Reliability Engineering and System Safety 131 (2014) 291–297

Contents lists available at ScienceDirect

Reliability Engineering and System Safety

journal homepage: www.elsevier.com/locate/ress

On selection of optimal stochastic model for accelerated life testing

P. Volf*, J. Timková

Institute of Information Theory and Automation, UTIA, Pod vodarenskou vezi 4, 18208 Praha 8, Czech Republic

ARTICLE INFO

Available online 2 May 2014

Keywords:
Reliability analysis
Accelerated life test
Cox's model
AFT model
Goodness-of-fit
Martingale residuals
Bayes statistics
MCMC

ABSTRACT

This paper deals with the problem of proper lifetime model selection in the context of statistical reliability analysis. Namely, we consider regression models describing the dependence of failure intensities on a covariate, for instance, a stressor. Testing the model fit is standardly based on the so-called martingale residuals. Their analysis has already been studied by many authors. Nevertheless, the Bayes approach to the problem, in spite of its advantages, is just developing. We shall present the Bayes procedure of estimation in several semi-parametric regression models of failure intensity. Then, our main concern is the Bayes construction of residual processes and goodness-of-fit tests based on them. The method is illustrated with both artificial and real-data examples.

1. Introduction

Accelerated life testing is a standard approach to gather information on the survival time of highly reliable devices. One of the goals of statistical analysis consists in the construction of a model of the time to failure dependence on the 'stressor' (in a quite wide sense). As a rule, the stressor is taken as a covariate in a regression model of the lifetime. The model should be selected in such a way that the information obtained under the over-stress could be extrapolated to standard stress conditions. These problems, including the test design, selection of models, procedures of statistical analysis, have been treated in a number of papers and books, for instance [9,10,4,5]. Nowadays, many authors prefer the Bayes approach, though mostly in the framework of parametrized (e.g. Weibull) models. Simultaneously, computations are supported by the Markov Chain Monte Carlo (MCMC) generation of posterior and predictive distribution, as in Van Dorp and Mazzuchi [14]. In the same context, Erto and Giorgio [6] accent the advantage of utilization of prior information, an experience from past tests as well as the expert knowledge. Wang et al. [16] model and analyze the process of degradation, instead of failure times directly, using a Gauss or gamma process as a baseline source of uncertainty. They provide the Bayes method and the MCMC procedure enabling one to combine accelerated laboratory tests with field data in order to analyze the reliability of system.

The selection of a proper stochastic model is just one of the steps of statistical analysis. The model criticism, including the goodness-of-fit tests, should follow. Therefore, the methods of goodness-of-fit statistical testing are in the center of our attention. In the present paper we consider three basic semi-parametric regression models describing the dependence of intensity of failures on covariates, in our context on the load, stress or other conditions of usage.

In the framework of intensity models for lifetime data, the goodness-of-fit tests are often based on the analysis of residual process (martingale residuals). The residual process is defined as a difference between estimated cumulated intensity and observed counting process of failures (see for instance [3]). Hence, the residual process is constructed from the observed data, its properties depend on the properties of the estimator of the cumulated hazard rate. In a case without regression, as well as in Aalen's additive regression model, residual processes are the martingales [15]. In some other cases, as is Cox's model or the accelerated failure time (AFT) model, the behavior of estimates, and therefore of residuals, is more complicated. That is why the tests are often performed just graphically [1]. Approximate critical regions for tests can also be obtained by random generation from asymptotic distribution of residual processes. Relevant theoretical results can be found for instance in Andersen et al. [3], Lin et al. [11], and Bagdonavicius and Nikulin [4]. In such cases, the Bayes approach can offer a reasonable alternative, especially when connected with the MCMC methods (an overview of the MCMC is given for instance in [7]). The present paper deals prevailingly with semi-parametric intensity models consisting of a parametric regression part and a nonparametric baseline hazard rate. For the Bayes solution, its representation can be made from piecewise-constant functions (or from splines or from other functional basis), in the way used in Arjas and Gasbarra [2]. Once a posterior sample of

* Corresponding author. Tel.: +420 266052431.
E-mail address: volf@utia.cas.cz (P. Volf).

http://dx.doi.org/10.1016/j.ress.2014.04.015
0951-8320/

hazard rate (i.e. representation of its posterior distribution obtained by the MCMC procedure) is available, we can construct a sample representing cumulated intensities and corresponding residuals.

Let us here also recall another approach to the Bayes analysis in the AFT model. It utilizes logarithmic model formulation. Instead of the baseline hazard rate of a baseline survival time T_0 it deals with the density for $\log T_0$. Often, its prior is constructed as a mixture of the Gauss densities with weights given by Dirichlet distributions (as for instance in [8]). However, complications are caused by censoring and have to be overcome by an additional generation of would-be non-censored values, i.e. by a data augmentation. It is actually a randomized version of the EM algorithm.

The present paper has the following structure: In the next section, the notion of martingale residuals is recalled, then the Bayes nonparametric approach to intensity modeling is described. While these sections are more-less introductory, the core of the paper lies in Sections 4–6 dealing with regression models, methods of analysis and their Bayesian counterparts. Utilization of the MCMC procedures leads to the Bayes 'empirical' construction of residual processes. The method is finally, in Section 7, illustrated with both artificial and real-data examples.

