D-optimal designs for lifetime experiments with exponential distribution and censoring

Christine H. Müller

Abstract The approach of Kiefer-Wolfowitz is used to construct D-optimal designs for lifetime experiments with exponential distribution and censoring. If the expected life time is simply reciprocal to the stress then the optimal design does not depend on the unknown parameter and the censoring. However, the situation is more complicated for the more often used assumption that the logarithm of the expected lifetime is linear in the stress. Conditions are given here where the locally D-optimal designs for experiments with censoring coincide with those in the classical approach of normally distributed errors. In particular this is the case when the censoring variable is not too small and the slope of the regression is not too large.

1 Introduction

Often the expected lifetime $E(T(s))$ of a product depends on the stress s via a given link function. Here it is assumed that this function is known up to a parametervector θ. We assume that N lifetime experiments at different stress levels $s_n \in \mathscr{S}$ for $n = 1,\ldots,N$ are executed and that the lifetime T_n of the product shall be observed in each lifetime experiment. However, if the stress is too low then often the lifetime cannot be observed since the time up to the event, the "death", is too long. Therefore usually a time c is fixed at which the lifetime experiment is stopped. Then the only information is that the product has survived the time c. Such observations are socalled censored observations. It is clear that the censored observations should also be used in an analysis of lifetime data. Therefore define

$$Y_n := \begin{cases} T_n, & \text{if } T_n \le c, \\ c, & \text{if } T_n > c, \end{cases} \quad \text{and} \quad D_n := \begin{cases} 1, & \text{if } T_n \le c, \\ 0, & \text{if } T_n > c. \end{cases}$$

Christine H. Müller
Faculty of Statistics, TU University Dortmund, Vogelpothsweg 87, D-44221 Dortmund, Germany,
e-mail: cmueller@statistik.tu-dortmund.de

Then $(Y_1, D_1, s_1), \ldots, (Y_N, D_N, s_N)$ are the available informations where D_n is the censoring variable. Let be t_n, y_n, and d_n the realizations of T_n, Y_n and D_n respectively and $y_* = (y_1, \ldots, y_N)^\top$, $d_* = (d_1, \ldots, d_N)^\top$, $s_* = (s_1, \ldots, s_N)^\top$. The likelihood function is then given by (see e.g. [9] p.75)

$$L_\theta(y_*, d_*, s_*) := \prod_{n=1}^{N} f_{\theta,s_n}(y_n)^{d_n} S_{\theta,s_n}(y_n)^{1-d_n}$$

if $f_{\theta,s}$ is the lifetime distribution density at stress s and

$$S_{\theta,s}(t) := \int_t^\infty f_{\theta,s}(u)\, du$$

the survival function at time t and stress s.

There is a vast literature on planning and analysis of lifetime experiments due to their importance in practice. Often the planning concerns only the construction of sampling plans like in [10] or the censoring mechanism by removing units after some failure events like in [12], [14], or [15]. In other cases, like in [3], the planning concerns the analysis of a specific lifetime experiment. Only few papers deal with the optimal planning of the stress variables. Bai and Chung [2] consider the construction of optimal two-point designs for experiments where the number of failures up to a given time point is observed. They use the Poisson distribution to model the number of failures. Ahmad et al. [1] determine numerically locally optimal designs $\alpha e_{s_1} + (1-\alpha)e_{s_2}$ with $s_1 \in (s_0, s_2)$, where s_0, s_2 are given, if the lifetime is modeled by the exponentiated Weibull distribution. However, they assume that only the scale parameter of the exponentiated Weibull distribution is unknown and that the lifetime is measured in k units. Bai and Chung as well as Ahmad et al. do not use the equivalence theorems of Kiefer and Wolfowitz [8] and Fedorov [5] for constructing D- and A-optimal designs, respectively. D-optimal designs for lifetime experiments are developed by Das and Lin [4]. But they assume a lognormal distribution for the lifetime and taking the logarithm of the lifetime they can use the usual theory for normally distributed observations. Their new contribution is the assumption of correlated errors. As far as the author knows, there are only two papers using the equivalence theorems of Kiefer and Wolfowitz [8] and Fedorov [5] for constructing D- and A-optimal designs for censored observations or lifetime experiments. López-Fidalgo and Garcet-Rodríguez [6] derive optimal designs when the independent variable is censored. But in lifetime experiments, as considered in this paper, the dependent variable is censored. Pal and Mandal [11] construct optimal designs for the stress strength reliability $P(X < Y|Z)$ where both, the stress X and the strength Y, have an exponential distribution.

