^2}{S_\epsilon(z)} \right) \left( [f_\epsilon(z)]_{(\infty)} - f_\epsilon(z) \right) \]

with

\[ I^{(1)}_{11} = E_{(1)} \left[ -\frac{\partial^2 LL(\mu, \theta; t, x)}{\partial \mu^2} \right] = \frac{1}{\sigma^2} \left( \left[ f'_\epsilon(z) \right]_{(-\infty)} + \frac{f_\epsilon(z) S_\epsilon(z)}{S_\epsilon(z)} + \int_{-\infty}^{z^2} f'_\epsilon(z)^2 f_\epsilon(z) dz \right) \]

as computed in [9]. On the other hand,

\[ E_{(2)} \left[ \frac{\partial LL(\mu, \theta; t, x)}{\partial \mu} \right] \]

\[ = E_{(1)} \left[ \frac{\partial LL(\mu, \theta; t, x)}{\partial \mu} \right] - \frac{\psi(x)}{\sigma \sqrt{n}} \int_{-\infty}^{\infty} \frac{\partial LL(\mu, \theta; t, x)}{\partial \mu} f'_\epsilon(z) dz + o \left( \frac{\psi(x)}{\sigma \sqrt{n}} \right) \]

\[ = \frac{\psi(x)}{\sigma^2 \sqrt{n}} \left( \int_{-\infty}^{z^2} \frac{f'_\epsilon(z)^2}{f_\epsilon(z)} dz - \frac{f_\epsilon(z) S_\epsilon(z)}{S_\epsilon(z)} \left( [f_\epsilon(z)]_{(\infty)} - f_\epsilon(z) \right) \right) + o \left( \frac{\psi(x)}{\sigma \sqrt{n}} \right) \]

and

\[ \text{Var}_{(2)} \left[ \frac{\partial LL(\mu, \theta; t, x)}{\partial \mu} \right] = E_{(2)} \left[ \left( \frac{\partial LL(\mu, \theta; t, x)}{\partial \mu} \right)^2 \right] - \left( E_{(2)} \left[ \frac{\partial LL(\mu, \theta; t, x)}{\partial \mu} \right] \right)^2 \]

\[ = E_{(1)} \left( \left( \frac{\partial LL(\mu, \theta; t, x)}{\partial \mu} \right)^2 \right) - \frac{\psi(x)}{\sigma \sqrt{n}} \int_{-\infty}^{\infty} \left( \frac{\partial LL(\mu, \theta; t, x)}{\partial \mu} \right)^2 f'_\epsilon(z) dz + o \left( \frac{\psi(x)}{\sigma \sqrt{n}} \right) \]

\[ = \frac{1}{\sigma^2} \left( \int_{-\infty}^{z^2} f'_\epsilon(z)^2 f_\epsilon(z) dz + f_\epsilon(z) S_\epsilon(z)^2 \right) - \frac{\psi(x)}{\sigma^3 \sqrt{n}} \left[ \int_{-\infty}^{z^2} f'_\epsilon(z)^3 f_\epsilon(z)^2 dz + f_\epsilon(z) S_\epsilon(z)^2 \left( [f_\epsilon(z)]_{(\infty)} - f_\epsilon(z) \right) \right] \]

\[ - \frac{\psi(x)}{\sigma^3 \sqrt{n}} \left( \int_{-\infty}^{z^2} f'_\epsilon(z)^2 f_\epsilon(z)^2 dz + f_\epsilon(z) S_\epsilon(z)^2 \left( [f_\epsilon(z)]_{(\infty)} - f_\epsilon(z) \right) \right) + o \left( \frac{\psi(x)}{\sigma \sqrt{n}} \right), \]

where the reader interested in obtaining the terms with subscript (1) is refereed to [9].