2. Martingale residuals

In order to introduce the notion of martingale residuals, we shall first consider a standard survival data case, without any dependence on covariates. Let us imagine that a set of i.i.d. random variables T_i, survival times of n objects of the same type, is observed. Alternatively, we may consider their counting processes $N_i(t)$, each having maximally 1 count (at the time of failure, T_i), or being censored without failure. Further, let us also consider indicator processes (of being at risk) $Y_i(t)$, $Y_i(t)=0$ after failure or censoring, $Y_i(t)=1$ otherwise. As the lifetimes are i.i.d., corresponding counting processes have the same common hazard rate $h(t)\geq 0$. The cumulated hazard rate is then $H(t)=\int_0^t h(s)\,ds$. It follows that the intensity of $N_i(t)$ is $a_i(t)=h(t)\cdot Y_i(t)$. Notice a difference between those two notions: the hazard rate is a characteristic of distribution, namely here $h(t)=-d(\ln \overline{F})(t)/dt$, where $\overline{F}(t)=1-F(t)$ is a survival function, complement to the distribution function, while the intensity depends on realizations of processes $Y_i(t)$. It is assumed that the data are observed on a finite time interval $t\in[0,T]$, $N_i(0)=0$.

Let us also define sums of individual characteristics, namely counting process $N(t)=\sum_{i=1}^n N_i(t)$ counting number of failures, further $Y(t)=\sum_{i=1}^n Y_i(t)$, cumulated intensities $A_i(t)=\int_0^t a_i(s)\,ds$ and $A(t)=\sum_{i=1}^n A_i(t)$, so that here $A(t)=\int_0^t h(s)Y(s)\,ds$.

In theoretical studies on lifetime models, many results are based on martingale—compensator decomposition of counting process, namely that $N_i(t)=A_i(t)+M_i(t)$, so that also $N(t)=A(t)+M(t)$, where $M_i(t)$ and $M(t)$ are martingales with zero means, conditional variance processes (conditioned by corresponding filtration, a nondecreasing set of σ-algebras $\mathcal{F}(t^-)$) are $\langle M_i\rangle(t)=A_i(t)$ and $\langle M\rangle(t)=A(t)$. Naturally, martingales have non-correlated increments, and $M_i(t)$ are also non-correlated mutually (for different i).

Then it is quite reasonable to consider a residual process (martingale residuals)

$$R(t)=N(t)-\hat{A}(t)=M(t)+A(t)-\hat{A}(t)$$

as a tool for testing model fit. Here $\hat{A}(t)$ is the estimated cumulated intensity. Hence, the residual process is constructed from the observed data, and its properties depend mainly on the properties of the estimator of the cumulated hazard rate, because $\hat{A}(t)=\int_0^t Y(s)\,d\hat{H}(s)$. Tests are then performed either graphically or numerically, critical borders for assessing the goodness-of-fit are based on the asymptotic properties of estimates.

2.1. Properties of residuals

The most common estimator of cumulated hazard rate $H(t)$ is the Nelson–Aalen estimator, which has the form

$$\hat{H}(t)=\int_0^t \sum_{i=1}^n \frac{dN_i(s)}{\sum_{j=1}^n Y_j(s)}=\int_0^t \frac{dN(s)}{Y(s)},$$

so that it is a piecewise constant function with jumps $d\hat{H}(s)=dN(s)/Y(s)$ at times where failures have occurred. Its asymptotic properties, namely uniform on $[0,T]$ consistency in probability and asymptotic normality when $n\to\infty$, are well known (for review of survival analysis, see for instance [10]). More precisely, the following convergence in distribution on $[0,T]$ to the Brown motion process $\mathcal{B}$ holds

$$\sqrt{n}(\hat{H}(t)-H(t))\to^d \mathcal{B}(V(t)),\quad V(t)=\int_0^t \frac{h(s)\,ds}{c_0(s)},$$

where we assume the existence of $c_0(s)=P-\lim Y(s)/n$, uniformly in $[0,T]$, $c_0(s)\geq\varepsilon>0$. Hence, it is possible to construct Kolmogorov–Smirnov type confidence bands for $H(t)$ as well as pointwise confidence intervals. Again, a consistent, uniformly in $[0,T]$, estimator of $V(t)$ is available: $\hat{V}(t)=\int_0^t n\,dN(s)/Y(s)^2$.

In the present contribution we are interested mainly in the properties of residual process $R(t)=N(t)-\hat{A}(t)$. Notice that here $\hat{A}(t)=N(t)$ directly, so that it is preferred to construct residuals in data subsets (strata), $S\subset\{1,..,n\}$. Thus, let us define

$$R_S(t)=N_S(t)-\hat{A}_S(t)=M_S(t)+A_S(t)-\hat{A}_S(t),$$

where we denote again $N(t)=\sum_{i=1}^n N_i(t)$, $N_S(t)=\sum_{i\in S}N_i(t)$, similarly for $Y(t), M(t), A(t), \hat{A}(t)$. As

$$\hat{A}_S(t)=\int_0^t \sum_{i\in S} d\hat{H}(r)Y_i(r)=\int_0^t \frac{dN(r)}{Y(r)}\cdot Y_S(r)$$
$$=\int_0^t \frac{dH(r)Y(r)+dM(r)}{Y(r)}\cdot Y_S(r)=A_S(t)+\int_0^t \frac{dM(r)}{Y(r)}\cdot Y_S(r),$$

we obtain that (with notation $\overline{S}$ – complement of S)

$$R_S(t)=M_S(t)-\int_0^t \frac{dM(r)}{Y(r)}\cdot Y_S(r)=\int_0^t \frac{dM_S(r)Y_{\overline{S}}(r)-dM_{\overline{S}}(r)Y_S(r)}{Y(r)}.$$

