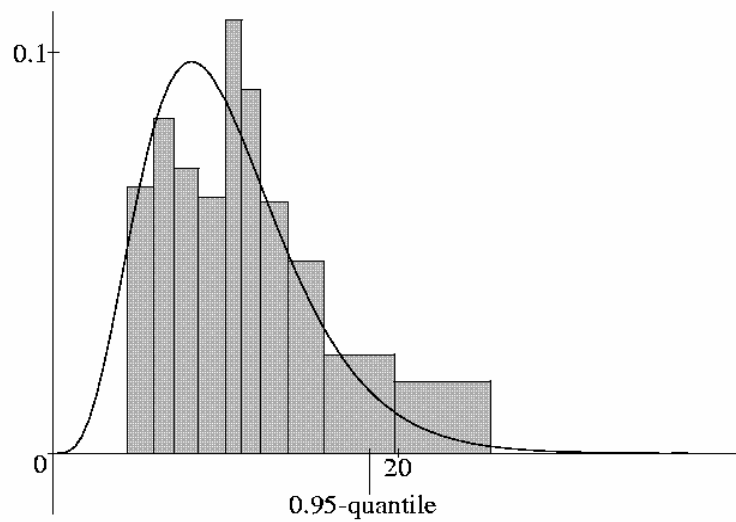




UNIVERSITÄT POTSDAM
Institut für Mathematik

Nonparametric Estimation and Testing in Survival Models

Henning Läter and Hannelore Liero



Mathematische Statistik und
Wahrscheinlichkeitstheorie

Universität Potsdam – Institut für Mathematik

Mathematische Statistik und Wahrscheinlichkeitstheorie

Nonparametric Estimation and Testing in Survival
Models

Henning Lauter

e-mail: laeuter@rz.uni-potsdam.de

Hannelore Liero

e-mail: liero@rz.uni-potsdam.de

Department of Mathematics, University of Potsdam, 14469 Potsdam, Germany

Preprint 2004/05

August 2004

Impressum

© Institut für Mathematik Potsdam, August 2004

Herausgeber: Mathematische Statistik und Wahrscheinlichkeitstheorie
am Institut für Mathematik

Adresse: Universität Potsdam
PF 60 15 53
14415 Potsdam

Telefon:
Fax: +49-331-977 1500
E-mail: +49-331-977 1578
neisse@math.uni-potsdam.de

ISSN 1613-3307

Nonparametric Estimation and Testing in Survival Models

Henning Lauter and Hannelore Liero
Institute of Mathematics, University of Potsdam
e-mail: laeuter@rz.uni-potsdam.de, liero@rz.uni-potsdam.de

The aim of this paper is to demonstrate that nonparametric smoothing methods for estimating functions can be an useful tool in the analysis of life time data. After stating some basic notations we will present a data example. Applying standard parametric methods to these data we will see that this approach fails - basic features of the underlying functions are not reflected by their estimates. Our proposal is to use nonparametric estimation methods. These methods are explained in section 2. Nonparametric approaches are better in the sense that they are more flexible, and misspecifications of the model are avoided. But, parametric models have the advantage that the parameters can be interpreted. So, finally, we will formulate a test procedure to check whether a parametric or a nonparametric model is appropriate.

1 Stating the Problem

We consider life or failure times of individuals or objects belonging to a certain group, the so-called population of interest. Examples are: survival times of patients in a clinical trial, lifetimes of machine components in industrial reliability or times taken by subjects to complete specified tasks in psychological tests. We assume that these life times can be modelled by a random variable Y with a distribution F , that is, we assume that the probability that an individual of the underlying population dies (fails) before time point t can be expressed in the form

$$P(Y \leq t) = F(t).$$

The probability that the individual survives the time point t is given by the survival function

$$S(t) = P(Y > t) = 1 - F(t).$$

Other functions of interest are the density $f(t) = F'(t)$ and the hazard or failure rate

$$\lambda(t) = \lim_{s \downarrow 0} \frac{1}{s} P(t < Y \leq t + s | Y \geq t)$$

describing the immediate risk attaching to an individual known to be alive at time point t .