In this paper, we consider also the exponential distribution which is the simplest lifetime distribution so that

$$f_{\theta,s}(t) = \lambda_\theta(s) \exp(-\lambda_\theta(s)t)$$

is the lifetime density. Here the link function $\lambda_\theta : \mathscr{S} \to (0,\infty)$ is known up to the parameter vector θ. The expected lifetime is then

$$E_\theta(T_n) = \frac{1}{\lambda_\theta(s_n)}.$$

Simple reasonable functions for λ_θ are the following:

$$\lambda_\theta(s) = \theta s, \;\; \theta \in (0,\infty), \tag{1}$$
$$\lambda_\theta(s) = \exp(\theta_0 + \theta_1 s), \;\; \theta = (\theta_0,\theta_1)^\top \in \Re \times (0,\infty). \tag{2}$$

Both functions ensure that the expected life time is decreasing with increasing stress s. The function given by (1) provides an infinite life time if there is no stress while function (2) is more flexible with a finite expected life time for no stress. Function (2), also used in [1], means that the logarithm of the expected life time is linear which is an assumption often used by engineers (see e.g. [7], p.25).

We consider here the problem of constructing optimal designs of the stress levels for the maximum likelihood estimator of θ

$$\widehat{\theta} := \arg\max_\theta L_\theta(y_*,d_*,s_*).$$

In Section 2, the information matrix for this estimator is given. Section 3 deals then with the optimal design for λ_θ given by (1) and Section 4 provides locally D-optimal designs for λ_θ given by (2).

2 The information matrix

Since the survival function for the exponential distribution satisfies $S_{\theta,s}(t) = \exp(-\lambda_\theta(s)t)$, the loglikelihood function has the form

$$\log L_\theta(y_*,d_*,s_*) = \sum_{d_n=1} (\log \lambda_\theta(s_n) - \lambda_\theta(s_n)y_n) + \sum_{d_n=0} (-\lambda_\theta(s_n)c) = \sum_{n=1}^N l(\theta,t_n,s_n)$$

with

$$l(\theta,t,s) := (\log \lambda_\theta(s) - \lambda_\theta(s)t)\, 1_{[0,c]}(t) - \lambda_\theta(s)c\, 1_{(c,\infty)}(t).$$

The maximum likelihood estimator $\widehat{\theta}$ is a solution of

$$\sum_{n=1}^N \dot{l}(\widehat{\theta},t_n,s_n) = 0,$$

where

$$\dot{l}(\theta,t,s) := \frac{\partial}{\partial\theta} l(\theta,t,s) = \frac{\partial}{\partial\theta}\lambda_\theta(s)\left[\left(\frac{1}{\lambda_\theta(s)} - t\right) 1_{[0,c]}(t) - c\, 1_{(c,\infty)}(t)\right].$$

Set also

$$\ddot{l}(\theta,t,s) := \frac{\partial}{\partial\theta}\dot{l}(\theta,t,s) = \frac{\partial^2}{\partial^2\theta}\lambda_\theta(s)\left[\left(\frac{1}{\lambda_\theta(s)} - t\right)1_{[0,c]}(t) - c\,1_{(c,\infty)}(t)\right] + \frac{\partial}{\partial\theta}\lambda_\theta(s)\frac{\partial}{\partial\theta}\lambda_\theta(s)^\top\left(-\frac{1}{\lambda_\theta(s)^2}\right)1_{[0,c]}(t).$$

Note that we have for any $\lambda > 0$

$$\int_0^c \left(\frac{1}{\lambda} - y\right)\lambda e^{-\lambda y}dy - c\int_c^\infty \lambda e^{-\lambda y}dy = 0, \tag{3}$$

$$\int_0^c \left(\frac{1}{\lambda} - y\right)^2 \lambda e^{-\lambda y}dy + c^2\int_c^\infty \lambda e^{-\lambda y}dy. = \frac{1}{\lambda^2}\left(1 - e^{-\lambda c}\right). \tag{4}$$