From its structure it follows that the process $R_S(t)$ has non-correlated increments, conditioned variance (by σ-algebras $\mathcal{F}(t^-)$) of $(1/\sqrt{n})\,dR_S(t)$ is

$$\frac{dH(t)}{nY(t)^2}(Y_{\overline{S}}(t)Y_S(t)^2+Y_{\overline{S}}(t)^2Y_S(t))\sim dH(t)\frac{c_S(t)c_{\overline{S}}(t)}{c_0(t)},$$

where we again assume that there exist P-limits $Y_S(t)/n\to c_S(t)$, $Y_{\overline{S}}(t)/n\to c_{\overline{S}}(t)$, $Y(t)/n\to c_0(t)$, uniform in $t\in[0,T]$, bounded away from zero. Then $(1/\sqrt{n})R_S(t)\to^d \mathcal{B}(V_R(t))$, i.e. it converges to the Brown motion process, too, and the asymptotic variance function $V_R(t)$ is consistently estimable by

$$\hat{V}_R(t)=\int_0^t \frac{d\hat{H}(r)Y_S(r)Y_{\overline{S}}(r)}{nY(r)}=\int_0^t \frac{dN(r)Y_S(r)Y_{\overline{S}}(r)}{nY(r)^2}.$$

Hence, if assumptions of our model hold, the process

$$\frac{1}{\sqrt{n}}\frac{R_S(t)}{(1+\hat{V}_R(t))}$$

should behave asymptotically as the Brown bridge process. It can be tested by the Kolmogorov–Smirnov criterion (or other similar criteria, as is the Cramer–von Mises test). Therefore, in such a simple case of survival model without any non-heterogeneity, the method can be used for assessing the model fit in different subsets S. However, in cases of regression models, the test are not so straightforward. That is why we shall continue by description of the Bayes variant of residual analysis.

3. Bayes residuals

The Bayes approach to statistical analysis assumes that all model components (i.e. the parameters as well as non-parametrized parts) are random quantities, initially with a prior probability distribution. The result of statistical analysis is then a posterior distribution of those model components, i.e. their estimate is a distribution. Actually it is the likelihood function 'modulated' by prior distribution.

From another point of view, it is possible to say that while the "standard statistics" studies the variation of data and its consequence when inserted to given functions (estimators), in the Bayes statistics the main concern is the variation of 'parameters', data are taken as fixed.

In the case considered here we deal with the nonparametric hazard rate. For the Bayes solution, its representation can be made from piece-wise constant functions, as in Arjas and Gasbarra [2]. Parameters are then the points of changes of hazard rate, also their number in $[0, T]$, and the levels of hazard rate in intervals between these points. Arjas and Gasbarra [2] show how the MCMC generation can follow the Gibbs sampler combined with an 'accept–reject' sampling method.

Once we have a posterior sample of 'hazard rates', $h^{(m)}(t)$ (i.e. last M representatives of posterior distribution obtained by the MCMC procedure), we can construct from them a sample of cumulated intensities in subgroup S and corresponding residuals:

$$A_S^{(m)}(t) = \int_0^t h^{(m)}(r) Y_S(r)\, dr, \quad R_S^{(m)}(t) = N_S(t) - A_S^{(m)}(t).$$

3.1. Bayes confidence regions

Point-wise (at each t) sample quantiles in a set $R_S^{(m)}(t)$ are obtained immediately, showing the so-called credibility intervals (Bayesian version of confidence intervals) for $R_S(t)$. Methods for construction of confidence bands (of Bayes type) on the whole interval $[0, T]$ are studied intensively nowadays. Theoretically, this problem is connected with the concept of 'depth of data' (see for instance [17]). Practically, the solution corresponds to the construction of multivariate quantiles, for instance in the following way: Let us consider a sample of functions $f^{(m)}(x)$, $m = 1, .., M$, given empirically by values at the same set of points $x_j, j = 1, ..., J$. For each $k < M/2$ point-wise sample k/M and $(M-k)/M$ quantiles (i.e. at each x_j) can be constructed. If we join them to bands, we can try to find such k that, approximately, a given proportion (95%, say) of functions lies inside. As an additional finer criterion we can compare numbers of points at which the quantiles are crossed.

4. Residuals in regression models

In the follow-up, it is assumed that the distribution of time-to failure depends on some covariates, hence we have to select a proper regression model of hazard rate. As it has been already said, we shall consider three basic types of regression models, namely the additive Aalen's model, the proportional hazard or Cox's regression model and the accelerated failure time (AFT) model. More details about regression models in reliability and survival analysis can be found in many monographs, let us mention here again [3,10].

