Applied Stochastic Models in Business and Industry

Discussion

(wileyonlinelibrary.com) DOI: 10.1002/asmb.2071 Published online in Wiley Online Library

Discussion of 'Methods for planning repeated measures accelerated degradation tests'

1. Introduction

The article *Methods For Planning Repeated Measures Accelerated Degradation Tests* written by B. P. Weaver and W. Q. Meeker provides an interesting extension in the area of optimal accelerated test planning. From the point of view of traditional accelerated design approaches where the lifetime is modeled as a function of stress parameters, it can be seen as a motivation to measure and model the degradation instead of waiting for a life-terminating event and analysing heavily censored data. Further, the applicability is limited to cases where the genesis of a failure can be reliably observed and reasonably measured.

The paper guides the reader in a structured way throughout the course. After introducing degradation models with mixed effects as linear functions of the acceleration variables, the authors derive a failure-time distribution, which allows one to estimate the quantile of interest of the time to failure distribution. In alignment with the classical design theory, the optimal design is based on the Fisher information matrix and takes practical boundaries into account considering design restrictions. The applicability of this planning procedure is demonstrated by two industrial examples.

2. The model

Usually, acceleration is applied to increase the effectivity of testing or (may be able) to assess stress-life relations. In particular, if the sample size is small—which is frequently the case in industrial applications—results may be heavily influenced by single observations. In such cases, the consideration of random effects allows to quantify the amount of dispersion caused by unit-to-unit variation.

The theoretical concepts applied in Weaver and Meeker's paper are well established (asymptotic normal theory, Fisher information, etc.). However, it's originality lies in the model approach, which considers the unit-to-unit variation by random effects. Another elegant aspect is the model formulation of the degradation Y as

$$Y = x_1\gamma_1 + x_2\gamma_2\tau + b_0 + b_1\tau + \epsilon$$

with $\epsilon \overset{iid}{\sim} N\left(0, \sigma^2\right)$. The unit-to-unit variability of the degradation is described by b_0 and b_1 where $\left(b_0, b_1\right)^T \sim N(\boldsymbol{\beta}, \mathbf{V})$ with mean vector $\boldsymbol{\beta} = \left(\beta_0, \beta_1\right)^T$ and covariance matrix $\mathbf{V}$ where $\mathbf{V}$ contains the parameters σ_{b_0}, σ_{b_1} and ρ. Moreover, $\left(b_0, b_1\right)^T$ is independent of ϵ and the ϵ values are independent over time. In the model, x_1 and x_2 can be occupied with the same variable, which is relevant for the frequent case of one stress variable, but also allows in principle the flexibility to consider two stress variables without increasing model complexity. Another aspect is the variance estimation based on the Fisher information matrix $\mathbf{I}$. The convenient way of composing the information of a design by adding the contributions of the tests has already been used in papers of Nelson and Kielpinski [1], Nelson and Meeker [2], Escobar and Meeker [3], and Escobar and Meeker [4].

3. The Objective

The objective of the test plan is to estimate the $p-$quantile of the time until the degradation measure exceeds a given degradation limit as precisely as possible. The authors present optimal designs and compromise designs in the style of Nelson [5], where the asymptotic variance $\mathrm{AVar}\left(\hat{t}_p\right)$ of the interesting quantile estimate $\hat{t}_p$ is minimized. Here, using the delta method, $\mathrm{AVar}\left(\hat{t}_p\right) = \mathbf{a}^T\hat{\mathbf{C}}\mathbf{a}$, with $\hat{\mathbf{C}}$ as large sample approximation of the covariance matrix of the maximum likelihood estimators and $a_i = \partial t_p / \partial \theta_i$ where vector $\boldsymbol{\theta}$ contains the unknown parameters of the degradation model. The authors apply a directional derivative according to Whittle's (1973) general equivalence theorem as $\Lambda(\eta, \nu) = \mathbf{a}^T (\mathbf{I}(\eta))^{-1} \mathbf{I}(\nu) (\mathbf{I}(\eta))^{-1} \mathbf{a} - \mathbf{a}^T (\mathbf{I}(\eta))^{-1} \mathbf{a}$. A design η provides smaller $\mathrm{AVar}(\hat{t}_p)$ than a design ν if $\Lambda(\eta, \nu) > 0$.

