

12-1-2017

Semi-Parametric Method to Estimate the Time-to-Failure Distribution and its Percentiles for Simple Linear Degradation Model

Laila Naji Ba Dakhn

Yarmouk University, Irbid, Jordan, la2010_math@yahoo.com


Mohammed Al-Haj Ebrahim

Yarmouk University, Irbid, Jordan, m_hassanb@hotmail.com

Omar Eidous

Yarmouk University, Irbid, Jordan, omarm@yu.edu.jo

Follow this and additional works at: <http://digitalcommons.wayne.edu/jmasm>

 Part of the [Applied Statistics Commons](#), [Social and Behavioral Sciences Commons](#), and the [Statistical Theory Commons](#)

Recommended Citation

Dakhn, L. N. B., Ebrahim, M. A.-H., & Eidous, O. (2017). Semi-Parametric Method to Estimate the Time-to-Failure Distribution and its Percentiles for Simple Linear Degradation Model. *Journal of Modern Applied Statistical Methods*, 16(2), 322-346. doi: 10.22237/jmasm/1509495420

This Regular Article is brought to you for free and open access by the Open Access Journals at DigitalCommons@WayneState. It has been accepted for inclusion in *Journal of Modern Applied Statistical Methods* by an authorized editor of DigitalCommons@WayneState.

Semi-Parametric Method to Estimate the Time-to-Failure Distribution and its Percentiles for Simple Linear Degradation Model

Laila Naji Ba Dakhn
Yarmouk University
Irbid, Jordan

Mohammed Al-Haj Ebrahim
Yarmouk University
Irbid, Jordan

Omar Eidous
Yarmouk University
Irbid, Jordan

Most reliability studies obtained reliability information by using degradation measurements over time, which contains useful data about the product reliability. Parametric methods like the maximum likelihood (ML) estimator and the ordinary least square (OLS) estimator are used widely to estimate the time-to-failure distribution and its percentiles. In this article, we estimate the time-to-failure distribution and its percentiles by using a semi-parametric estimator that assumes the parametric function to have a half-normal distribution or an exponential distribution. The performance of the semi-parametric estimator is compared via simulation study with the ML and OLS estimators by using the mean square error and length of the 95% bootstrap confidence interval as the basis criteria of the comparison. An application to real data is given. In general, if there are assumptions on the random effect parameter, the ML estimator is the best; otherwise the kernel semi-parametric estimator with half-normal distribution is the best.

Keywords: Degradation model, semi-parametric estimator, maximum likelihood estimator, ordinary least square estimator

Introduction

Meeker and Escobar (1998) defined the reliability of a unit as the probability that a unit will perform its intended function until a specified point of time under encountered use conditions. There are many proposed applications to measure the reliability of any product. One of these applications is the estimation of the time-to-failure distribution and its percentiles. In estimation, traditional life tests are

Ms. Laila Naji Ba Dakhn is a graduate student in the Department of Statistics. Email her at: la2010_math@yahoo.com. Prof. Mohammed Al-Haj Ebrahim is a Professor of Statistics. Email them at: m_hassanb@hotmail.com. Prof. Omar Eidous is a Professor of Statistics. Email them at: omarm@yu.edu.jo.

often not the most efficient way to obtain reliability information because few failure time data are observed by the end of the test; it is then difficult to use the traditional reliability analysis that records only failure time data to analyze life time data. Thus, it is possible to get failure data by degradation measurements over time which may contain useful data about product reliability.

Degradation is usually measured as a function of time T . Let $D(t)$ denote the actual sequence or path of the degradation of a particular unit over time t for each sample unit that will be observed, and let D_f denote the critical level for the degradation path where failure has occurred. The focus is on the linear degradation model for estimating the $100r^{\text{th}}$ percentile of the time-to-failure distribution.

Gertsbakh and Kordonskiĭ (1966/1969) discussed the degradation problem from an engineering point of view. They presented the Bernstein distribution, which describes the time-to-failure distribution for a simple linear model with random intercept and random slope. Amster and Hooper (1983) proposed a simple degradation model for single, multiple, and step-stress life tests. They explain how to use this model to estimate the central tendency of the time-to-failure distribution. Lu, Meeker, and Escobar (1996) compared the degradation analysis and traditional failure time analysis in terms of asymptotic efficiency. They demonstrated that the degradation analysis gives more precision than the traditional failure time analysis in general.

Al-Haj Ebrahim, Eidous, and Kmail (2009) proposed the nonparametric classical kernel method to estimate the time-to-failure distribution and its percentiles for the simple linear degradation model. They compared the performance of this method with the existing parametric methods like ML and OLS. They gave the time-to-failure distribution based on the classical kernel method (by assuming Gaussian kernel), which is

$$F_T(t) = 1 - \frac{1}{n} \sum_{i=1}^n \Phi \left(\frac{\frac{D_f}{t} - \beta_i}{h} \right)$$

where Φ is the distribution function of the standard normal distribution. They compute the bandwidth using the formula (Silverman, 1986)

$$h_{\text{int}} = 0.9An^{-1/5} \tag{1}$$

where $A = \min\{\text{SD}, \text{IQR}/1.34\}$, SD is the sample standard deviation, and IQR is the sample inter-quartile range.

The kernel function K is taken to be the Gaussian function

$$K(u) = \frac{1}{\sqrt{2\pi}} e^{-u^2/2}, \quad -\infty < u < \infty \quad (2)$$

and the smoothing parameter of the nonparametric estimator is computed using the formula in (1).