4.1. Additive regression model

In the additive (also Aalen's) model, the hazard function is specified as $h(t, z) = z' \cdot \beta(t)$, where z represents the values of covariates, $\beta(t)$ are functions of time, both z and β are p-dimensional. Their domains should ensure that $h(t, z) \geq 0$. As a rule, the first covariate component is taken fixed to 1, so that $\beta_1(t)$ has the meaning of a 'baseline' hazard function. In the sequel, by index i, $i = 1, .., n$, we shall denote individual objects, while by k, $k = 1, .., p$, components of vectors β, z.

The covariates $Z_i(t)$ are different for each object and can change in time. Individual intensity of $N_i(t)$ is then

$$a_i(t) = Z_i(t)' \cdot \beta(t) \cdot Y_i(t), \quad i = 1, ..., n.$$

Cumulated functions $B_k(t) = \int_0^t \beta_k(s)\, ds$ are estimated by a weighted least squares method. As $dN_i(t) = X_i(t)' dB(t) + dM_i(t)$, where $X_i(t) = Z_i(t) \cdot Y_i(t)$

$$\hat{B}(t) = \int_0^t (X(r)' W(r) X(r))^{-1} X(r)' W(r)\, dN(r),$$

where $W(r)$ is a matrix of weights; the simplest choice is $W(r) = I_n$, an identity matrix, optimal weights are $W(r) = diag\{1/a_i(r)\}$, in practice, $\hat{a}_i(r)$ are used, computation is iterated.

Consistency and asymptotic normality of $\hat{B}(t)$ are straightforward, it holds that the term

$$\sqrt{n}(\hat{B}(t) - B(t)) = \sqrt{n} \int_0^t \overline{X}(r)\, dM(r),$$

where $\overline{X}(r) = (X(r)' W(r) X(r))^{-1} X(r)' W(r)$ is asymptotically distributed as a Gauss process with independent increments (Brown motion process), its covariance function is estimable by an empirical version of

$$n \int_0^t \overline{X}(s) D(s, B(s)) \overline{X}'\, ds,$$

where $D(s, B(s))$ is a diagonal matrix with components $a_i(s)$.

It follows that the case is similar to the case of non-parametrized hazard rate treated in the preceding part. Therefore it is possible to derive tractable asymptotic distribution of residuals. It is described in detail in Volf [15]. Then, the Bayes residual analysis can follow the same scheme as in the preceding section, each function $\beta_k(t)$ has to be modelled separately, again for instance by the approach of Arjas and Gasbarra [2].

5. Cox's regression model

The case differs in certain aspects from the preceding one, which is caused by more complicated asymptotic properties of estimates. The hazard rate is specified as $h(t, z) = h_0(t) \cdot \exp(z' \cdot \beta)$, with processes of covariates $Z_i(t)$ and parameter β (both p-dimensional), $h_0(t)$ is a baseline hazard rate, a nonnegative function. The intensity of i-th process $N_i(t)$ is then

$$a_i(t) = h(t, Z_i(t)) \cdot Y_i(t).$$

Parameter β is estimated from the so-called partial log-likelihood

$$L_p = \sum_{i=1}^n \int_0^T \log \left\{ \frac{\exp(Z_i(t)'\beta)}{\sum_{k=1}^n \exp(Z_k(t)'\beta) \cdot Y_k(t)} \right\} dN_i(t),$$

by an iterative procedure (of Newton–Raphson, as a rule), cumulated baseline hazard $H_0(t) = \int_0^t h_0(r)\, dr$ is then estimated as

$$\hat{H}_0(t) = \int_0^t \frac{dN(r)}{\sum_{k=1}^n \exp(Z_k(t)'\hat{\beta}) \cdot Y_k(t)}.$$

Theory on the properties of estimates is collected elsewhere [3,10]. Estimates are consistent, asymptotically normal, however, neither $\sqrt{n}(\hat{H}_0(t)-H_0(t))$ nor residual process is the martingale.

5.1. Residuals in Cox's model

Residuals are sometimes formulated more generally, as

$$dR(t)=\sum_{i=1}^{n}K_i(t)\cdot(dN_i(t)-d\hat{A}_i(t)),$$

with some (convenient) 'weight' processes $K_i(t)$, for instance if $K_i(t)=Z_i(t)$ (p-dimensional), $R(t)$ is then the estimated score process (the first derivative) of L_p, while $K_i(t)=1[i\in S]$ yields stratified residuals. Stratified residuals (the simplest case) are then expressed as

$$dR_S(t)=dM_S(t)+dH_0(t)C_S(\beta_0,t)-d\hat{H}_0(t)\cdot C_S(\hat{\beta},t),$$

where

$$d\hat{H}_0(t)=\frac{dN(t)}{C(\hat{\beta},t)},$$
$$C_S(\beta,t)=\sum_{i\in S}\exp(Z_i(t)'\beta)\cdot Y_i(t),$$
$$C(\beta,t)=\sum_{i=1}^{n}\exp(Z_i(t)'\beta)\cdot Y_i(t).$$

If we take approximately $\hat{\beta}\sim\beta_0$, we obtain expression similar to cases without regression. The exact approach uses the Taylor expansion of the last term at β_0. $R_S(t)/\sqrt{n}$ is then expressed with the aid of a martingale and a nonrandom function, with asymptotic distribution of a Gauss process, however with rather complicated covariance structure (compare [4, Chapter 12]). Hence, random generation of would-be residual processes under the hypothesis of model fit is possible, but not easy. It is actually based on a bootstrapping, by which we obtain a sample of 'ideal' residual processes. Then, certain characteristics of generated residuals are compared with the same characteristics obtained from the data. That is why practical tests of Cox's model fit are often performed just graphically, comparing visually how far are residuals in group S from zero line, or, equivalently, $\hat{A}_S(t)$ from $N_S(t)$, as in Arjas [1]. Thus, it seems that in the case of Cox's model the Bayes analysis could offer an easiest tool for the model fit assessing.