Taking into account above calculations and

\[ E_{(2)} \left[ \frac{\partial LL(\mu, \theta; t, x)}{\partial \beta} \right] = E_{(2)} \left[ \frac{\partial LL(\mu, \theta; t, x)}{\partial \mu} \right] \left( \frac{1}{x} \right), \]

obtaining the expectation of the score function is direct.
Appendix B

The FIM for model (1) and a single point \( x \in \chi \) is

\[
I(\mu, \theta; t, x) = \frac{1}{3\sigma^2} \left( 1 - S_\ell(z_d)^3 \right) \left( \begin{array}{c} 1 \\ x \\ xx' \end{array} \right), x' \in \chi.
\]

The FIM for model (2) and a single point \( x \in \chi \) is

\[
I^{(2)}(\mu, \theta; t, x) = \left[ \frac{1}{3\sigma^2} \left( 1 - S_\ell(z_d)^3 \right) + o \left( \frac{g(x)}{\sigma \sqrt{n}} \right) \right] \left( \begin{array}{c} 1 \\ x \\ xx' \end{array} \right), x' \in \chi.
\]

because the component 11 of the FIM for model (2) is

\[
I^{(2)}_{11} = I_{11} + \frac{g(x)}{\sigma^3 \sqrt{n}} \left[ \frac{f'_0(z_d)^2}{f_0(z_d)} - \left[ \frac{f'_0(z)^2}{f_0(z)} \right]_{(-\infty)} - \int_{-\infty}^{\infty} f'_0(z)^2 dz \right] - \frac{g(x)}{\sigma^3 \sqrt{n}} \left[ \frac{-e^{2z_d}}{(e^{z_d} + 1)^4} \right] + o \left( \frac{g(x)}{\sigma \sqrt{n}} \right)
\]

\[
= \frac{1}{3\sigma^2} \left( 1 - S_\ell(z_d)^3 \right) + \frac{g(x)}{\sigma^3 \sqrt{n}} \left[ \frac{e^{2z_d}}{(e^{z_d} + 1)^3} + \frac{e^{2z_d} (e^{2z_d} + 3)}{3(e^{z_d} + 1)^3} \right] + o \left( \frac{g(x)}{\sigma \sqrt{n}} \right)
\]

On the other hand:

\[
E(2) \left[ \frac{\partial LL(\mu, \theta; t, x)}{\partial \mu} \right]
\]

\[
= \frac{g(x)}{\sigma^2 \sqrt{n}} \left[ \frac{e^{2z_d}}{(e^{z_d} + 1)^3} + \frac{e^{2z_d} (e^{2z_d} + 3)}{3(e^{z_d} + 1)^3} \right] + o \left( \frac{g(x)}{\sigma \sqrt{n}} \right)
\]

\[
= \frac{g(x)}{\sigma^2 \sqrt{n}} \left[ \frac{1}{3} - \frac{1}{3(e^{z_d} + 1)^3} \right] + o \left( \frac{g(x)}{\sigma \sqrt{n}} \right)
\]

\[
= \frac{g(x)}{3\sigma^2 \sqrt{n}} \left( 1 - S_\ell(z_d)^3 \right) + o \left( \frac{g(x)}{\sigma \sqrt{n}} \right) = \frac{g(x)}{\sqrt{n}} I_{11} + o \left( \frac{g(x)}{\sigma \sqrt{n}} \right)
\]

Appendix C

Proof. For \( \xi = \begin{pmatrix} 0 \\ \omega \end{pmatrix}, \) the determinant of the MSE matrix is

\[
\det \text{MSE}(\xi, \psi) = \frac{81\sigma^8}{(1 - S_\ell^2(z_{0,d}))(1 - S_\ell^2(z_{1,d}))} \frac{1 + (1 - \omega)(1 - S_\ell^2(z_{1,d})) \cdot \psi^2(1)}{\omega - \omega^2}
\]
where \( z_{x,d} = \frac{\log(d) - \mu + \delta x}{\sigma} \) and \( \det \text{MSE}(\hat{\beta}) = \det \text{MSE}(\xi, \psi) \) abusing notation. Since the maximization is performed over \( \Psi_1 \) and taking into account that
\[
\det \text{MSE}(\xi, \psi) \propto \frac{1 + (1 - \omega)(1 - S_2^3(z_{1,d})) \cdot \psi^2(1)}{\omega - \omega^2} \leq \frac{1 + (1 - \omega)(1 - S_2^3(z_{1,d})) \cdot \tau^2}{\omega - \omega^2},
\]
from first-order derivative results
\[
(1 - S_2^3(z_{1,d})) \cdot \tau^2 - \psi^2(1) - \omega \cdot ((1 - S_2^3(z_{1,d}) \cdot \tau^2)) + (1 + ((1 - S_2^3(z_{1,d}) \cdot \tau^2)) = 0
\]
\[
\Leftrightarrow \omega_{1,2} = \frac{(1 + ((1 - S_2^3(z_{1,d}) \cdot \tau^2)) \pm \sqrt{1 + (1 - S_2^3(z_{1,d}) \cdot \tau^2)}}{(1 - S_2^3(z_{1,d}) \cdot \tau^2)}.
\]