4. Sensitivity and robustness of the optimal design

The optimal and (at least to some extent) the compromise design generated that way depend both on the model as well as on the input data. Among several properties, which could have been investigated, we focused on two aspects: at first, we investigated the sensitivity by comparing a grid of candidate designs with the design chosen by Weaver and Meeker. In a second step, we evaluated the robustness of the optimal solutions regarding the influence of the input data in using bootstrap simulations. For both tasks, we examined the examples of Weaver and Meeker's paper by using a collection of continuous designs $\nu = (\mathbf{x}, \boldsymbol{\pi})$. With fixed boundary stress levels x_L (low-stress level) and x_H (high-stress level) as well as fixed proportions at the bounds, π_L and π_H, we varied the inner level x_M so that $x_L \leqslant x_M \leqslant x_H$ and $\sum_{j \in \{L,M,H\}} \pi_j = 1$ with $\pi_j > 0 \; \forall j$.

4.1. Sensitivity analysis of example 1

Example 1 from Weaver and Meeker's paper deals with an example of carbon resistors used in Shiomi and Yanagisawa (1979). Here, the degradation model for unit i measured at time point j for acceleration level k is the resistance in Ohms $y_{ijk} = \gamma_2 x_k \tau_{ij} + b_{0i} + b_{1i}\tau_{ij} + \epsilon_{ijk}$ with $\tau = \sqrt{\text{Time in hours}}$, and stress variable $x = -11605/(\text{T in } {}^oC + 273.15)$ and $e_{ij} \overset{iid}{\sim} N(0, \sigma^2)$ motivated by the Arrhenius temperature model.

With the correspondingly given original data, the maximum likelihood estimation has been carried out with the `R` function `lme()` and leads to the following output where τ is denoted by `sqrt(tme)` and x as `xTmp2`:

```
Linear mixed-effects model fit by maximum likelihood
 Data: d_D1
       AIC      BIC    logLik
  449.0865 469.9236 -217.5433

Random effects:
 Formula: ~I(sqrt(tme)) | as.factor(unitNb)
 Structure: General positive-definite, Log-Cholesky parametrization
             StdDev     Corr
(Intercept)  1.89720371 (Intr)
I(sqrt(tme)) 0.03815121 0.544
Residual     0.52305075

Fixed effects: yRes ~ I(sqrt(tme)) + I(sqrt(tme) * xTmp2)
                         Value Std.Error  DF  t-value p-value
(Intercept)          218.51308 0.3632262 114 601.5896       0
I(sqrt(tme))           0.65548 0.0690617 114   9.4912       0
I(sqrt(tme) * xTmp2)   0.02032 0.0023735 114   8.5626       0
 Correlation:
                     (Intr) I(s())
I(sqrt(tme))         0.052
I(sqrt(tme) * xTmp2) 0.000  0.994

Standardized Within-Group Residuals:
       Min         Q1        Med         Q3        Max
-4.1537723 -0.2613070  0.1079996  0.3532090  2.8586171

Number of Observations: 145
Number of Groups: 29
```

This result shows highly significant fixed effects and the correlation between the second and the third parameter estimator, the effects of τ and $\tau{:}x$, is 0.994. Here, the discussion arises whether both terms would be required in the model, respectively, whether the effect of τ should be modeled as a random effect. Hence, we discussed different linear alternatives where we observed either a high correlation between two parameter estimators or a substantial lack of fit, if one of the parameters is excluded.

We are now interested in the sensitivity of Weaver and Meeker's optimal 3-level design η with $\mathbf{x} = (50, 83, 173)$ and $\boldsymbol{\pi} = (0.05, 0.711, 0.239)$ when varying x_M and π_M. Considering of the boundary conditions stated in the example, we investigated the 3-level design where $\nu = (\mathbf{x}, \boldsymbol{\pi}) = \left(\left(50, x_M, 173\right), \left(0.05, \pi_M, 1 - 0.05 - \pi_M\right)\right)$ with $x \in [50, 173]$ and $\pi_x \in [0.05, 0.9]$.

The black dot shows Weaver and Meeker's optimal design with $x_M = 83$ and $\pi_M = 0.711$. Figure 1 exhibits that their optimal design has the highest stress level x_M of all designs on the isoline $\Lambda(\eta, \nu) = 0$. An interesting aspect is that for all $x_M \in [63, 83]$, two different optimal designs exist with different π_M, that is, the two designs $\mathbf{x} = (50, 70, 173)$, $\boldsymbol{\pi} = (0.05, 0.459, 0.491)$ and $\mathbf{x} = (50, 70, 173)$, $\boldsymbol{\pi} = (0.05, 0.855, 0.095)$ appear equal with respect to Λ, although from a practical perspective, the second solution is more unbalanced and would probably require more execution time due to the reciprocal relation between stress time to failure. Following the isoline towards lower π_M, we end up at $x_M = 50$ with $\pi_M = 0.316$, which shows in fact the equivalent 2-level design $\mathbf{x} = (50, 173), \boldsymbol{\pi} = (0.366, 0.634)$. Again, practical reasons may call for a 3-level design and thus avoid the 2-level design solution. Therefore, Weaver and Meeker's optimal design for example 1 seems to be adequate from a theoretical as well as a practical perspective.