5.2. Bayes procedure in Cox's model

The procedure of Bayes analysis in the Cox model setting consists of two steps, similarly as a standard estimation method. First, the samples representing posterior distributions of β are obtained with the aid of the Metropolis–Hastings algorithm. In this framework, values of β are proposed from a prior (for instance from a sufficiently wide uniform distribution) and accepted or rejected with the use of partial likelihoods proportion. Then, to each β, a representation of $h_0(t)$ is generated, similarly as in preceding parts, i.e. from a piecewise constant prior. In such a way, a sample of both $\beta^{(m)}$ and $h_0^{(m)}(t)$, $m=1,..,M$, are obtained, from them the intensities and residuals (in a group S, say) can be derived,

$$A_S^{(m)}(t)=\int_0^t h_0^{(m)}(r)\sum_{i\in S}\exp(Z_i(r)\beta^{(m)})\cdot Y_i(r))\,dr,$$

$$R_S^{(m)}(t)=A_S^{(m)}(t)-N_S(t),$$

and used for assessing the model fit.

6. AFT regression model

The accelerated failure time model is often considered as an alternative to Cox's proportional hazard model, when the proportionality of hazards does not hold, cf. Newby [12]. The model assumes that individual speed of ageing is changed by a factor depending on covariates. Quite commonly this factor has the form $\exp(\alpha' z)$, where z is a covariate vector, constant in time. It follows that the distribution of time to failure T_i of an object with covariate value z has the distribution function $F(t)=F_0(t\cdot e^{\alpha' z})$, where F_0 characterizes a baseline distribution (of a random variable T_0 with covariate $z=0$). It also means that $T_{0i}=T_i\cdot\exp(\alpha' z_i)$ is an i.i.d. representation of T_0. Logarithmic transformation yields that

$$\log T_i=-\alpha'\cdot z_i+\log T_{0i}. \tag{1}$$

Statistical inference based on (1) has to deal with unknown distribution of $\log(T_0)$, analysis is not straightforward and could be complicated further by the presence of censored data. Therefore we shall prefer here the approach based on hazard rates, similarly as in the case of Cox's model. Namely, let $\{T_i,z_i,d_i,\ i=1,..,n\}$ be observed times of failures or censoring of i-th object, their covariates, indicators of censoring, respectively, then the likelihood reads

$$L=\prod_{i=1}^{n}h_i(T_i)^{d_i}\cdot\exp\left(-\int_0^{T_i}h_i(t)\,dt\right), \tag{2}$$

where $h_i(t)=h_0(t\cdot\exp(\alpha' z_i))\cdot\exp(\alpha' z_i)$ is the hazard rate of i-th object at time t, h_0 is the baseline hazard rate of T_0. Theory of estimation and asymptotic properties are derived in Lin et al. [11] and further also developed in Bagdonavicius and Nikulin [4, Chapter 6]. Nevertheless, as the practical computation of asymptotic characteristics is rather complicated, the Bayes approach can again offer a reasonable alternative.

Lin et al. (1993) have showed that instead of an exact score function for α (i.e. obtained by derivation of log-likelihood (2)), it is possible to use approximate score functions. Namely, the score function has the form

$$U(\alpha)=\sum_{i=1}^{n}\left\langle z_iW_i(s)-\frac{\sum_j z_jW_j(s)Y_j^*(T_{0i})}{\sum_k Y_k^*(T_{0i})}\right\rangle d_i, \tag{3}$$

where $Y_k^*(t)$ are indicators of risk in the scale of $T_{0i}=T_i\cdot\exp(\alpha z_i)$. While exact weights $W_i(s)$ depend actually also on $h_0(s),h_0'(s)$, they may be substituted by a set of simpler functions, among them also by $W_I(t)=1$ for all i and t. Lin et al. [11] have proved that the corresponding estimator of α retains good asymptotic properties. Hence, from such a score function it is possible to estimate α without knowledge of $h_0(t)$, similarly like with the aid of partial likelihood in Cox's model case. Recently, Novák [13] has proposed a method of goodness-of-fit test based on the random generation of residual processes, and has studied the test behavior in various situations.