Now, suppose that we have obtained data from the underlying population. How we can use these data to estimate the survival function or the hazard rate?

Assuming a parametric model for the distribution the survival times we have to estimate parameters. It is well-known, that the maximum likelihood method provides good estimates.

For example, if we assume that our data are realizations of exponential distributed random variables Y_1, \dots, Y_n , that is, the survival function is given by

$$S(t) = \exp(-t\beta),$$

with parameter $\beta > 0$, then the problem of estimating the function S is simply the problem of estimating the parameter β . And the maximum likelihood estimator (m.l.e.) is given by

$$\hat{\beta} = \frac{1}{n} \sum_{i=1}^n Y_i.$$

Assuming a Weibull distribution with parameters β and ν , i.e

$$S(t) = P(Y > t) = \exp(-(t/\beta)^\nu),$$

we obtain that the m.l.e. of the two-dimensional parameter is a solution of

$$\begin{aligned} \hat{\beta}^{\hat{\nu}} &= \frac{1}{n} \sum_{i=1}^n Y_i^{\hat{\nu}} \\ \frac{\sum_{i=1}^n Y_i^{\hat{\nu}} \log Y_i}{\hat{\beta}^{\hat{\nu}}} &= \frac{n}{\hat{\nu}} + \sum_{i=1}^n \log Y_i. \end{aligned} \quad (1)$$

If the assumed parametric model is a good description of the of the underlying population, then parametric estimators and test procedures based on these estimators provide good results. But if the parametric model is not appropriate such an approach can lead to wrong conclusions. This is demonstrated in the following:

Suppose that a mixture of two Weibull distributions is considered. The first group is characterized by parameters β_1, ν_1 and the second with β_2, ν_2 , and let p be the portion of the first group. Then the survival function is given by

$$S^*(t) = (1 - p) \exp(-(t/\beta_1)^{\nu_1}) + p \exp(-(t/\beta_2)^{\nu_2}) \quad (2)$$

For $\beta_1 = 1, \beta_2 = 4, \nu_1 = 2, \nu_2 = 4$ and $p = 0.05$ the figures show S^* , the density f^* and the hazard rate λ^* of the mixture (solid line). Further the main part of the mixture, i.e. $\exp(-(t/\beta_1)^{\nu_1})$ is given in Figure (a). In (b) and (c) you see not only this term of the mixture but also the minor one.

In such a case with a small p one can interpret the first Weibull distribution as a disturbance of the second one and one would hope that the fit with a single Weibull distribution is sufficiently well. Simulated data with 100 observations

from the disturbed Weibull model were used to estimate the parameters β and ν in a single Weibull model with

$$S(t) = \exp(-(t/\beta)^\nu) \quad \text{and} \quad \lambda(t) = \frac{t^{\nu-1}}{\beta^\nu},$$

which was assumed neglecting the inhomogeneity of the population.