4.2. *Sensitivity analysis of example 2*

Example 2 from Weaver and Meeker's paper shows the wear resistance of metal plates encountered load situations with different weights, known from Meeker and Escobar (1998). The degradation model for unit i $(i = 1, \ldots, 12)$ measured at time point j $(j = 1, \ldots, 8)$ for acceleration level k $(k = 1, 2, 3)$ is the log of the scar width in microns: $y_{ijk} = \gamma_1 x_{1k} + \gamma_2 x_{2k} \tau_{ij} + b_{0i} + b_{1i}\tau_{ij} + \epsilon_{ijk}$ with $\tau = \log(\text{Time in kilocycles})$, stress variable $x_1 = x_2$ as weight in grams and $\epsilon_{ijk} \overset{iid}{\sim} N\left(0, \sigma^2\right)$.

The correlation between time and interaction between time and weight is here -0.823, which is strong, but compared with the corresponding effect of example 1, somewhat weaker. We investigated again the sensitivity, but this time, of Weaver and Meeker's optimal 2-level design η with $\mathbf{x} = (10, 100)$ and $\boldsymbol{\pi} = (0.95, 0.05)$ when varying x_M and π_M. Although the lower bound of the testing weight was given with 10g, for comparison, we additionally investigated candidate designs $(\mathbf{x}, \boldsymbol{\pi}) = \left(\left(x_M \geqslant 5, 100\right), \left(\pi_M \geqslant 0.05, 1 - \pi_M\right)\right)$ with $\pi_L = 0$.

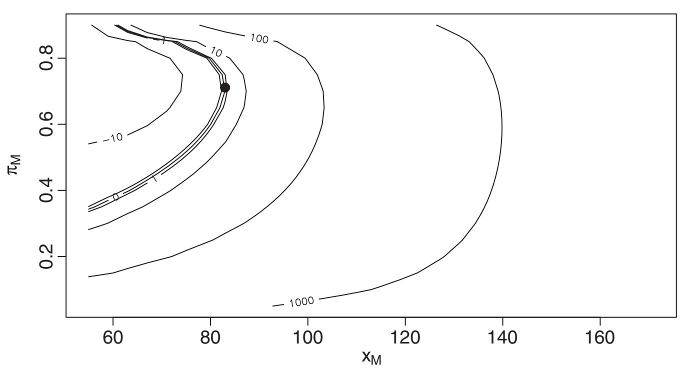

Figure 1. Directional derivative $\Lambda(\eta, \nu)$ for designs ν with $x_M \in [50, 173]$ and $\pi_M \in [0.05, 0.9]$ compared with η as Weaver and Meeker's optimal design for example 1.

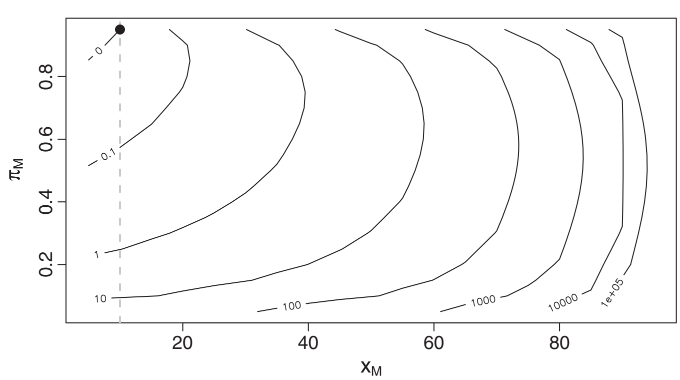

Figure 2. Directional derivative $\Lambda(\eta, \nu)$ for designs ν with $x_M \in [5, 100]$ and $\pi_M \in [0.05, 0.95]$ compared to η as Weaver and Meeker's optimal design for example 2.