6.1. Bayes analysis in the AFT model

Similarly as in Cox's model case, we employ the MCMC procedure consisting of two steps. The first step generates α in order to minimize $|U(\alpha)|$ from (3) with $W_I(t)=1$. New α is proposed from a prior, the Metropolis–Hastings algorithm uses the acceptance rule based on the proportion of $\exp(-|U(\alpha)|)$ with new and current α–s. Then, in the second step, to each $\alpha^{(m)}$ obtained in Step 1, a representation of baseline hazard rate $h_0(t)$ is updated from the likelihood

$$L=\prod_{i=1}^{n}[h_0(T_{0i})\cdot e^{\alpha z_i}]^{d_i}\cdot\exp\left(-\int_0^{T_{0i}}h_0(t)\,dt\right),$$

where $T_{0i} = T_i \cdot \exp(\alpha z_i)$, T_i are observed times. Again, the method of Arjas and Gasbarra [2] can be utilized.

In such a way, the MCMC procedure yields a sequence of estimates $\alpha^{(m)}, h_0(t)^{(m)}$. Corresponding m-th estimate of the intensity of failure for i-th object equals $a_i^{(m)}(t) = h_0^{(m)}(t \cdot \exp(\alpha^{(m)} z_i)) \cdot \exp(\alpha^{(m)} z_i)$ on $[0, T_i]$ and equals zero for $t > T_i$. Cumulated intensities are then $A_i^{(m)}(t) = \int_0^t a_i^{(m)}(s)\, ds$, and, finally, residual processes in a subset $S \subset \{1, 2, \ldots, n\}$ are again the differences

$$R_S^{(m)}(t) = \sum_{i \in S} [A_i^{(m)}(t) - N_i(t)], \tag{4}$$

where $N_i(t)$ is the counting process of failure of i-th object, i.e. with maximally one step $+1$ at T_i provided $d_i = 1$.

7. Examples

In order to illustrate the performance of proposed approach, we shall present two examples. The first analyzes randomly generated data, while the second deals with real data study.

7.1. Artificial example

A sample of $n = 100$ data was generated randomly from the following AFT model: Baseline distribution of $\log T_0$ followed normal distribution with $\mu = 0, \sigma = 0.5$, values of covariate Z were distributed uniformly in $(0, 2)$, the corresponding accelerating parameter was set to $\alpha = 1$. Further, values $T_i^* = T_{0i} \exp(-\alpha' z_i)$ were randomly right-censored by i.i.d. variables distributed uniformly in $(\min T_i^*, \max T_i^*)$. Final censored data (in log scale) are displayed in Fig. 1.

Then, data were analyzed in the framework of both AFT and Cox's models, by the Bayes approach described in preceding parts. In the AFT setting, 5000 MCMC iterations of α were performed, last 2000 were used for the analysis. Posterior representation of α had the mean 1.0105 and the standard deviation 0.0309. Then, to each $\alpha^{(m)}$ 200 instances of $H_0(t)$ were generated, we always took just the last of them. In such a way, $M = 2000$ 'models' were obtained representing the posterior distribution of the AFT model. Fig. 2 shows the characteristics of corresponding posterior sample of residual processes, namely their point-wise medians and then also approximate 95% credibility bands, i.e. such a region that approximately 95% of residual processes lie fully inside. The bands are dashed, they are plotted against counts $N_S(t)$, separately for two groups $Z < 1, Z > 1$. It is seen that graphs are concentrated around zero, thus assessing good AFT model fit. Several trajectories of residual processes are displayed by dots.

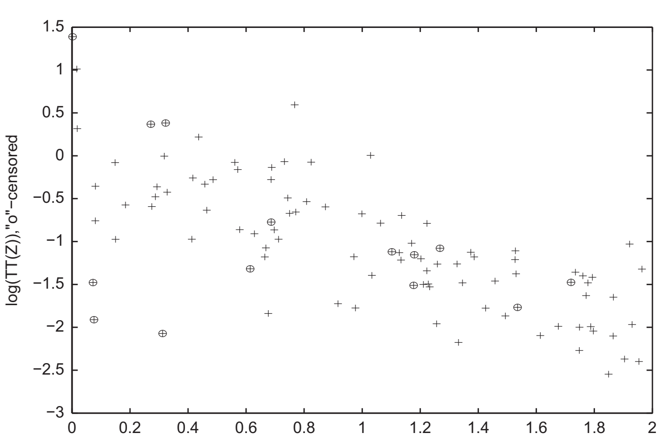

Fig. 1. Data of Example 7.1: Covariate is on the x-axis, log of survival on y axis, censored items are denoted by 'o'.

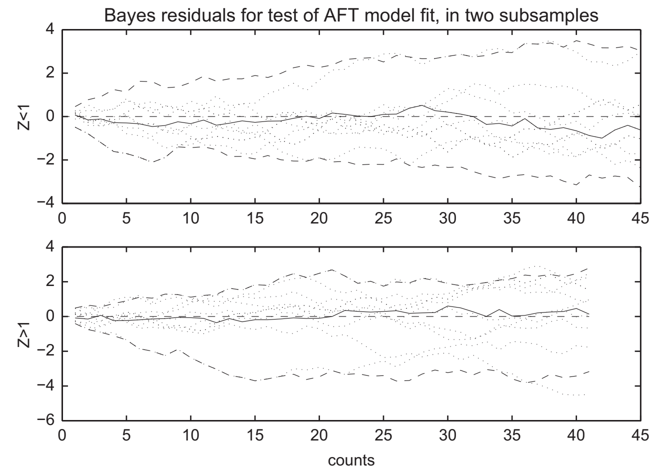

Fig. 2. Characteristics of residuals in the AFT model, in two subgroups with $Z < 1$ and $Z > 1$.