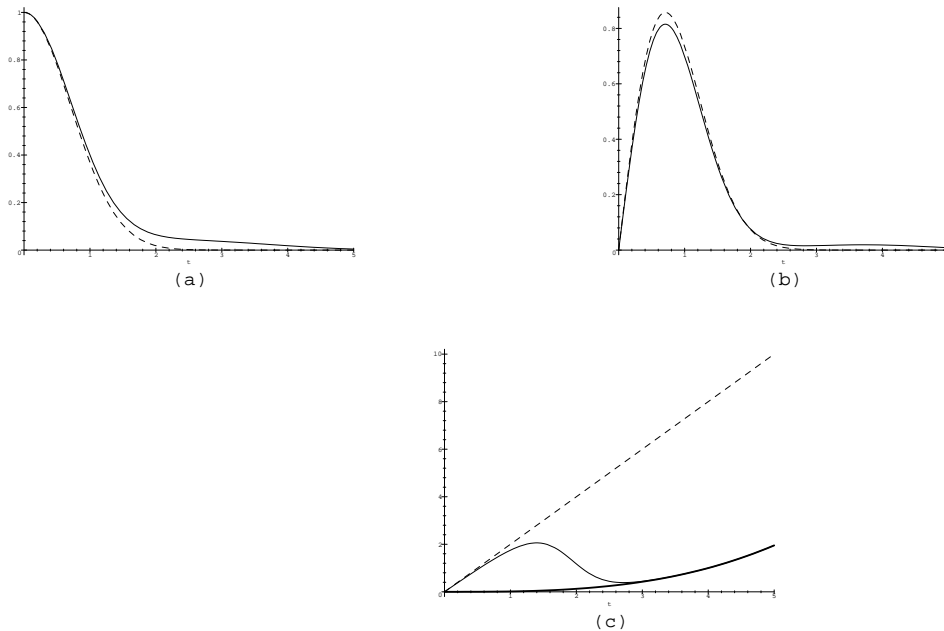
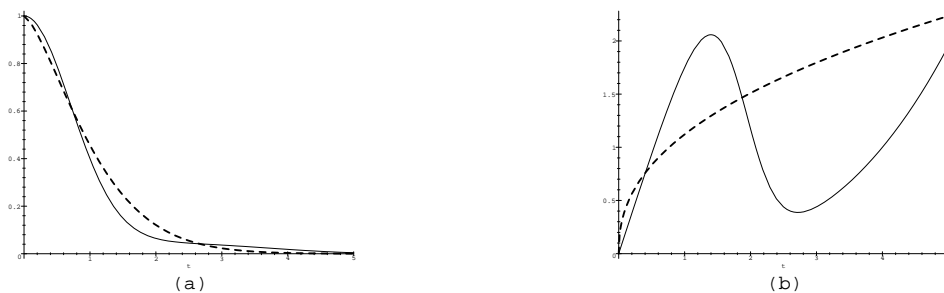


Figure 1 (a) Survival functions, (b) Densities, (c) Hazard rates, for the main component (dashed line), for the mixture (thin solid line), in (c) the hazard rate for the minor component (bold solid line)

The maximum likelihood estimates, computed according to (1), are: $\hat{\beta} = 1.057$ and $\hat{\nu} = 1.422$. Replacing these estimates into the functions S and λ we get Figure 2.



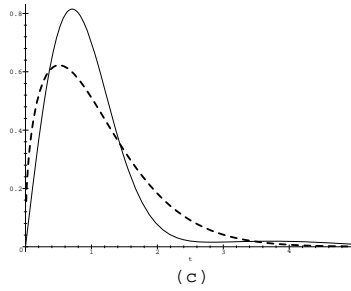


Figure 2 (a) Survival functions, (b) Hazard rates, (c) Densities, for the estimated single Weibull model (bold dashed line), for the mixture (thin solid line)

We see: The estimators using the single Weibull model are wrong estimators. This model is unable to detect the features of the underlying functions! Such a mixed distribution one meets if the underlying population is not homogenous. A latent factor, which is not observed divides the population into (for simplicity) two groups. Further, assume that both groups can be characterized by a Weibull distribution: the first with parameters β_1, ν_1 and the second with β_2, ν_2 , and let p be the portion of the first group. Latent factors can be: a not observed underlying disease (depression), different litter in an animal experiment or different producer of a technical component.

2 Nonparametric Estimators

2.1 Model with censoring

Very often, in practical applications the life times Y_i 's are subject to random right censoring, i.e. some individuals may not be observed for the full time to failure. Thus, our observations are values of r. v.'s T_i which are censored or uncensored. Here we assume a random censoring scheme characterized by i.i.d. r. v.'s C_i which are independent of the Y -sequence. Thus, we observe (T_i, δ_i) , $i = 1, \dots, n$ with

$$T_i = \min(Y_i, C_i) \quad \text{and} \quad \delta_i = 1(Y_i \leq C_i).$$

The distribution of the observations is described by the distribution function and the subdistribution function of the uncensored observations

$$H(t) := P(T_i \leq t) \quad \text{and} \quad H^U(t) := P(T_i \leq t, \delta_i = 1).$$