Equation (3) implies $E_\theta(\dot{l}(\theta,T_n,s_n)) = 0$ for all stress levels s_n so that the maximum likelihood estimator is not biased. If the stress is a random variable S with distribution δ and the conditional distribution of the life time T given $S = s$ is an exponential distribution with parameter $\lambda_\theta(s)$, then equations (3) and (4) imply

$$E_\theta\left(\dot{l}(\theta,T,S)\,\dot{l}(\theta,T,S)^\top\right) = \int \frac{1}{\lambda_\theta(s)^2}\left(1 - e^{-\lambda_\theta(s)c}\right)\frac{\partial}{\partial\theta}\lambda_\theta(s)\frac{\partial}{\partial\theta}\lambda_\theta(s)^\top\delta(ds) = -E_\theta\left(\ddot{l}(\theta,T,S)\right).$$

If the concrete design measure $\delta_N = \sum_{n=1}^N e_{s_n}$, where e_s is the Dirac measuere on s, is converging to the design measure δ then the maximum likelihood estimator has an asymptotic normal distribution with variance $E_\theta(\dot{l}(\theta,T_n,s_n)\dot{l}(\theta,T_n,s_n)^\top)^{-1}$, i.e.

$$\sqrt{N}(\widehat{\theta} - \theta) \xrightarrow{N\to\infty} \mathcal{N}\left(0,\, I_\theta(\delta)^{-1}\right),$$

where

$$I_\theta(\delta) := \int \frac{1}{\lambda_\theta(s)^2}\left(1 - e^{-\lambda_\theta(s)c}\right)\frac{\partial}{\partial\theta}\lambda_\theta(s)\frac{\partial}{\partial\theta}\lambda_\theta(s)^\top\delta(ds)$$

is the information matrix at the design δ (see e.g. [13] p.421–428).

3 Optimal designs if $\lambda_\theta(s) = \theta s$

Here the information is

$$I_\theta(\delta) = \int \frac{1}{\theta^2}\left(1 - e^{-\theta sc}\right)\delta(ds).$$

Since $1-e^{-\theta sc}$ is strictly increasing in s, the information is maximized if the design puts all its mass on the largest possible value for the stress, i.e. the optimal design on a design region $\mathscr{S}=[S_l,S_u]$ uses only the upper value S_u. However, as soon as there is no censoring, i.e. $c=\infty$, then it does not matter which stress levels are used.

4 Locally D-optimal designs if $\lambda_\theta(s)=\exp(\theta_0+\theta_1 s)$

Setting

$$x_\theta(s):=\sqrt{1-e^{-\exp(\theta_0+\theta_1 s)\,c}}\begin{pmatrix}1\\ s\end{pmatrix}=\sqrt{1-e^{-k\exp(\theta_1 s)}}\begin{pmatrix}1\\ s\end{pmatrix}$$

with $k:=c\exp(\theta_0)$, the information matrix can be expressed here by

$$I_\theta(\delta)=\int\left(1-e^{-k\exp(\theta_1 s)}\right)\begin{pmatrix}1 & s\\ s & s^2\end{pmatrix}\,\delta(ds)=\int x_\theta(s)x_\theta(s)^\top\delta(ds).$$

To derive locally D-optimal two-point designs on $[0,S_u]$, let be $0\le s_1<s_2\le S_u$ and set $X_\theta:=\begin{pmatrix}x_\theta(s_1)^\top\\ x_\theta(s_2)^\top\end{pmatrix}$. Then $\delta_{s_1,s_2}:=\frac{1}{2}e_{s_1}+\frac{1}{2}e_{s_2}$, where e_s is the Dirac measure on s, is D-optimal within all designs with support s_1 and s_2 since with the equivalence theorem of D-optimality [8] we have

$$x_\theta(s_i)^\top I_\theta(\delta_{s_1,s_2})^{-1}x_\theta(s_i)=u_i^\top X_\theta\left(\frac{1}{2}X_\theta^\top X_\theta\right)^{-1}X_\theta^\top u_i=2$$

for $i=1,2$ (here u_i denotes the i'th unit vector in $\Re^2$). The determinant of the information matrix of a design δ_{s_1,s_2} is given by

$$\det(I_\theta(\delta_{s_1,s_2}))=\frac{1}{4}\left(1-e^{-k\exp(\theta_1 s_1)}\right)\left(1-e^{-k\exp(\theta_1 s_2)}\right)[s_2-s_1]^2.$$