The positive root of the numerator leads to weights greater than one. Therefore, the positive root is rejected. The final result is direct after some algebra. \( \Box \)

Appendix D

Proof. Define
\[
r(x) = \frac{\eta}{\sqrt{n}} G^{1/2}(\xi) \{ a(x) I - A^{-1} M(\xi) \} g(x)
\]
so that
\[
\int_X r(x) r'(x) d\xi(x) = \frac{\eta^2}{n} I,
\]
\[
\int_X g(x) g'(x) d\xi(x) = 0,
\]
\[
\int_X \psi(x) r(x) d\xi(x) = \frac{\eta^2}{n} G^{-1/2}(\xi) [b(\xi, \psi),
\]
\[
\int_X g(x) r'(x) d\xi(x) = \frac{\eta^2}{n} G^{-1/2}(\xi),
\]
are satisfied. The rest of the proof is obtained in a manner very similar to that used in Theorem 1 of Wiens (1992) \[19]\. \( \Box \)

References

1. Collett, D. *Modelling Survival Data in Medical Research*; CRC Press: Boca Raton, FL, USA, 2015.
2. Wei, L.J. The accelerated failure time model: A useful alternative to the cox regression model in survival analysis. *Stat. Med.* 1992, 11, 1871–1879. [CrossRef] [PubMed]
3. Escobar, L.A.; Meeker, W.Q. A Review of Accelerated Test Models. *Stat. Sci.* 2006, 21, 552–577. [CrossRef]
4. Gökkuş, Z.; Alakuş, K.; Bilien, S. USE of group sequential test methods in the survival analysis in aquaculture. *Aquaculture* 2020, 515, 734568. [CrossRef]
5. Bai, D.; Chung, S. An optimal design of accelerated life test for exponential distribution. *Reliab. Eng. Syst. Saf.* 1991, 31, 57–64. [CrossRef]
6. Ahmad, N.; Islam, A.; Salam, A. Analysis of optimal accelerated life test plans for periodic inspection: The case of exponentiated Weibull failure model. *Int. J. Qual. Reliab. Manag.* 2006, 23, 1019–1046. [CrossRef]
7. Das, R.N.; Lin, D.K. On D-optimal robust designs for lifetime improvement experiments. *J. Stat. Plan. Inference* 2011, 141, 3753–3759. [CrossRef]
8. Müller, C.H. D-Optimal Designs for Lifetime Experiments with Exponential Distribution and Censoring. In mODa 10—Advances in Model-Oriented Design and Analysis; Uciniski, D., Atkinson, A.C., Patan, M., Eds.; Springer International Publishing: Heidelberg, Germany, 2013; pp. 179–186. [CrossRef]
9. Rivas-López, M.J.; López-Fidalgo, J.; Campo, R.d. Optimal experimental designs for accelerated failure time with Type I and random censoring. *Biom. J.* 2014, 56, 819–837. [CrossRef]
10. Rivas-López, M.; Yu, R.; López-Fidalgo, J.; Ruiz, G. Optimal experimental design on the loading frequency for a probabilistic fatigue model for plain and fibre-reinforced concrete. *Comput. Stat. Data Anal.* 2017, 113, 363–374. [CrossRef]
11. Mariñas Collado, I.; Rivas-López, M.J.; Rodríguez-Díaz, J.M.; Santos-Martín, M.T. A New Compromise Design Plan for Accelerated Failure Time Models with Temperature as an Acceleration Factor. *Mathematics* 2021, 9, 836. [CrossRef]
12. Liang, T.; Chen, L.S.; Yang, M.C. Efficient Bayesian sampling plans for exponential distributions with random censoring. J. Stat. Plan. Inference 2012, 142, 537–551. [CrossRef]  
13. Park, S.; Ng, H.K.T. Missing information and an optimal one-step plan in a Type II progressive censoring scheme. Stat. Probab. Lett. 2012, 82, 396–402. [CrossRef]  
14. Ryeznik, Y.; Sverdlov, O.; Hooker, A.C. Adaptive optimal designs for dose-finding studies with time-to-event outcomes. AAPS J. 2018, 20, 1–12. [CrossRef] [PubMed]  