2.2 The Nelson-Aalen estimator for the cumulative hazard function

Starting point of the construction of an estimator for the hazard function λ and the survival function S is an estimator for Λ , the cumulative hazard function defined by

$$\Lambda(t) = \int_0^t \lambda(s) ds.$$

Using standard transformations we can write this estimator in the following form

$$\Lambda(t) = \int_0^t \frac{dF(y)}{S(y)} = \int_0^t \frac{dH^U(y)}{1-H(y)}. \quad (3)$$

The idea for the estimation of Λ goes back to Beran (1981). He proposed to replace the functions H and H^U in (3) by their empirical versions

$$\hat{H}_n^U(t) = \sum_{i=1}^n \mathbf{1}(T_i \leq t, \delta_i = 1), \quad \hat{H}_n(t) = \sum_{i=1}^n \mathbf{1}(T_i \leq t). \quad (4)$$

The resulting estimator is the so-called Nelson-Aalen type estimator

$$\hat{\Lambda}_n(t) := \int_0^t \frac{d\hat{H}_n^U(s)}{1-\hat{H}_n(s-)}.$$

The explicit formula of $\hat{\Lambda}_n$ is given by

$$\hat{\Lambda}_n(t) = \sum_{i=1}^n \frac{\mathbf{1}(T_{(i)} \leq t) \delta_{[i]}}{n-i+1}.$$

Here $T_{(1)} \leq \dots \leq T_{(n)}$ is the order statistic, and $\delta_{[i]} = \delta_j$ if $T_j = T_{(i)}$.

From this estimator we get the well-known Kaplan-Meier product limit estimator by the transformation

$$\hat{F}_n(t) = 1 - \exp(-\hat{\Lambda}_n(t)).$$

Asymptotic properties of these estimators were investigated by several authors, for example by Horvath(1981), Lo/Singh(1986) and Major/Rejtó (1988).

2.3 A kernel estimator for the hazard function

The hazard function λ is the derivative of the cumulative hazard Λ . But the estimator $\hat{\Lambda}_n$ is not differentiable. So, we follow the same line as in the case of nonparametric density estimation. Let us estimate λ at point t . Consider a small interval $[t-b, t+b)$ of length $2b$ around t . We can approximate $\lambda(t)$ in the following way:

$$\lambda(t) \sim \frac{\int_{t-b}^{t+b} \lambda(s) ds}{2b} = \frac{\Lambda(t+b) - \Lambda(t-b)}{2b} \sim \frac{\hat{\Lambda}_n(t+b) - \hat{\Lambda}_n(t-b)}{2b}. \quad (5)$$

The last term in (5) can be written in the form

$$\frac{1}{b} \sum_{i=1}^n K^* \left(\frac{t - T_{(i)}}{b} \right) \frac{\delta_{(i)}}{n-i+1},$$

where

$$K^*(u) = \begin{cases} \frac{1}{2} & \text{for } -1 \leq u \leq 1 \\ 0 & \text{otherwise} \end{cases}.$$

The first approximation step in (5) yields a systematic error, which becomes small if the length of the interval is small. At the other hand, if b is small, then the second approximation error, the stochastic error, is large, because we have not enough observations for stability. To take these tendencies into account, we have to choose b depending on the sample size n , $b = b_n$, such that

$$b_n \rightarrow 0 \quad \text{and} \quad nb_n \rightarrow \infty. \quad (6)$$

Further, it is useful to take instead of the function K^* a more general function K , a function giving small weights to observations $T_{(i)}$ far away from the point t and large weights to observations very near to the point, at which we estimate. This is realized, for example, by taking a symmetric density function for K . So, finally we arrive at the following definition:

$$\hat{\lambda}_n(t) = \frac{1}{b_n} \sum_{i=1}^n K\left(\frac{t - T_{(i)}}{b_n}\right) \frac{\delta_{(i)}}{n - i + 1}. \quad (7)$$