Theorem 1. *Let be $k:=c\exp(\theta_0)>0$. Then $\delta_{0,S_u}=\frac{1}{2}e_0+\frac{1}{2}e_{S_u}$ is the D-optimal design within all two-point designs on $\mathscr{S}=[0,S_u]$ if and only if $\theta_1\le\frac{2}{kS_u}(e^k-1)$.*

Proof. Since $1-e^{-k\exp(\theta_1 s)}$ is strictly increasing in s, $\det(I_\theta(\delta_{s_1,s_2}))$ is maximized with respect to $s_2\in(s_1,S_u]$ for any given $s_1\in[0,S_u]$ if and only if $s_2=S_u$. Therefore we have only to determine $s\in[0,S_u]$ so that $\det(I_\theta(\delta_{s,S_u}))$ is maximized. This is equivalent of maximizing

$$g(s)=\left(1-e^{-k\exp(\theta_1 s)}\right)[S_u-s]^2.$$

Since we have

$$g'(s) = e^{-k\exp(\theta_1 s)}\, k\,\theta_1\, \exp(\theta_1\, s)[S_u - s]^2 - 2\left(1 - e^{-k\exp(\theta_1 s)}\right)[S_u - s],$$

δ_{0,S_u} can be only D-optimal if

$$0 \geq g'(0) = e^{-k}\, k\,\theta_1\, S_u^2 - 2\left(1 - e^{-k}\right) S_u \Longleftrightarrow e^{-k}\, k\,\theta_1\, S_u \leq 2\left(1 - e^{-k}\right).$$

This is equivalent with $\theta_1 \leq \frac{2}{kS_u} e^k \left(1 - e^{-k}\right) = \frac{2}{kS_u}(e^k - 1)$. Hence δ_{0,S_u} is not D-optimal if $\theta_1 > \frac{2}{kS_u}(e^k - 1)$. To prove that δ_{0,S_u} is indeed the D-optimal two-point design for $\theta_1 \leq \frac{2}{kS_u}(e^k - 1)$, it is sufficient to prove that g is strictly decreasing on $[0, S_u]$. To show $g'(s) < 0$, we need the monotonicity of some auxiliary functions:
a) We have for $h_1(k) := 1 - e^k + k\,e^k - k^2\,e^k$

$$h_1'(k) = -e^k + e^k + k\,e^k - 2\,k\,e^k - k^2\,e^k = -k\,e^k - k^2\,e^k < 0$$

so that h_1 is strictly decreasing for $k > 0$. Since obviously $h_1(0) = 0$, it holds $h_1(k) < 0$ for all $k > 0$.
b) Now consider $h_2(k) := \frac{2}{k}(e^k - 1) - 1 - 2e^k$. The rule of L'Hospital provides $\lim_{k\downarrow 0} h_2(k) = -1$. Then $h_2(k) < 0$ for all $k \geq 0$ follows with a) from

$$h_2'(k) = -\frac{2}{k^2}(e^k - 1) + \frac{2}{k}\, e^k - 2\, e^k = 2\, k^2\, h_1(k) < 0.$$

c) $\theta_1 \leq \frac{2}{kS_u}(e^k - 1)$ and b) imply for $g_1(s) := \theta_1 [S_u - s] - 1 - 2e^{k\exp(\theta_1 s)}$

$$\begin{aligned} g_1(0) &= \theta_1\, S_u - 1 - 2\, e^k \\ &\leq \frac{2}{k\, S_u}(e^k - 1)\, S_u - 1 - 2\, e^k = \frac{2}{k}(e^k - 1) - 1 - 2e^k = h_2(k) < 0 \end{aligned}$$

for all $k \geq 0$. Because of

$$g_1'(s) = -\theta_1 - 2\, e^{k\exp(\theta_1 s)}\, k\,\theta_1\, \exp(\theta_1 s) < 0,$$

we have $g_1(s) < 0$ for all $k \geq 0$, $s \geq 0$.
d) $\theta_1 \leq \frac{2}{kS_u}(e^k - 1)$ implies for $g_2(s) := k\,\theta_1\, \exp(\theta_1\, s)[S_u - s] + 2 - 2\, e^{k\exp(\theta_1 s)}$

$$g_2(0) = k\,\theta_1\, S_u + 2 - 2\, e^k \leq k\,\frac{2}{k\, S_u}(e^k - 1)\, S_u + 2 - 2\, e^k = 2\, e^k - 2 + 2 - 2\, e^k = 0.$$