Applied Stochastic Models in Business and Industry

The black dot in Figure 2 shows again Weaver and Meeker's optimal design with $x_M = 10$ and $\pi_M = 0.95$ $(\pi_L = 0)$. Compared with example 1, the range of designs, which appear statistically equivalent to the optimal design, is smaller here. Considering the given lower bound for the test of 10g in Weaver and Meeker's paper, the optimum is indeed the unique solution with $x_M = 10$ and $\pi_M = 0.95$. If the lower bound for testing is extended down to the use weight of 5g, alternative optimal designs can be found. These lie between $\mathbf{x} = (5, 100)$, $\boldsymbol{\pi} = (0.853, 0.147)$ and $\mathbf{x} = (10, 100)$, $\boldsymbol{\pi} = (0.95, 0.05)$. In particular, the solution with $x_M = 5$ provides the least unbalance and the biggest proportion at x_H among the competitive optimal designs.

5. Robustness of the designs

The quantile estimator's variance is a function of the Fisher information, which is evaluated at the maximum likelihood estimate of the model parameters. Thus, $\mathrm{AVar}(\hat{t}_p)$ strongly depends on the input data. By using a blocked bootstrap simulation, we wanted to see how strongly the parameter estimators and consequently the optimal test plan depend on the variability in the input data. The following steps have been executed per simulation replicate:

- Select randomly m_k data records with replication on each stress level k where m_k is the number of specimens tested at stress level k in the original data set. To avoid an artificial bias concerning degradation, a data record is the complete measurement series of one specimen. E.g., for example 2 of Weaver and Meeker's paper, data record i contains the eight corresponding design points including the measurements of the response of item i, $i = 1, \dots, 12$.
- Carry out the maximum likelihood estimation for the generated data set according to the model described in Weaver and Meeker's paper
- Estimate $\mathrm{AVar}(t_p)$ for each candidate design $(\mathbf{x}, \boldsymbol{\pi})$ on the grid as initialized earlier in the sensitivity analysis
- Check the directional derivative on the grid with respect to Weaver and Meeker's optimal design.

The simulation study has been carried out again for both examples presented in Weaver and Meeker's paper.

5.1. Robustness of example 1

The simulation shows different variability in the parameters as exhibited in Figure 3. Although b_0 can be estimated very precisely, the coefficients of variation CV for the other model parameters vary between 10% and 30%. In this context, we define the CV of a parameter estimator by $100\% * \text{standard error}/\text{estimate}$. What is remarkable is that each single simulation run leads exactly to the same optimal design, that is, in fact, CV = 0 for x_M and π_M.

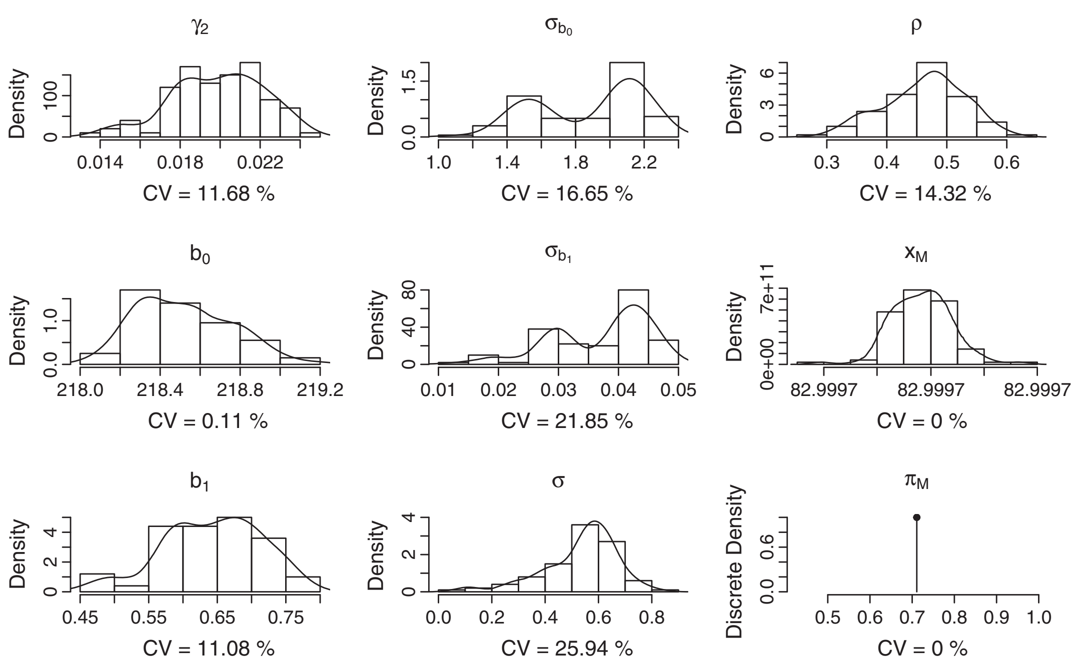

Figure 3. Distribution of parameter estimates and optimal designs for example 1 checked by directional derivative.