Here $K : \mathbb{R} \rightarrow \mathbb{R}$ is the kernel function and $\{b_n\}$ the sequence of bandwidths satisfying (6). The estimator (7) can be written shortly as

$$\hat{\lambda}_n(t) = \frac{1}{b_n} \int K\left(\frac{t - s}{b_n}\right) d\hat{\Lambda}_n(s).$$

Several properties of this estimator are known. Let us mention here papers of Singpurwalla and Wong (1983), Tanner and Wong (1983) and the results of Diehl and Stute (1986). In these papers conditions for consistency are derived and asymptotic expressions for the bias and the variance are given. Diehl and Stute considered an approximation for the difference between the estimator $\hat{\lambda}_n$ and a smoothed hazard rate by a sum of i.i.d. r.v.'s. On the basis of such a representation limit theorems can be derived.

The following picture shows a nonparametric kernel estimate for the data generated in the simulated model (2). Here the kernel function is the Gaussian kernel, the bandwidth is $b_n = 0.2$. We see, that this estimate reflects the features of the underlying hazard function much better than the parametric estimator.

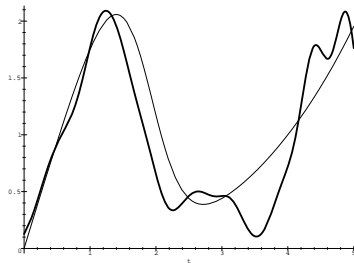


Figure 3 True underlying hazard rate (thin) and nonparametric estimate (bold)

3 Testing the Hazard Rate

Nonparametric estimators of a curve are an appropriate tool in the analysis of data. But, sometimes in practical situations it seems to be useful to have a parametric model. The advantage of a parametric model is that the parameters have a some meaning, very often they can be interpreted. Of course, this holds only, if the chosen parametric model is appropriate. Thus, the question arises, whether the choice of a certain parametric model can be justified by the data. In this section we propose a test procedure for checking whether a hypothetical model fits the data, that is we consider the following hypothesis

$$\mathcal{H}: \lambda \in \mathcal{L} \quad \text{vs.} \quad \mathcal{K}: \lambda \notin \mathcal{L},$$

where \mathcal{L} is the class of parametric hazard functions

$$\mathcal{L} = \{\lambda(t; \vartheta) \mid \vartheta \in \Theta \subset \mathbb{R}^k\}$$

An example for such an parametric class \mathcal{L} is the set of all Weibull hazards. Further parametric models are given in the book of Bagdonavičius, V. and Nikulin, M. (2002).

At the first view one would choose as test statistic the deviation of the nonparametric estimator $\hat{\lambda}_n$, which is a good estimator under the alternative, from a hypothetical hazard with estimated parameter $\hat{\vartheta}$, i.e. from $\lambda(t; \hat{\vartheta})$. Here $\hat{\vartheta}$ is an appropriate estimator of the unknown parameter. But the nonparametric $\hat{\lambda}_n$ is a result of smoothing procedure. Remember formulae (5) - it is an unbiased estimator of

$$\frac{1}{b_n} \int K\left(\frac{t-z}{b_n}\right) \lambda(z) dz,$$

and not unbiased for the underlying hazard rate. So, it seems to be natural to compare $\hat{\lambda}_n$, which smoothes the data, with a smoothed version of the hypothesis. Thus, we will take the difference between $\hat{\lambda}_n$ and $\tilde{\lambda}_n$ defined by

$$\tilde{\lambda}_n(t; \hat{\vartheta}) = \frac{1}{b_n} \int K\left(\frac{t-z}{b_n}\right) \lambda(z; \hat{\vartheta}) dz.$$

Generally speaking, one can take as deviation measures L_p -distances for functions. Here we will consider a L_2 -type distance, namely

$$Q_n = \int (\hat{\lambda}_n(t) - \tilde{\lambda}_n(t; \hat{\vartheta}))^2 a(t) dt$$

The function a is a known weight function, it is introduced to control the region of integration.