Model and Time-to-Failure Distribution

Consider the following simple linear degradation model to estimate the time-to-failure distribution:

$$y_{ij} = \beta_i t_{ij} + \varepsilon_{ij} \quad (3)$$

where y_{ij} is the observed degradation measurement of the i^{th} unit at time t_{ij} , β_i is the random effect parameter (the slope of the linear degradation model for unit i), t_{ij} is the failure time for the degradation model, and ε_{ij} is the random error term, where the ε_{ij} are iid with $N(0, \sigma_\varepsilon^2)$.

In general, the time-to-failure distribution can be written as a function of the degradation model parameters. The failure time T is defined as the time when the actual path $D(t)$ crosses the critical degradation level D_f , i.e. T is the solution of

$$D_f = D(t)$$

By considering the simple linear degradation model (3),

$$D_f = \beta T$$

Then the distribution function of the time-to-failure is

$$\begin{aligned}
 F_T(t) &= P(T \leq t) \\
 &= P\left(\frac{D_f}{\beta} \leq t\right) \\
 &= P\left(\beta \geq \frac{D_f}{t}\right) \\
 &= 1 - G_\beta\left(\frac{D_f}{t}\right), \quad t > 0
 \end{aligned} \tag{4}$$

where β is a random effect parameter and $G(\cdot)$ is the distribution function of β .

Let the 100^rth percentile of the time-to-failure distribution be denoted by t_r . To find t_r , we need to solve

$$r = F_T(t_r) = 1 - G_\beta\left(\frac{D_f}{t_r}\right)$$

with respect to t_r . This gives

$$t_r = \frac{D_f}{G_\beta^{-1}(1-r)} \tag{5}$$

It is clear that, for a fixed value of D_f , the distribution of T and the 100^rth percentile depend on the distribution of β , the random effect parameter. In some simple cases, a closed-form expression for $F(t)$ could be obtained, but for most practical path models, it is necessary to evaluate $F(t)$ using numerical methods. For example, consider the linear degradation model (3) with random effect distributed as $N(\mu, \sigma^2)$.

From equation (4),

$$F_T(t) = \Phi\left(\frac{t_r\mu - D_f}{t_r\sigma}\right)$$

and, from equation (5),

$$t_r = \frac{D_f}{\mu - \sigma\Phi^{-1}(r)}$$

where, $\Phi(z)$ is the standard normal cumulative distribution function. As another example, if $\beta \sim \exp(\theta)$, then

$$F_r(t) = e^{-D_f/t^\theta}$$

and

$$t_r = \frac{-D_f}{\theta \ln(r)}$$

For the above two examples, the parameters μ , σ , and θ can be estimated using the ML or OLS methods, or even by any good statistical method.

Estimating Percentiles of Time-to-Failure Distribution Using Semi Parametric Density Method

If $\hat{\theta}$ is an estimator of θ , we can construct the following semi-parametric estimator of $f(x)$:

$$\tilde{f}_{SP}(x) = \frac{f(x; \hat{\theta})}{nh} \sum_{i=1}^n K\left(\frac{x - X_i}{h}\right) \left(f(X_i; \hat{\theta})\right)^{-1}, \quad x \in \mathbb{R} \quad (6)$$

This estimator is semi-parametric because it combines a nonparametric estimator, the classical kernel estimator, and a parametric estimator, $f(x, \theta)$. In this work, two versions of $\tilde{f}_{SP}(x)$ are considered and studied. The first one assumes $f(x, \theta)$ follows a half-normal distribution and the other assumes $f(x, \theta)$ follows an exponential distribution.

If the degradation model is a simple linear as in (3), and if $\beta_1, \beta_2, \dots, \beta_n$ is a random sample from unknown pdf ($g_\beta(b)$), then we proposed – in this section – the semi-parametric estimator for the time-to-failure distribution and its percentiles.

The Half-Normal Distribution

The semi-parametric estimator of $g_\beta(b)$ that depends on the half-normal distribution is

$$\hat{g}_{\text{SPH}}(b) = \frac{v_1(b)}{nh} \sum_{j=1}^n K\left(\frac{b-\beta_j}{h}\right) (v_1(\beta_j))^{-1}$$

where

$$v_1(b) = \sqrt{\frac{2}{\pi\sigma^2}} e^{-\frac{b^2}{2\sigma^2}}$$

with

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^n \beta_i^2}{n}$$

Taking the kernel function $K(u)$ to be a Gaussian function,

$$\begin{aligned} \hat{g}_{\text{SPH}}(b) &= \frac{\sqrt{\frac{2}{\pi\sigma^2}} e^{-\frac{b^2}{2\sigma^2}}}{nh} \sum_{j=1}^n \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{b-\beta_j}{h}\right)^2} \left(\sqrt{\frac{2}{\pi\sigma^2}} e^{-\frac{\beta_j^2}{2\sigma^2}} \right)^{-1} \\ &= \frac{e^{-\frac{b^2}{2\sigma^2}}}{nh} \sum_{j=1}^n \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{b-\beta_j}{h}\right)^2} \left(e^{-\frac{\beta_j^2}{2\sigma^2}} \right)^{-1} \\ &= \frac{1}{n} \sum_{j=1}^n \frac{\left(e^{-\frac{\beta_j^2}{2\sigma^2}} \right)^{-1}}{h\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{b-\beta_j}{h}\right)^2 - \frac{b^2}{2\sigma^2}} \\ &= \frac{1}{n} \sum_{j=1}^n \frac{\left(e^{-\frac{\beta_j^2}{2\sigma^2}} \right)^{-1}}{h\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2 h^2}(\hat{\sigma}^2(b-\beta_j)^2 + h^2 b^2)} \\ &= \frac{1}{n} \sum_{j=1}^n \frac{\left(e^{-\frac{\beta_j^2}{2\sigma^2}} \right)^{-1}}{h\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2 h^2} \left(h^2 + \hat{\sigma}^2 \right) \left(b - \frac{\hat{\sigma}^2 \beta_j}{h^2 + \hat{\sigma}^2} \right)^2 + \frac{\hat{\sigma}^2 h^2 \beta_j^2}{h^2 + \hat{\sigma}^2}} \end{aligned}$$