Not only the optimum is identical but also the isolines for $\Lambda(\eta, \nu) = 0$ are identical, which is shown in Figure 4. There, isolines out of 1000 simulations coincide in one solution. Thus, for example 1 the optimal test plan appears robust against the variability of the input data. A technical reason for this result is that the Fisher information is composed of elements $\mathbf{X}^T\mathbf{V}\mathbf{X}$ where $\mathbf{X}$ contains the design aspects and $\mathbf{V}$ the covariance structure of the model. Because the t-test statistics of all fixed effects are > 8, the residual sum of squares is very small compared to the regression sum of squares in each replicate of the blocked bootstrap simulation. Thus, the design aspect is dominating over the aspect of stochastic uncertainty, which justifies the approach proposed by Weaver and Meeker.

5.2. Robustness of Example 2

The simulation shows for example 2 even more different variability in the parameters as for example 1. Whereas here the fixed effects γ_1 and γ_2 as well as b_0 can be estimated very precisely (CV $\approx$ 2%), the coefficients of variation for the variance

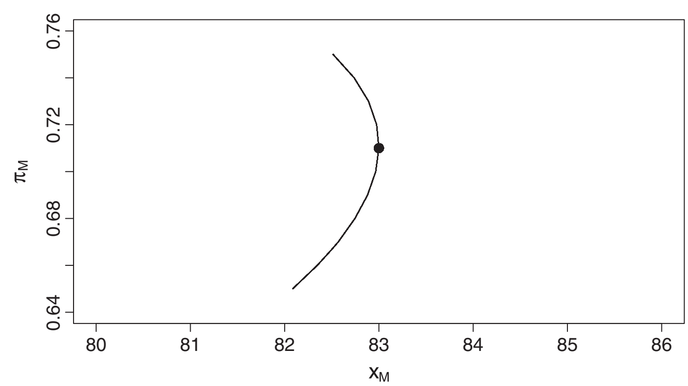

Figure 4. Directional derivative $\Lambda(\eta, \nu)$ for 1000 bootstrap-simulated designs ν with $x_M \in [10, 173]$ and $\pi_M \in [0.05, 0.95]$ compared with η as Weaver and Meeker's optimal design for example 1, which is represented by the black dot.