3.1 An asymptotic α -test

To formulate a test based on this statistic we have to derive the distribution of Q_n , or at least the limiting distribution under the hypothesis. The theory about the asymptotical distributional behavior of quadratic forms yields the following limit statement. Under

- regularity conditions on the kernel K and the bandwidth b_n ,
- smoothness of the functions H and H^U and
- conditions ensuring that the estimator $\hat{\vartheta}_n$ is \sqrt{n} -consistent

the distribution of the standardized Q_n converges to the standard normal distribution, that is

$$\frac{nb_n^{1/2}}{\sigma} (Q_n - \mu_n) \xrightarrow{\mathcal{D}} \mathbf{N}(0, 1)$$

where

$$\begin{aligned} \mu_n &= (nb_n)^{-1} \kappa_1 \int \frac{\lambda(t; \hat{\vartheta}_n)}{1 - H(t)} a(t) dt \\ \sigma^2 &= 2 \kappa_2 \int \left(\frac{\lambda(t; \hat{\vartheta}_n)}{1 - H(t)} \right)^2 a^2(t) dt \end{aligned} \quad (8)$$

with $\kappa_1 = \int K^2(x) dx$ and $\kappa_2 = \int (K * K)^2(x) dx$ and "*" denotes the convolution.

The only unknown term in this limit statement is the distribution H of the observations. Replacing this by the empirical distribution \hat{H}_n we obtain the following asymptotic α -test: Reject \mathcal{H} , iff

$$Q_n \geq \frac{z_\alpha \hat{\sigma}_n}{nb_n^{1/2}} + \hat{\mu}_n. \quad (9)$$

Here z_α is the $(1 - \alpha)$ -quantile of the standard normal distribution and $\hat{\mu}_n$ and $\hat{\sigma}_n^2$ are defined as in (8), where H is replaced by \hat{H}_n .

3.2 Application to the example

Now, let us apply the proposed test to the example considered in Section 1. The nonparametric estimator of the hazard rate in the Weibull mixture model and the smoothed hypothetical hazard function, that is a hazard rate in a Weibull model with parameter $\hat{\vartheta} = (1.057, 1.422)$, are given in Figure 4.

We compute the integrated quadratic distance over the interval $[0, 4]$, and get the following values for the test statistic and the standardizing terms

$$\begin{aligned} Q_n &= 2.8161 \\ \hat{\mu}_n &= 1,461 \\ \hat{\sigma}_n^2 &= 1853.717 \end{aligned}$$

With these values the test procedure yields for $\alpha = 0.05$: Reject \mathcal{H} . The p -value is 0.0025.



Figure 4 (a) Densities, (b) Hazard rates. Hypothetical single Weibull model (dashed line), nonparametric estimate (bold solid line), in (b) true underlying mixture model (thin solid line)

3.2.1 Conclusions

1. There are two possible points of view. The first is to consider the minor part of the mixture as a disturbance. That is, one is interested in the main part, for which the parametric model is justified. Then the nonparametric estimate of the hazard rate shows that the population is not homogenous, or in other words, our data are not appropriate for the estimation of both parameters. Further, we see that the hazard rate reflects this deviation much better than the survival function. Hence, in this case the application of a nonparametric estimator for the hazard rate is helpful for detecting outliers.
2. A second point of view is, that one is interested in the distribution of the population, that is the data are correct in the sense, that they represent the population we are interested in. Then our nonparametric approach shows that the chosen parametric model is not appropriate. Thus, the nonparametric estimator can be helpful for stating a better parametric model. Of course a parametric mixture model with unknown parameter p is a complicated matter.
3. In both cases we see that the hazard rate is more sensitive. The deviation of a hazard rate from a hypothetical one, which can be seen very clearly, is smoothed away when we consider the corresponding survival functions.

4 Some further remarks

1. The proposed test is consistent, that is, if the distribution of the data does not belong to the hypothetical class, then the probability that the test rejects the hypothesis tends to one. This is not a very strong property.