SEMI-PARAMETRIC ESTIMATION FOR LINEAR DEGRADATION MODEL

$$\begin{aligned}
 &= \frac{1}{n} \sum_{j=1}^n \frac{e^{-\frac{1}{2\hat{\sigma}^2 h^2} \left(\frac{\hat{\sigma}^2 h^2 \beta_j^2}{h^2 + \hat{\sigma}^2} \right)}}{e^{-\frac{\beta_j^2}{2\sigma^2}}} \frac{1}{h\sqrt{2\pi}} e^{-\frac{h^2 + \hat{\sigma}^2}{2\hat{\sigma}^2 h^2} \left(b - \frac{\hat{\sigma}^2 \beta_j}{h^2 + \hat{\sigma}^2} \right)^2} \\
 &= \frac{1}{n} \sum_{j=1}^n e^{\frac{h^2 \beta_j^2}{2\hat{\sigma}^2 (h^2 + \hat{\sigma}^2)}} \frac{1}{h\sqrt{2\pi}} e^{-\frac{h^2 + \hat{\sigma}^2}{2\hat{\sigma}^2 h^2} \left(b - \frac{\hat{\sigma}^2 \beta_j}{h^2 + \hat{\sigma}^2} \right)^2}
 \end{aligned}$$

The estimation of the distribution function of the time-to-failure is

$$\begin{aligned}
 \hat{F}_T(t) &= 1 - \hat{G}_{SPH} \left(\frac{D_f}{t} \right) \\
 &= 1 - \int_{-\infty}^{\frac{D_f}{t}} \hat{g}_{SPH}(b) db \\
 &= 1 - \int_{-\infty}^{\frac{D_f}{t}} \frac{1}{n} \sum_{j=1}^n e^{\frac{h^2 \beta_j^2}{2\hat{\sigma}^2 (h^2 + \hat{\sigma}^2)}} \frac{1}{h\sqrt{2\pi}} e^{-\frac{h^2 + \hat{\sigma}^2}{2\hat{\sigma}^2 h^2} \left(b - \frac{\hat{\sigma}^2 \beta_j}{h^2 + \hat{\sigma}^2} \right)^2} db \\
 &= 1 - \frac{1}{n} \sum_{j=1}^n e^{\frac{h^2 \beta_j^2}{2\hat{\sigma}^2 (h^2 + \hat{\sigma}^2)}} \int_{-\infty}^{\frac{D_f}{t}} \frac{1}{h\sqrt{2\pi}} e^{-\frac{h^2 + \hat{\sigma}^2}{2\hat{\sigma}^2 h^2} \left(b - \frac{\hat{\sigma}^2 \beta_j}{h^2 + \hat{\sigma}^2} \right)^2} db \\
 &= 1 - \frac{1}{n} \sqrt{\frac{\hat{\sigma}^2}{h^2 + \hat{\sigma}^2}} \sum_{j=1}^n e^{\frac{h^2 \beta_j^2}{2\hat{\sigma}^2 (h^2 + \hat{\sigma}^2)}} \text{Pro} \left(Q \leq \frac{D_f}{t} \right)
 \end{aligned}$$

where Q is a random variable distributed as

$$\text{N} \left(\frac{\hat{\sigma}^2 \beta_j}{h^2 + \hat{\sigma}^2}, \frac{\hat{\sigma}^2 h^2}{(h^2 + \hat{\sigma}^2)} \right)$$

Then

$$\hat{F}_T(t) = 1 - \frac{1}{n} \sqrt{\frac{\hat{\sigma}^2}{h^2 + \hat{\sigma}^2}} \sum_{j=1}^n e^{\frac{h^2 \beta_j^2}{2\hat{\sigma}^2(h^2 + \hat{\sigma}^2)}} \Phi \left(\frac{D_f - \frac{\hat{\sigma}^2 \beta_j}{h^2 + \hat{\sigma}^2}}{\sqrt{\frac{\hat{\sigma}^2 h^2}{(h^2 + \hat{\sigma}^2)}}} \right) \quad (7)$$

where $\Phi(\cdot)$ is a standard normal distribution. To estimate the 100^rth percentiles (denoted by $\hat{t}_{r\text{-SPH}}$), we should solve $\hat{F}_{T\text{-SPH}}(\hat{t}_{r\text{-SPH}}) = r$ numerically for $\hat{t}_{r\text{-SPH}}$.