Moreover, with c) we obtain

$$\begin{aligned} g_2'(s) &= k\,\theta_1^2\, \exp(\theta_1\, s)[S_u - s] - k\,\theta_1\, \exp(\theta_1\, s) - 2\, e^{k\exp(\theta_1 s)}\, k\,\theta_1\, \exp(\theta_1 s) \\ &= k\,\theta_1\, \exp(\theta_1\, s)\left[\theta_1\, [S_u - s] - 1 - 2\, e^{k\exp(\theta_1 s)}\right] = k\,\theta_1\, \exp(\theta_1\, s)\, g_1(s) < 0 \end{aligned}$$

so that g_2 is strictly decreasing from a value $g_2(0) \leq 0$ which implies $g_2(s) < 0$ for all $k > 0$, $s > 0$.

e) Finally, we have

$$
\begin{aligned}
& g'(s) = [S_u - s]\, e^{-k\exp(\theta_1 s)}\, [k\,\theta_1 \exp(\theta_1 s)[S_u - s] + 2] - 2\,[S_u - s] < 0 \\
\iff & \\
& \quad e^{-k\exp(\theta_1 s)}\, [k\,\theta_1 \exp(\theta_1 s)[S_u - s] + 2] < 2 \\
\iff & \\
& \quad k\,\theta_1 \exp(\theta_1 s)[S_u - s] + 2 < 2\, e^{k\exp(\theta_1 s)} \iff g_2(s) < 0
\end{aligned}
$$

so that d) provides the assertion. □

To prove that δ_{0,S_u} is D-optimal within all designs on $\mathscr{S} = [0, S_u]$, the property

$$
\begin{aligned}
2 \;&\geq\; x_\theta(s)^\top I_\theta(\delta_{0,S_u})^{-1} x_\theta(s) \\
&= \frac{2}{S_u^2}\left(1 - e^{-k\exp(\theta_1 s)}\right)\left(\frac{(S_u - s)^2}{1 - e^{-k}} + \frac{s^2}{1 - e^{-k\exp(\theta_1 S_u)}}\right)
\end{aligned}
\tag{5}
$$

must be shown for all $s \in [0, S_u]$ according to [8] where equality holds only for $s = 0$ and $s = S_u$. The equality is indeed always satisfied for $s = 0$ and $s = S_u$. Set

$$
q(s) := \left(1 - e^{-k\exp(\theta_1 s)}\right)\left(\frac{(S_u - s)^2}{1 - e^{-k}} + \frac{s^2}{1 - e^{-k\exp(\theta_1 S_u)}}\right).
$$

A necessary condition for the D-optimality of δ_{0,S_u} is then $q'(0) \leq 0$.

Lemma 1. *$q'(0) \leq 0$ if and only if $\theta_1 \leq \frac{2}{kS_u}(e^k - 1)$.*

Hence the condition $\theta_1 \leq \frac{2}{kS_u}(e^k - 1)$ implies not only that δ_{0,S_u} is the locally D-optimal design within all two-point designs on $[0, S_u]$ but also the necessary condition for D-optimality of δ_{0,S_u} within all designs on $[0, 1]$. Several plots of $q(s)$ for different values of θ_1 and k with $\theta_1 \leq \frac{2}{kS_u}(e^k - 1)$ showed that q is first decreasing and then increasing on $[0, S_u]$ so that (5) should be satisfied. However, the author was not able to prove it until now.