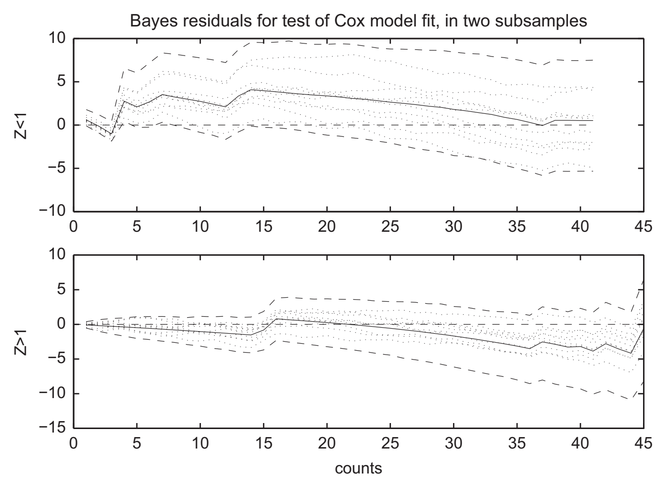

Fig. 3. Characteristics of residuals in the Cox model, in two subgroups with $Z < 1$ and $Z > 1$.

The same data (i.e. generated in the AFT model) were then analyzed in Cox's model framework, following the procedure of Section 4.2. Again, 5000 β-s were generated, and last 2000 were taken as a representation of posterior distribution of β. It had the mean 2.3665 and standard deviation 0.2688. Further, to each $\beta^{(m)}$ 200 instances of $H_0(t)$ were generated, we took the last of them. Thus, a representation with $M = 2000$ members was obtained. Fig. 3 shows again the characteristics of posterior sample of residual processes, i.e. their point-wise medians and approximate 95% credibility region, for two groups with $Z < 1, Z > 1$. Similarly as above, several trajectories of residual processes are displayed by dots. Departures of sample of residuals from the zero level is now rather significant, especially for low times in the first group. Graphs indicate that Cox's model overestimates the failure intensity for small covariate values and also underestimate it for larger covariate values.

In Cox's model setting, standard analysis was performed, too. It yielded the estimate $\beta = 2.3687$ with asymptotic standard deviation 0.2786.

7.2. Real data example

Data collected in Table 1 have their origin in accelerated testing of resistance of certain steel parts made for suspension system of

Table 1
Data: Load in N/cm^2, survival in cycles of vibration, indicator of censoring.

Load	Survival	Cens.
45.4	41.2	1
35.3	202.4	1
27.75	919.3	1
25.0	1036.2	1
22.7	2020.0	1
20.2	10,064.8	0
35.6	470.1	1
27.7	620.0	1
27.75	2119.9	1
25	11,390.0	0
22.7	2065.9	1
17.66	9219.0	0
35.6	865.8	1
27.75	884.9	1
27.5	1998.9	1
25	14,443.9	0
20.2	3231.4	1

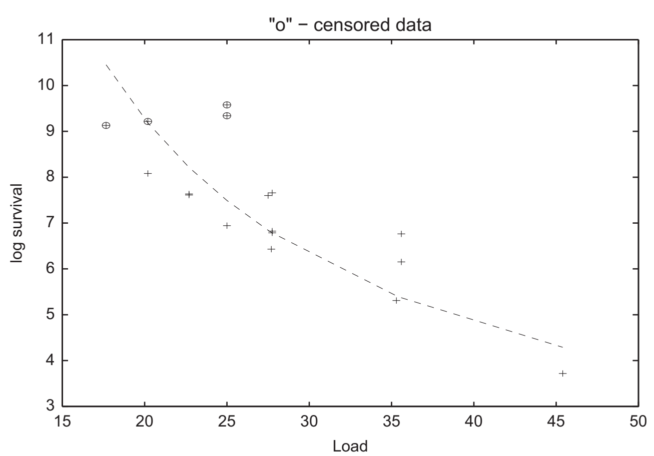

Fig. 4. Data and resulting model of Example 7.2: Load is on the x-axis, log of survival on y axis, censored data are denoted by 'o'.

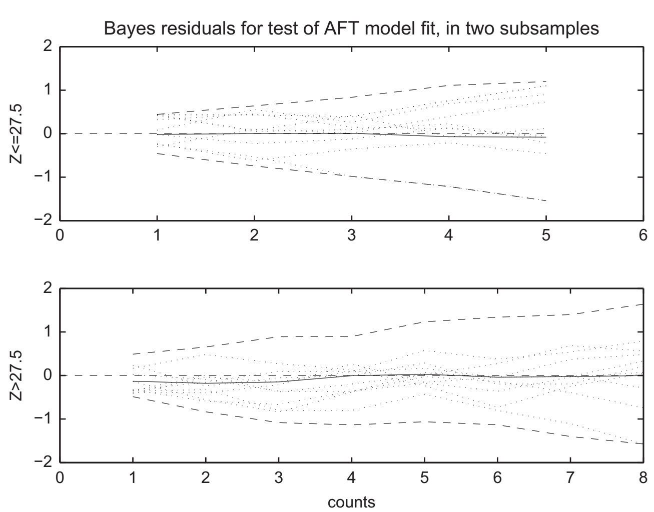

Fig. 5. Characteristics of residuals in two subgroups with $Z \leq 27.5$ and $Z > 27.5$.

trucks Tatra. The parts were tested, in a laboratory, by cycles of high-frequency vibrations under different loads. Load remained constant during each experiment. It is given in N/cm^2, survival in cycles, censoring indicator is 0 when the experiment was terminated before any defect occurred.