The Exponential Distribution

The semi-parametric estimator of $g_{\beta}(b)$ based on the exponential distribution is

$$\hat{g}_{\text{SPE}}(b) = \frac{v_2(b)}{nh} \sum_{j=1}^n K\left(\frac{b - \beta_j}{h}\right) (v_2(\beta_j))^{-1}$$

where

$$v_2(b) = \frac{e^{-\frac{b}{\hat{\alpha}}}}{\hat{\alpha}}$$

with

$$\hat{\alpha} = \frac{\sum_{i=1}^n \beta_i}{n}$$

By taking $K(u)$ to be a Gaussian function, we obtain

$$\begin{aligned} \hat{g}_{\text{SPE}}(b) &= \frac{e^{-\frac{b}{\hat{\alpha}}}}{\hat{\alpha}nh} \sum_{j=1}^n \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{b-\beta_j}{h}\right)^2} \left(\frac{e^{-\frac{\beta_j}{\hat{\alpha}}}}{\hat{\alpha}} \right)^{-1} \\ &= \frac{1}{nh} \sum_{j=1}^n \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{b-\beta_j}{h}\right)^2} e^{-\frac{b}{\hat{\alpha}}} \left(\frac{e^{-\frac{\beta_j}{\hat{\alpha}}}}{\hat{\alpha}} \right)^{-1} \end{aligned}$$

SEMI-PARAMETRIC ESTIMATION FOR LINEAR DEGRADATION MODEL

$$\begin{aligned}
 &= \frac{1}{n} \sum_{j=1}^n \frac{\left(e^{-\frac{\beta_j}{\hat{\alpha}}}\right)^{-1}}{h\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{b-\beta_j}{h}\right)^2 - \frac{b}{\hat{\alpha}}} \\
 &= \frac{1}{n} \sum_{j=1}^n \frac{\left(e^{-\frac{\beta_j}{\hat{\alpha}}}\right)^{-1}}{h\sqrt{2\pi}} e^{-\frac{1}{2h^2}\left((b-\beta_j)^2 + 2h^2\frac{b}{\hat{\alpha}}\right)} \\
 &= \frac{1}{n} \sum_{j=1}^n \frac{\left(e^{-\frac{\beta_j}{\hat{\alpha}}}\right)^{-1}}{h\sqrt{2\pi}} e^{-\frac{1}{2h^2}\left(\left(b-\left(\beta_j-\frac{h^2}{\hat{\alpha}}\right)\right)^2 + \frac{2h^2}{\hat{\alpha}}\left(\beta_j-\frac{h^2}{\hat{\alpha}}\right)\right)} \\
 &= \frac{1}{n} \sum_{j=1}^n \frac{\left(e^{-\frac{\beta_j}{\hat{\alpha}}}\right)^{-1}}{h\sqrt{2\pi}} e^{-\frac{1}{2h^2}\left(b-\left(\beta_j-\frac{h^2}{\hat{\alpha}}\right)\right)^2 - \frac{1}{\hat{\alpha}}\left(\beta_j-\frac{h^2}{\hat{\alpha}}\right)} \\
 &= \frac{1}{n} \sum_{j=1}^n e^{\frac{h^2}{2\hat{\alpha}^2}} \frac{1}{h\sqrt{2\pi}} e^{-\frac{1}{2h^2}\left(b-\left(\beta_j-\frac{h^2}{\hat{\alpha}}\right)\right)^2}
 \end{aligned}$$

The estimation of the distribution function of the time-to-failure by using $\hat{g}_{\text{SPE}}(b)$ is

$$\begin{aligned}
 \hat{F}_T(t) &= 1 - \hat{G}_{\text{SPE}}\left(\frac{D_f}{t}\right) \\
 &= 1 - \int_{-\infty}^{\frac{D_f}{t}} \hat{g}_{\text{SPE}}(b) db \\
 &= 1 - \int_{-\infty}^{\frac{D_f}{t}} \frac{1}{n} \sum_{j=1}^n e^{\frac{h^2}{2\hat{\alpha}^2}} \frac{1}{h\sqrt{2\pi}} e^{-\frac{1}{2h^2}\left(b-\left(\beta_j-\frac{h^2}{\hat{\alpha}}\right)\right)^2} db \\
 &= 1 - \frac{1}{n} \sum_{j=1}^n e^{\frac{h^2}{2\hat{\alpha}^2}} \int_{-\infty}^{\frac{D_f}{t}} \frac{1}{h\sqrt{2\pi}} e^{-\frac{1}{2h^2}\left(b-\left(\beta_j-\frac{h^2}{\hat{\alpha}}\right)\right)^2} db \\
 &= 1 - \frac{1}{n} \sum_{j=1}^n e^{\frac{h^2}{2\hat{\alpha}^2}} \text{Pro}\left(S \leq \frac{D_f}{t}\right)
 \end{aligned}$$

where S is a random variable distributed as

$$N\left(\beta_j - \frac{h^2}{\hat{\alpha}}, h^2\right)$$

Therefore

$$\hat{F}_T(t) = 1 - \frac{1}{n} \sum_{j=1}^n e^{\frac{h^2}{2\hat{\alpha}^2}} \Phi\left(\frac{\frac{D_f}{t} - \left(\beta_j - \frac{h^2}{\hat{\alpha}}\right)}{h}\right) \quad (8)$$

where $\Phi(\cdot)$ is a standard normal distribution. To estimate the $100r^{\text{th}}$ percentiles (denoted by $\hat{t}_{r\text{-SPH}}$) we should solve $\hat{F}_{T\text{-SPH}}(\hat{t}_{r\text{-SPH}}) = r$ numerically with respect to $\hat{t}_{r\text{-SPH}}$.