As soon as $\theta_1 > \frac{2}{kS_u}(e^k - 1)$ holds, then the locally D-optimal two-point design is of the form $\delta_{s(\theta_1,k),S_u}$ with $0 < s(\theta_1, k) < S_u$. The lower points $s(\theta_1, k)$ depending on θ_1 are shown in Fig. 1 for $k = 0.5, 1, 2, 3$ and $S_u = 1$. The condition $\theta_1 > \frac{2}{kS_u}(e^k - 1)$ is in particular satisfied if k is small. The quantity $k := c\exp(\theta_0)$ is small if the censoring variable c or the regression parameter θ_0 is small. A small θ_0 means a high expected lifetime at $s = 0$ which provides a high probability of censoring. Then it is reasonable to make the observations at higher stress levels $s(\theta_1, k) > 0$ so that the probability of censoring is smaller. But since $\frac{2}{kS_u}(e^k - 1) \geq \frac{2}{S_u}$ for all $k \geq 0$, the censoring variable as well as θ_0 have no influence on the D-optimal design as soon as $\theta_1 \leq \frac{2}{S_u}$. The condition $\theta_1 > \frac{2}{kS_u}(e^k - 1)$ is also satisfied if θ_1 is large. In this case, the expected lifetime decreases so rapidly with growing stress that observations at $s(\theta_1, k) > 0$ provide more information than at 0 where observations are censored with higher probability.

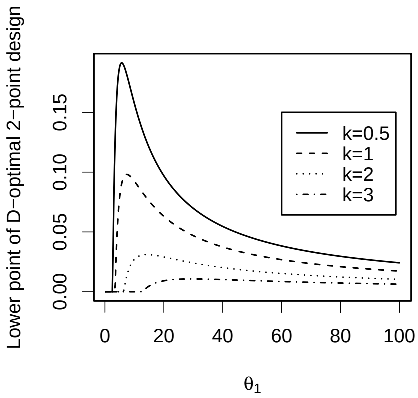

Fig. 1 Lower points $s(\theta_1, k)$ of the D-optimal two-point designs on [0,1].

References

1. Ahmad, N., Islam, A., Salam, A.: Analysis of optimal accelerated life test plans for periodic inspection: The case of exponentiated Weibull failure model. International Journal of Quality & Reliability Management **23**, 1019–1046 (2006).
2. Bai, D.S., Chung, S.W.: An optimal design of accelerated life test for exponential distribution. Reliab. Eng. Syst. Saf. **31**, 57–64 (1991).
3. Barriga, G.D.C., Ho, L.L., Cancho, V.G.: Planning accelerated life tests under Exponentiated-Weibull-Arrhenius model. International Journal of Quality & Reliability Management **25**, 636–653 (2008).
4. Das, R.N., Lin, D.K.J.: On D-optimal robust designs for lifetime improvement experiments. J. Statist. Plann. Inference **141**, 3753-3759 (2011).
5. Fedorov, V.V.: Design of experiments for linear optimality criteria. Theory Probab. Appl. **16**, 189–195 (1971).
6. López-Fidalgo, J., Garcet-Rodríguez, S.: Optimal experimental designs when an independent variable is potentially censored. Statistics **45** , 143–154 (2011).
7. Haibach, E.: Betriebsfestigkeit. Verfahren und Daten zur Bauteilberechnung. Springer, Berlin (2006).
8. Kiefer, J. and Wolfowitz, J.: The equivalence of two extremum problems. Can. J. Math. **12**, 363–366 (1960).
9. Klein, J.P., Moeschberger, M.L.: Survival Analysis: Techniques for Censored and Truncated Data. Springer, New York (2003).
10. Liang, T., Chen, L.-S., Yang, M.-C.: Efficient Bayesian sampling plans for exponential distributions with random censoring. J. Statist. Plann. Inference **142**, 537-551 (2012).
11. Pal, M., Mandal, N.K.: Optimum designs for stress strength reliability. J. Indian Soc. Agric. Stat. **63**, 305–310 (2009).
12. Park, S., Ng, H.K.T.: Missing information and an optimal one-step plan in a Type II progressive censoring scheme. Stat. Probab. Lett. **82**, 396–402 (2012).
13. Schervish, M.J.: Theory of Statistics. Springer, New York (1995).
14. Tsai, T.-R., Lin, C.-W.: Acceptance sampling plans under progressive interval censoring with likelihood ratio. Statistical Papers **51**, 259–271 (2010).
15. Wu, S.-J., Huang, S.-R.: Optimal progressive group-censoring plans for exponential distribution in presence of cost constraint. Statistical Papers **51**, 431–443 (2010).