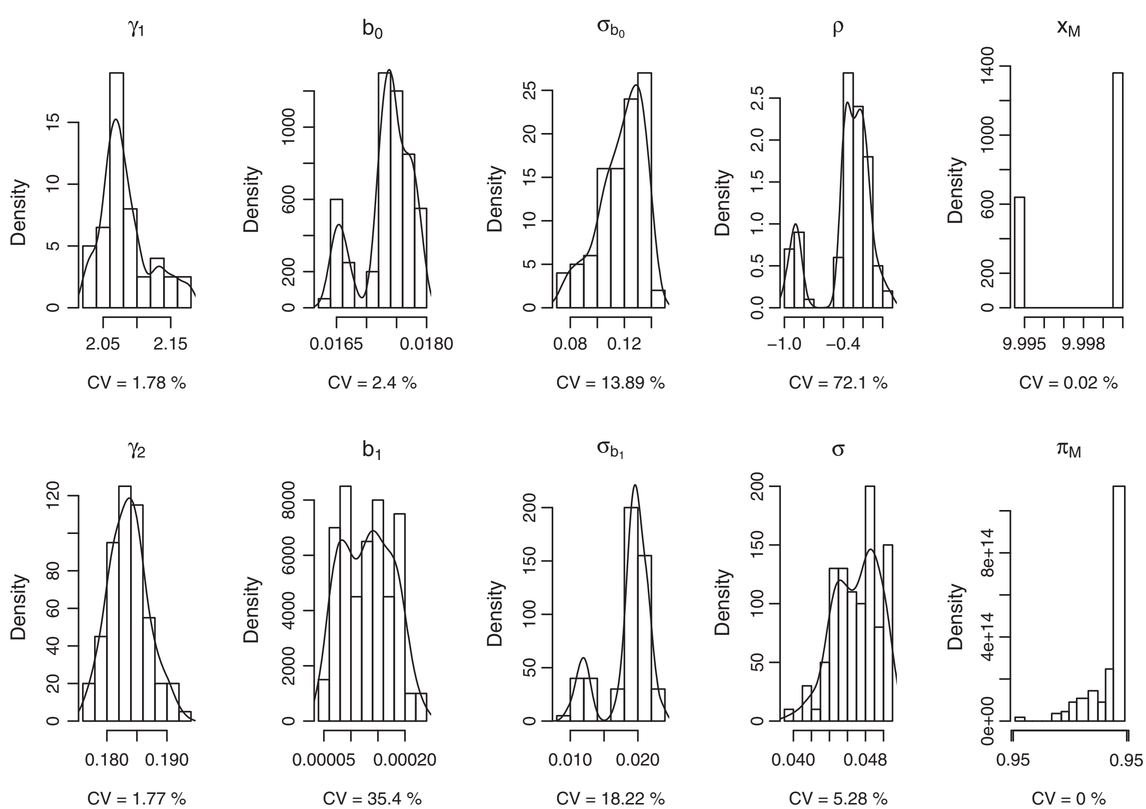

Figure 5. Distribution of parameter estimators and optimal designs for example 2 checked by directional derivative.

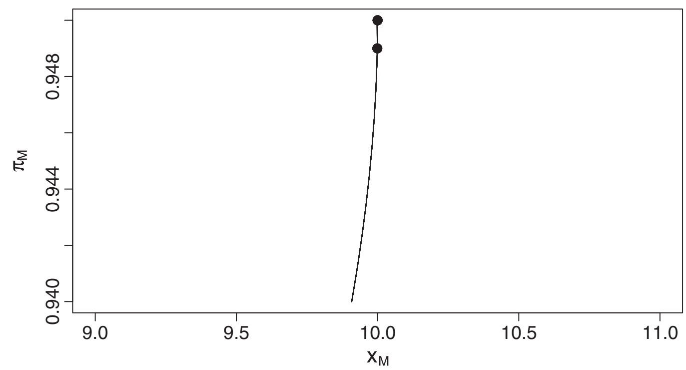

Figure 6. Directional derivative $\Lambda(\eta, \nu)$ for 1000 bootstrap-simulated designs ν with $x_M \in [10, 173]$ and $\pi_M \in [0.05, 0.95]$ compared with η as Weaver and Meeker's optimal design for example 2, which is represented by the black dots.

related parameters vary between 5% and 72% for the correlation ρ between b_0 and b_1. What is remarkable in this case is the rather large variability of the estimator of b_1 as well as a bimodal shape of the simulation distributions of b_0, σ_{b_1} and ρ which is shown in Figure 5.

Although the estimates of σ_{b_1} and ρ correlate they do not explain the small variation in x_M. Among all simulation runs, we only found two different optimal designs, which are located very closely to each other $\left(\pi_{M_1} = \pi_{M_2} = 0.95 \text{ and } x_{M_1} = 9.995, x_{M_2} = 10\right)$ as Figure 6 indicates. Thus, the planning procedure also for example 2 provides an optimal design, which is robust against the variability in the input data.

6. Conclusion

The planning of accelerated tests based on random effect models is a process that shows various advantages such as a sensitive optimization and robustness against the variability of the input data. The model properties are not perfect because we observed high correlations between some effect estimates, which might indicate a potential for simplification. Finally, we congratulate B. Weaver and W. Meeker on their stimulating paper.

NIKOLAUS HASELGRUBER
CIS Consulting in Industial Statistics
A-4170 Haslach
E-mail: *nh@cis-on.com*

ERNST STADLOBER
Graz University of Technology
Institute of Statistics
TU Graz, A-8010 Graz

References

1. Nelson W, Kielpinski TJ. Theory for optimum censored accelerated tests for normal and log-normal life distributions. *Technometrics* 1976; **18**(1):105–114.
2. Nelson W, Meeker WQ. Theory for optimum accelerated censored life tests for Weibull and extreme value distributions. *Technometrics* 1978; **20**(2):171–177.
3. Escobar LA, Meeker WQ. Elements of the Fisher information matrix for the smallest extreme value distribution and censored data. *Applied Statistics* 1986; **35**(1):80–91.
4. Escobar LA, Meeker WQ. Fisher information matrix for the ex-treme value, normal and logistic distributions and censored data. *Applied Statistics* 1994; **43**(3):533–540.
5. Nelson W. *Accelerated Testing*. John Wiley & Sons, Inc.: New York, 1990.