Estimating Percentiles of Time-to-Failure Distribution Using Maximum Likelihood (ML) Estimator Method

Consider the simple linear degradation model

$$y_{ij} = \beta_i t_{ij} + \varepsilon_{ij}, \quad i = 1, \dots, n; j = 1, \dots, m$$

where y_{ij} is the observed degradation measurement of the i^{th} unit at time t_{ij} , β_i is the random effect parameter (the slope of the linear degradation model for unit i), t_{ij} is the soft failure for the degradation model, and ε_{ij} is the random error term, where the ε_{ij} are iid with $N(0, \sigma_\varepsilon^2)$. By using the formula of the time-to-failure distribution in (4), we will construct the ML estimator of t_r for the following distributions:

The Half-Normal Distribution

If $\beta_i \sim \text{half normal}(\sigma^2)$, the time-to-failure distribution is

$$F_T(t) = 1 - \int_0^{\frac{D_f}{t}} \sqrt{\frac{2}{\pi\sigma^2}} e^{-\frac{b^2}{2\sigma^2}} db \quad (9)$$

By the Leibniz integral rule, and by differentiating both sides of (9) with respect to t , we obtain the pdf of the time-to-failure distribution, which is

$$f_T(t; \sigma^2) = \frac{D_f}{t^2} \sqrt{\frac{2}{\pi\sigma^2}} e^{-\frac{D_f^2}{2\sigma^2 t^2}} \quad (10)$$

Now, to find the ML estimator of σ^2 , let t_1, t_2, \dots, t_n be a random sample from (10); then the natural logarithm of the likelihood function of σ^2 is

$$\begin{aligned} \ln(L(\sigma^2; t_1, t_2, \dots, t_n)) &= \ln\left(\prod_{i=1}^n f_T(t_i; \sigma^2)\right) \\ &= \ln\left(\left(\sqrt{\frac{2}{\pi\sigma^2}} D_f\right)^n \left(\prod_{i=1}^n \frac{1}{t_i^2}\right) e^{-\frac{D_f^2}{2\sigma^2} \sum_{i=1}^n \frac{1}{t_i^2}}\right) \\ &= n \ln\left(\sqrt{\frac{2}{\pi}} D_f\right) - \frac{n}{2} \ln \sigma^2 - \frac{D_f^2}{2\sigma^2} \sum_{i=1}^n \frac{1}{t_i^2} - \ln\left(\prod_{i=1}^n t_i^2\right) \end{aligned}$$

Now, by differentiating with respect to σ^2 , we obtain

$$\frac{d}{d\sigma^2} \left(\ln(L(\sigma^2; t_1, t_2, \dots, t_n)) \right) = -\frac{n}{2\sigma^2} + \frac{D_f^2}{2\sigma^4} \sum_{i=1}^n \frac{1}{t_i^2}$$

By solving

$$\frac{d}{d\sigma^2} \left(\ln(L(\sigma^2; t_1, t_2, \dots, t_n)) \right) = 0$$

we obtain the ML estimator of σ^2 , which is

$$\sigma^2_{MLH} = \frac{D_f^2}{n} \sum_{i=1}^n \frac{1}{t_i^2} \quad (11)$$

The ML estimator of t_r (denoted by \hat{t}_{MLH}) is obtained by solving the following equation with respect to \hat{t}_{MLH} :

$$r = 1 - \int_0^{\hat{t}_{MLH}^{D_f}} \sqrt{\frac{2}{\pi \sigma_{MLH}^2}} e^{-\frac{b^2}{2\sigma_{MLH}^2}} db \quad (12)$$

The Exponential Distribution

If $\beta_i \sim \exp(\alpha)$, then the time-to-failure distribution is

$$F_T(t) = e^{-\frac{D_f}{\alpha t}} \quad (13)$$

By differentiating both sides of (13) with respect to t , we obtain the pdf of the time-to-failure distribution, which is

$$f_T(t; \alpha) = \frac{D_f}{\alpha t^2} e^{-\frac{D_f}{\alpha t}} \quad (14)$$

To find the ML estimator of α let t_1, t_2, \dots, t_n be a random sample from (14); then the natural logarithm of the likelihood function of α is

$$\begin{aligned} \ln(L(\alpha; t_1, t_2, \dots, t_n)) &= \ln\left(\prod_{i=1}^n f_T(t_i; \alpha)\right) \\ &= \ln\left(\left(\frac{D_f}{\alpha}\right)^n \left(\prod_{i=1}^n \frac{1}{t_i^2}\right) e^{-\frac{D_f}{\alpha} \sum_{i=1}^n \frac{1}{t_i}}\right) \\ &= n \ln(D_f) - n \ln(\alpha) - \frac{D_f}{\alpha} \sum_{i=1}^n \frac{1}{t_i} - \ln\left(\prod_{i=1}^n t_i^2\right) \end{aligned}$$

Therefore

$$\frac{d}{d\alpha} \left(\ln(L(\alpha; t_1, t_2, \dots, t_n)) \right) = -\frac{n}{\alpha} + \frac{D_f}{\alpha^2} \sum_{i=1}^n \frac{1}{t_i}$$

By solving

$$\frac{d}{d\alpha} \left(\ln(L(\alpha; t_1, t_2, \dots, t_n)) \right) = 0$$

we obtain the ML estimator of α , which is given by

$$\hat{\alpha}_{\text{MLE}} = \frac{D_f}{n} \sum_{i=1}^n \frac{1}{t_i} \quad (15)$$