The data, with a log scale of survival, are displayed also in Fig. 4. We considered the AFT model, with a log-hyperbolic trend, actually a variant of Arrhenius model, namely

$$\log\ T = -\beta \cdot (C - 1/Z) + \log\ T_0,$$

where Z was the load, $\beta > 0$ was an unknown parameter and we selected $C = 0.1$. It means that T_0 corresponded to load $Z = 10\ \mathrm{N/cm^2}$. There were two reasons for such a choice: The data show that the log-hyperbolic trend is more likely than log-linear, the shift by C gives a reasonable meaning to T_0. Naturally, $Z = 10$ is here just a reference value, during real use both frequency and load of vibrations vary.

The method of solution followed the Bayes approach described in Section 6.1. In the MCMC procedure, 5000 instances of β-s were generated, last 2000 considered as representation of posterior distribution of β, with the mean 178.34, median 178.12 and standard deviation 0.70. Fig. 4 also contains the median curve med $\langle -\beta^{(m)} \cdot (C - 1/Z_i) + \log\ T_{0,i}^{(m)} \rangle$, $m = 1, 2, \ldots, 2000$. The graphical result of the goodness-of-fit test is displayed in Fig. 5. The figure again contains point-wise medians of generated residual processes, in two subgroups (full curves), and also approximate 95% credibility bands (dashed). Some of residual processes are displayed by dots. As the graphs are concentrated around zero, we can conclude that the selected model fits well.

8. Conclusion

In the field of statistical reliability analysis, the models for lifetime often have to reflect a dependence on covariates, for instance, on the load, degradation or conditions of usage. The present paper was devoted to the method of corresponding regression models selection, with the aid of the goodness-of-fit based on martingale residuals. While in the case of Aalen's additive regression the residual process retains the martingale property and therefore the critical region for the test can be derived easily. However, it is not the case of Cox's and AFT regression models. That is why the main objective was to propose the Bayes variant of the martingale residuals construction and to derive a test procedure based on them. It was shown that in the cases of mentioned models the Bayes approach offers a reasonable alternative to standard methods.

Acknowledgement

The research was supported by the Project of MŠMT ČR No. 1M06047, The Center for Quality and Reliability.

References

[1] Arjas E. A graphical method for assessing goodness of fit in Cox's proportional hazard model. J Am Stat Assoc 1988;83:204–12.
[2] Arjas E, Gasbarra D. Nonparametric Bayesian inference from right censored survival data using Gibbs sampler. Stat Sin 1994;4:505–24.
[3] Andersen PK, Borgan O, Gill RD, Keiding N. Statistical models based on counting processes. New York: Springer; 1993.
[4] Bagdonavicius V, Nikulin M. Accelerated life models, modeling and statistical analysis. Boca Raton: Chapman& Hall/CRC; 2002.
[5] Elsayed EA, Zhang H. Design of PH-based accelerated life testing plans under multiple-stress-type. Reliab Eng Syst Saf 2007;92:286–92.
[6] Erto P, Giorgio M. Assessing high reliability via Bayesian approach and accelerated tests. Reliab Eng Syst Saf 2002;76:301–10.
[7] Gamerman D. Markov Chain Monte Carlo: stochastic simulation for Bayesian inference. Boca Raton: Chapman & Hall/CRC; 1997.
[8] Gelfand AE, Kottas A. Bayesian semiparametric regression for median residual life. Scand J Stat 2003;30:651–65.

[9] Hoyland A, Rausand M. System reliability theory: models and statistical methods. New York: Wiley; 1994.
[10] Kalbfleisch JD, Prentice RL. The statistical analysis of failure time data. New York: Wiley; 2002.
[11] Lin DY, Wei LJ, Ying Z. Accelerated failure time models for counting processes. Biometrika 1998;85:605–18.
[12] Newby M. Accelerated failure time models for reliability data analysis. Reliab Eng Syst Saf 1988;20:187–97.
[13] Novák P. Goodness-of-fit test for the accelerated failure time model based on martingale residuals. Kybernetika 2013;49:40–59.
[14] Van Dorp JR, Mazzuchi TA. A general Bayes Weibull inference model for accelerated life testing. Reliab Eng Syst Saf 2005;90:140–7.
[15] Volf P. Analysis of generalized residuals in hazard regression models. Kybernetika 1996;32:501–10.
[16] Wang L, Pan R, Li X, Jiang T. A Bayesian reliability evaluation method with integrated accelerated degradation testing and field information. Reliab Eng Syst Saf 2013;112:38–47.
[17] Zuo Y, Serfling R. General notion of statistical depth function. Ann Stat 2000;28:461–82.