So, it seems to be useful to consider the power of the test under so-called local alternatives. For testing a density function nonparametrically such considerations were done in Liero, Läuter and Konakov (1998). The results for the hazard rate are similar. Roughly speaking one obtains, that the test is sensitive against alternatives tending to the hypothetical hazard function at the rate $\sqrt{nb_n^{1/2}}$.

2. The problem of the application of the nonparametric estimator and the test is the choice of the bandwidth b_n . If the bandwidth is chosen large, the systematic error becomes large. At the first view this is not crucial, because we compare the smooth nonparametric estimator $\hat{\lambda}_n$ with the smoothed hypothetical function $\tilde{\lambda}_n$. But the approximation of the distribution of the standardized test statistic Q_n by the normal distribution is worse for large b_n . Simulation results show that in this case the test has the tendency to accept the hypothesis. At the other hand, if b_n is chosen to small, then the resulting estimator is wiggly, and the power of the becomes worse.

5 About the Extension to the Model with Covariates

The approach described above can be generalized to the model with covariates. In applications often we observe in addition to the life times some covariates. These covariates can be e.g. the dosis of a drug, the temperature or other factors of influence. That is, we have observations (T_i, X_i, δ_i) , where X_i is the covariate taking values in \mathbb{R} or more general in \mathbb{R}^k . We can consider these covariates as fixed design points, or as random values. In both cases we are interested in statistical inference about the survival function $S(t|x)$, the density $f(t|x) = -\frac{dS(t|x)}{dt}$ and the hazard function $\lambda(t|x) = \frac{f(t|x)}{S(t|x)}$. Here $S(t|x)$ is the probability that an individuum or item survives the time point t given the covariate takes the value x . We do not want to go into further details, the basic idea is to estimate the distribution functions $H(\cdot|x)$ and $H^U(\cdot|x)$ not by the empirical distribution functions given in (4), but by *weighted* empirical distribution functions

$$\hat{H}_n^U(t) = \sum_{i=1}^n w_{ni}(X, x; h_n) \mathbf{1}(T_i \leq t, \delta_i = 1) \quad \hat{H}_n(t) = \sum_{i=1}^n w_{ni}(X, x; h_n) \mathbf{1}(T_i \leq t).$$

Here, the weights $w_{nj}(X, x)$ depend on the observed covariates $X = (X_1, \dots, X_n)$, on x and on a smoothing parameter h_n . They are chosen such that the T_j gets a large weight in counting all the T_i 's, which are smaller or equal t , if the corresponding covariate X_j is near x . Appropriate weights are kernel weights of Gasser-Müller type for fixed covariates or Nadaraya-Watson kernel weights for random X_i 's. The resulting estimator of the hazard rate has then the following form

$$\hat{\lambda}_n(t|x) = \frac{1}{b_n} \sum_{i=1}^n K\left(\frac{t - T_{(i)}}{b_n}\right) \frac{\delta_{(i)} w_{n(i)}(X, x; h_n)}{1 - \sum_{j=1}^{i-1} w_{n(j)}(X, x; h_n)}.$$

Properties of nonparametric estimators for the hazard rate, the cumulative hazard function and the survival functions for models with covariates are derived, for example, in papers by González-Manteiga and Cadarso-Suarez (1996) and Van Keilegom and Veraverbeke (1997, 2001, 2002).

For testing the hypothesis that $\lambda(\cdot|x)$ is equal to a given hazard function $\lambda^*(\cdot|x)$ we propose (for fixed covariates) the following test statistic

$$S_n = \frac{1}{n} \sum_{k=1}^n \int (\hat{\lambda}_n(t|x_k) - \tilde{\lambda}_n^*(t|x_k))^2 a(t) dt$$

Here $\tilde{\lambda}_n^*(\cdot|x_k)$ is the smoothed hypothetical hazard function at fixed covariate x_k . In Liero (2003a) it is shown that under certain conditions on K , b_n , the weights w_{ni} and h_n and on the smoothness of the underlying distribution functions that the (appropriate standardized) S_n is asymptotically normally distributed. Based on this limit statement a test procedure can be derived. Moreover, for testing the hypothesis, that $\lambda(\cdot|x)$ lies in a prespecified parametric class a test statistic with estimated parameters can be applied.