Then the ML estimator of t_r (denoted by \hat{t}_{MLE}) is

$$\begin{aligned} \hat{t}_{\text{MLE}} &= -\frac{D_f}{\hat{\alpha}_{\text{MLE}} \ln(r)} \\ &= -\frac{n}{\ln(r) \sum_{i=1}^n \frac{1}{t_i}} \end{aligned} \quad (16)$$

Estimating Percentiles of Time-to-Failure Distribution Using OLS Estimator Method

By considering the same degradation model that was studied in the previous section, and by letting $\beta_1, \beta_2, \dots, \beta_n$ be a random sample of size n from the probability density function $g_\beta(b; \mu)$ and the distribution function $G_\beta(b; \mu)$, the OLS estimator of μ (denoted by $\hat{\mu}_{\text{OLS}}$) will be obtained as follows:

$$\begin{aligned} Q(\mu) &= \sum_{i=1}^n \sum_{j=1}^m (y_{ij} - E(y_{ij}))^2 \\ &= \sum_{i=1}^n \sum_{j=1}^m (y_{ij} - t_{ij} E(\beta_i))^2 \end{aligned} \quad (17)$$

where $E(\beta_i)$ is a function of μ .

By minimizing (17) with respect to μ we get the OLS estimator of μ . Then the OLS estimator for the time-to-failure distribution is

$$\hat{F}_{\text{OLS}}(t) = 1 - G_\beta \left(\frac{D_f}{\hat{t}_{r-\text{OLS}}} \right) \quad (18)$$

where \hat{t}_{r-OLS} is the OLS estimator of t_r that is given by solving

$$r = 1 - G_{\beta} \left(\frac{D_f}{\hat{t}_{r-OLS}} \right)$$

By using the formula of the time-to-failure distribution (18) we obtain the OLS estimator for the following distributions:

The Half-Normal Distribution

If $\beta_i \sim \text{half normal}(\sigma^2)$, then the OLS estimator of σ^2 (denoted by σ^2_{OLSH}) is obtained by minimizing

$$Q(\sigma^2) = \sum_{i=1}^n \sum_{j=1}^m (y_{ij} - E(y_{ij}))^2$$

where

$$E(y_{ij}) = \sqrt{\frac{2}{\pi} \sigma^2 t_{ij}}$$

Therefore

$$Q(\sigma^2) = \sum_{i=1}^n \sum_{j=1}^m \left(y_{ij} - \sqrt{\frac{2}{\pi} \sigma^2 t_{ij}} \right)^2$$

and

$$\frac{dQ(\sigma^2)}{d\sigma^2} = \frac{-2}{\sqrt{2\pi\sigma^2}} \sum_{i=1}^n \sum_{j=1}^m \left(y_{ij} - \sqrt{\frac{2}{\pi} \sigma^2 t_{ij}} \right) t_{ij}$$

Equating the above derivative to zero, we get

$$\sum_{i=1}^n \sum_{j=1}^m \left(y_{ij} - \sqrt{\frac{2}{\pi}} \sigma^2 t_{ij} \right) t_{ij} = 0$$

Now, solving the last equation with respect to σ^2 , we obtain

$$\sigma^2_{\text{OLSH}} = \frac{\pi}{2} \left(\frac{\sum_{i=1}^n \sum_{j=1}^m y_{ij} t_{ij}}{\sum_{i=1}^n \sum_{j=1}^m t_{ij}^2} \right)^2 \quad (19)$$

and the $\hat{t}_{r\text{-OLSH}}$ is given by solving the following equation with respect to $\hat{t}_{r\text{-OLSH}}$:

$$r = 1 - G_{\beta} \left(\frac{D_f}{\hat{t}_{r\text{-OLSH}}} \right) \quad (20)$$

The Exponential Distribution

If $\beta_i \sim \exp(\alpha)$, then the OLS estimator of α (denoted by $\hat{\alpha}_{\text{OLSE}}$) is obtained by minimizing

$$Q(\alpha) = \sum_{i=1}^n \sum_{j=1}^m \left(y_{ij} - E(y_{ij}) \right)^2$$

where $E(y_{ij}) = \alpha t_{ij}$. Thus

$$Q(\alpha) = \sum_{i=1}^n \sum_{j=1}^m \left(y_{ij} - \alpha t_{ij} \right)^2$$

and

$$\frac{dQ(\alpha)}{d\alpha} = -2 \sum_{i=1}^n \sum_{j=1}^m \left(y_{ij} - \alpha t_{ij} \right) t_{ij}$$

By equating the above derivative to zero, we get

$$\sum_{i=1}^n \sum_{j=1}^m (y_{ij} - \alpha t_{ij}) t_{ij} = 0$$

The OLS estimator of α is obtained by solving the last equation with respect to α . Thus,

$$\hat{\alpha}_{\text{OLSE}} = \frac{\sum_{i=1}^n \sum_{j=1}^m y_{ij} t_{ij}}{\sum_{i=1}^n \sum_{j=1}^m t_{ij}^2} \tag{21}$$

The estimator $\hat{t}_{r-\text{OLSE}}$ of t_r is given by

$$\begin{aligned} \hat{t}_{r-\text{OLSE}} &= -\frac{D_f}{\hat{\alpha}_{\text{OLSE}} \ln(r)} \\ &= -\frac{D_f}{\ln(r)} \frac{\sum_{i=1}^n \sum_{j=1}^m t_{ij}^2}{\sum_{i=1}^n \sum_{j=1}^m y_{ij} t_{ij}} \end{aligned} \tag{22}$$