15. Pascual, F.G. Accelerated Life Test Plans Robust to Misspecification of the Stress—Life Relationship. Technometrics 2006, 48, 11–25. [CrossRef]  
16. Ginebra, J.; Sen, A. Minimax approach to accelerated life tests. IEEE Trans. Reliab. 1998, 47, 261–267. doi: 10.1109/24.740498. [CrossRef]  
17. Pascual, F.; Montepiedra, G. On minimax designs when there are two candidate models. J. Stat. Comput. Simul. 2002, 72, 841–862. [CrossRef]  
18. Box, G.E.P.; Draper, N.R. A Basis for the Selection of a Response Surface Design. J. Am. Stat. Assoc. 1959, 54, 622–654. [CrossRef]  
19. Wiens, D.P. Minimax designs for approximately linear regression. J. Stat. Plan. Inference 1992, 31, 353–371. [CrossRef]  
20. Wiens, D.P. Designs for approximately linear regression: Maximizing the minimum coverage probability of confidence ellipsoids. Can. J. Stat. 1993, 21, 59–70. [CrossRef]  
21. Wiens, D.P. Robust designs for approximately linear regression: M-estimated parameters. J. Stat. Plan. Inference 1994, 40, 135–160. [CrossRef]  
22. Xu, X. Robust prediction and extrapolation designs for censored data. J. Stat. Plan. Inference 2009, 139, 486–502. [CrossRef]  
23. Konstantinou, M.; Biedermann, S.; Kimber, A. Model robust designs for survival trials. Comput. Stat. Data Anal. 2017, 113, 239–250. [CrossRef]  
24. Nelson, W.B. Accelerated Testing: Statistical Models, Test Plans, and Data Analysis; John Wiley & Sons: Hoboken, NJ, USA, 2009; Volume 344.  
25. Huber, P.J. Robustness and designs. In A Survey of Statistical Design and Linear Models; North Holland: Amsterdam, The Netherlands, 1975; pp. 287–303.  
26. Qi, J. Comparison of Proportional Hazards and Accelerated Failure Time Models. Ph.D. Thesis, Department of Mathematics and Statistics, University of Saskatchewan, Saskatoon, SK, Canada, 2009.  
27. Garcia-Camacha Gutiérrez, I.; Martín Martín, R.; Sanz Argent, J. Optimal-robust selection of a fuel surrogate for homogeneous charge compression ignition modeling. PLoS ONE 2020, 15, e0234963. [CrossRef] [PubMed]  
28. Garcia-Camacha Gutiérrez, I. Diseño Óptimo de Experimentos para Modelos de Mezclas Aplicados en la Ingeniería y las Ciencias Experimentales. Ph.D. Thesis, University of Castilla-La Mancha, Ciudad Real, Spain, 2017. Available online: http://hdl.handle.net/10578/16595 (accessed on 28 September 2021).  
29. Pesotchinsky, L. Optimal robust designs: Linear regression in $\mathbb{R}^k$. Ann. Stat. 1982, 10, 511–525. [CrossRef]  
30. Li, K.C. Robust regression designs when the design space consists of finitely many points. Ann. Stat. 1984, 12, 269–282. [CrossRef]  
31. García-Camacha Gutiérrez, I.; Martín Martín, R.; Sanz Argent, J. Optimal-robust selection of a fuel surrogate for homogeneous charge compression ignition modeling. PLoS ONE 2020, 15, e0234963. [CrossRef] [PubMed]  
32. Wiens, D.P. I-robust and D-robust designs on a finite design space. Stat. Comput. 2018, 28, 241–258. [CrossRef]  
33. Kong, L.; Wiens, D.P. Model-Robust Designs for Quantile Regression. J. Am. Stat. Assoc. 2015, 110, 233–245. [CrossRef]  
34. Rücker, G.; Messerer, D. Remission duration: An example of interval-censored observations. Stat. Med. 1988, 7, 1139–1145. [CrossRef]  
35. Law, C.; Brookmeyer, R. Effects of mid-point imputation on the analysis of doubly censored data. Stat. Med. 1992, 11, 1569–1578. [CrossRef]