Appendix: Formulation of the Limit Theorem

This theorem is formulated not only for the behavior under the null hypothesis, but for general hazard rate λ . We define

$$\tilde{\lambda}_n(t) := \int K_{b_n}(t-s) \lambda(s) ds.$$

$$Q_n = \int (\hat{\lambda}_n(t) - \tilde{\lambda}_n(t))^2 a(t) dt$$

Further, let T_H be the right end point of the distribution H .

Theorem 1 *Suppose that*

- (i) K is a continuous density function vanishing outside the interval $[-L, L]$ for some $L > 0$.
- (ii) λ and H are Lipschitz continuous.
- (iii) The function a is continuous and $a(t) \equiv 0$ for all $t > T_H$.
- (iv) $b_n \rightarrow 0$ and $nb_n^2 \rightarrow \infty$.

Then for $n \rightarrow \infty$

$$\frac{nb_n^{1/2}}{\sigma} (Q_n - \mu_n) \xrightarrow{\mathcal{D}} \mathbf{N}(0, 1) \tag{10}$$

where

$$\mu_n = (nb_n)^{-1} \int \frac{\lambda(t)}{1-H(t)} a(t) dt \kappa_1 \quad \sigma^2 = 2 \int \left(\frac{\lambda(t)}{1-H(t)} \right)^2 a^2(t) dt \kappa_2$$

The proof of this theorem is given in Liero (2003b).

References

1. Beran, R. (1981). Nonparametric regression with randomly censored survival times. Technical Report, Univ. California, Berkeley.
2. Bagdonavičius, V. and Nikulin, M. (2002) Accelerated Life Models; Modeling and Statistical analysis. Boca Raton.; Chapman and Hall /CRC
3. Diehl, S. and Stute, W. (1988). Kernel density and hazard function estimation in the presence of censoring, *J. Multivariate Anal.*, **25**, 299–310
4. González-Manteiga, W. and Cadarso-Suarez, C. (1996). Asymptotic properties of a generalized Kaplan-Meier estimator with some applications, *J. Nonparametric Statistics* , **4**, 65–78
5. Liero, H. and Läuter, H. and Konakov, V. D. (1998). Nonparametric versus parametric goodness of fit, *Statistics*, **31**, 115–149
6. Liero, H. (2003a). Goodness of fit tests of L_2 -type, in: Statistical Inference for Semiparametric Models and Applications, Ed. Nikulin, M., Publisher Birkhäuser
7. Liero, H. (2003b). Testing the hazard rate, *Preprint, Institut für Mathematik, Universität Potsdam*
8. Lo, S.-H. and Singh, K. (1986). The product-limit estimator and the bootstrap: Some asymptotic representations, *Probab. Theory Related Fields*, **71**, 455–465
9. Singpurwalla, N. D. and Wong, M. Y. (1983). Estimation of the failure rate - a survey of nonparametric methods, part I: Non- Bayesian methods, *Commun. Statist.- Theory and Meth.*, **12**, 559–588
10. Tanner, M. A. and Wong, W. H. (1983). The estimation of the hazard function from randomly censored data by the kernel method, *Ann. Statist.*, **11**, 989–993
11. Van Keilegom, I. and Veraverbeke, N. (1997). Estimation and bootstrap with censored data in fixed design nonparametric regression , *Ann. Inst. Statist. Math.* , **49**, 467–401
12. Van Keilegom, I. and Veraverbeke, N. (2001). Hazard rate estimation in nonparametric regression with censored data, *Ann. Inst. Statist. Math.*, **53**, 730–745
13. Van Keilegom, I. and Veraverbeke, N. (2002). Density and hazard estimation in censored regression models, *Bernoulli* , **8**,607–625