Simulation Study and Results

Consider the performance of the four estimators of t_r . The bandwidth for each estimator is computed by using formula (1). The bias (B), the mean square error (MSE), and the length of 95% bootstrap confidence interval using the bootstrap percentile method (Efron & Tibshirani, 1993; Racine & MacKinnon, 2007) of each estimator are computed from the data of size n that is simulated from the selected distributions, half normal(σ^2) or exp(α). Let $\beta_1, \beta_2, \dots, \beta_n$ be a random sample of size n generated from one of the above distributions. To compute B and MSE for each of the four estimators of t_r , find the exact value of t_r .

Case 1

If $\beta_i \sim$ half normal(σ^2), then by using equation (9), the distribution function of the time-to-failure is

$$F_T(t) = 1 - \int_0^{\frac{D_f}{t}} \sqrt{\frac{2}{\pi\sigma^2}} e^{-\frac{b^2}{2\sigma^2}} db$$

SEMI-PARAMETRIC ESTIMATION FOR LINEAR DEGRADATION MODEL

and, based on equation (5), the exact value of t_r is

$$t_r = \frac{D_f}{H^{-1}(1-r)}$$

where $H^{-1}(\cdot)$ is the inverse distribution function of the half-normal distribution.

Case 2

If $\beta_i \sim \exp(\alpha)$, then based on equation (13),

$$F_T(t) = 1 - \left(1 - e^{-\frac{D_f}{\alpha t}}\right) = e^{-\frac{D_f}{\alpha t}}$$

and, based on equation (5), the exact value of t_r is

$$t_r = -\frac{D_f}{\alpha \ln(r)}$$

In simulation, the initial values are,

- Sample size $n = 20, 40, \text{ or } 60$, $r = 0.1, 0.2, \text{ or } 0.3$, and σ^2 or $\alpha = 10$.
- The critical level of degradation $D_f = 20$ for each sample of size n .
- The number of iterations to compute B and MSE is $N = 2000$.
- The number of bootstrap iterations are $M = 1000$.

Simulation results are presented in Tables 1-6.

Table 1. *B*, MSE, and length of 95% bootstrap confidence interval of estimate t_r from half normal(10) with $n = 20$

<i>r</i>	t_r	SPH estimator		SPE estimator		ML estimator		OLS estimator		Length of 95% bootstrap CI			
		B	MSE	B	MSE	B	MSE	B	MSE	SPH	SPE	ML	OLS
0.1	1.2159	0.0514	0.0566	0.2687	0.1518	0.0436	0.0463	0.0359	0.0502	1.3591	1.6369	0.8280	1.2687
0.2	1.5606	0.0579	0.0865	0.3128	0.2269	0.0665	0.0789	0.0474	0.0800	1.7436	2.1057	0.9522	1.6558
0.3	1.9297	0.0670	0.1389	0.4047	0.4195	0.0773	0.1176	0.0627	0.1257	2.3842	3.0097	1.3241	1.9469

Table 2. *B*, MSE and length of 95% bootstrap confidence interval of estimate t_r from half normal(10) with $n = 40$

<i>r</i>	t_r	SPH estimator		SPE estimator		ML estimator		OLS estimator		Length of 95% bootstrap CI			
		B	MSE	B	MSE	B	MSE	B	MSE	SPH	SPE	ML	OLS
0.1	1.2159	0.0250	0.0253	0.1975	0.0721	0.0232	0.0208	0.0152	0.0242	0.9616	1.1542	0.5720	0.8045
0.2	1.5606	0.0376	0.0418	0.2326	0.1097	0.0256	0.0325	0.0214	0.0390	1.0938	1.3996	0.7565	1.0598
0.3	1.9297	0.0271	0.0613	0.2710	0.1695	0.0397	0.0535	0.0196	0.0496	1.4412	1.7483	0.8858	1.3863

Table 3. *B*, MSE and length of 95% bootstrap confidence interval of estimate t_r from half normal(10) with $n = 60$

<i>r</i>	t_r	SPH estimator		SPE estimator		ML estimator		OLS estimator		Length of 95% bootstrap CI			
		B	MSE	B	MSE	B	MSE	B	MSE	SPH	SPE	ML	OLS
0.1	1.2159	0.0156	0.0159	0.1689	0.0485	0.0184	0.0136	0.0072	0.0139	0.7482	0.8255	0.4627	0.6795
0.2	1.5606	0.0146	0.0267	0.1820	0.0689	0.0177	0.0228	0.0185	0.0234	0.9139	1.0511	0.5979	0.8825
0.3	1.9297	0.0186	0.0449	0.2301	0.1182	0.0243	0.0347	0.0171	0.0363	1.3115	1.5285	0.7201	1.0552

SEMI-PARAMETRIC ESTIMATION FOR LINEAR DEGRADATION MODEL

Table 4. B, MSE and length of 95% bootstrap confidence interval of estimate t_r from $\exp(10)$ with $n = 20$

r	t_r	SPH estimator		SPE estimator		ML estimator		OLS estimator		Length of 95% bootstrap CI			
		B	MSE	B	MSE	B	MSE	B	MSE	SPH	SPE	ML	OLS
0.1	0.8690	0.0248	0.0641	0.3105	0.1937	0.0525	0.0527	0.0553	0.0535	1.4330	1.7700	0.8370	1.2080
0.2	1.2430	0.0359	0.1100	0.3664	0.2933	0.0732	0.0970	0.0663	0.0999	2.0460	2.4300	1.2230	1.7210
0.3	1.6610	0.0214	0.1794	0.4704	0.5168	0.0998	0.1892	0.0821	0.1815	2.8010	3.7720	1.4590	2.4160

Table 5. B, MSE and length of 95% bootstrap confidence interval of estimate t_r from $\exp(10)$ with $n = 40$

r	t_r	SPH estimator		SPE estimator		ML estimator		OLS estimator		Length of 95% bootstrap CI			
		B	MSE	B	MSE	B	MSE	B	MSE	SPH	SPE	ML	OLS
0.1	0.8690	0.0061	0.0269	0.2424	0.0972	0.0260	0.0219	0.0239	0.0227	0.9510	1.1250	0.5590	0.7840
0.2	1.2430	0.0064	0.0545	0.2783	0.1522	0.0312	0.0449	0.0210	0.0455	1.3440	1.5050	0.8460	1.1600
0.3	1.6610	-0.0086	0.0889	0.3434	0.2480	0.0193	0.0734	0.0435	0.0762	1.7670	2.1250	1.0930	1.4220

Table 6. B, MSE and length of 95% bootstrap confidence interval of estimate t_r from $\exp(10)$ with $n = 60$

r	t_r	SPH estimator		SPE estimator		ML estimator		OLS estimator		Length of 95% bootstrap CI			
		B	MSE	B	MSE	B	MSE	B	MSE	SPH	SPE	ML	OLS
0.1	0.8690	-0.0017	0.0183	0.2038	0.0683	0.0150	0.0131	0.0160	0.0136	0.7760	0.8680	0.4490	0.6230
0.2	1.2430	-0.0001	0.0333	0.2404	0.0994	0.0230	0.0284	0.0217	0.0285	1.0520	1.2010	0.6710	0.8720
0.3	1.6610	-0.0224	0.0592	0.2801	0.1590	0.0300	0.0512	0.0242	0.0479	1.3600	1.5970	0.8650	1.1750

From Tables 1-6, the following conclusions may be made:

- The MSE for each estimator decrease as n increases.
- The MSE for each estimator increase as r increases.
- The length of the 95% bootstrap confidence interval is decrease as n increases.
- By comparing the MSE of the four estimators, the ML and OLS estimators have the smallest values of the MSE and the semi parametric half normal estimator close to them. ML estimator has the smallest value of the MSE and the shortest length of a 95% bootstrap confidence interval for each distribution and different sample size, so ML estimator has the best performance.

Estimating $t_{0.5}$ Using Data with Misspecified Density

In this section, we will study and compare the performance of the semi-parametric method, OLS, and ML estimators for t_r when the distribution of the random effect is not chosen correctly. To perform this comparison, generate the random effect from Weibull(2, 15) and assume this generated sample is from half normal(15) or exp(15). The true value of $t_{0.5}$ is

$$t_{0.5} = \frac{20}{W^{-1}(0.5)} = 1.6015$$

where, $W^{-1}(\cdot)$ is the inverse distribution function of the Weibull(2, 15) distribution. Under this misspecification, estimate $t_{0.5}$ using four estimators.

Table 7. Estimating $t_{0.5}$ for a sample from a Weibull(2, 15) distribution that is misspecified as a half normal(15) distribution

Estimator	Bias (B)	Mean Square Error (MSE)
$\hat{t}_{0.5\text{-SPH}}$	0.0353	0.0188
$\hat{t}_{0.5\text{-SPE}}$	0.0977	0.0319
$\hat{t}_{0.5\text{-ML}}$	0.3889	0.1681
$\hat{t}_{0.5\text{-OLS}}$	0.1874	0.0500

SEMI-PARAMETRIC ESTIMATION FOR LINEAR DEGRADATION MODEL

Table 8. Estimating $t_{0.5}$ for a sample from a Weibull(2, 15) distribution that is misspecified as an exp(15) distribution

Estimator	Bias (B)	Mean Square Error (MSE)
$\hat{t}_{0.5\text{-SPH}}$	0.03530	0.01880
$\hat{t}_{0.5\text{-SPE}}$	0.09770	0.03190
$\hat{t}_{0.5\text{-ML}}$	0.57644	0.35365
$\hat{t}_{0.5\text{-OLS}}$	0.58674	0.36570

In this simulation, the initial values are

- Sample size $n = 60$, σ^2 or $\alpha = 15$, and $r = 0.5$.
- The critical level of degradation $D_f = 20$.
- The number of iterations to compute B and MSE is $N = 2000$.

Simulation results are presented in Tables 7 and 8.

From Tables 7 and 8 we conclude the following:

- The ML and OLS estimators perform poorly when the random effect distribution is misspecified.
- The semi-parametric half-normal estimator has the best performance.

Real Data Application

The laser degradation data gives the percent increase in laser operating current for GaAs lasers tested at 80°C which is presented in Table c.17 of Meeker and Escobar (1998, p. 642). In this article, failure is assumed to occur at the critical degradation level $D_f = 5$. Figure 1 shows percent increase in operating current for GaAs lasers tested at 80°C.

Data Analysis

Consider the data to estimate the percentiles of the time-to-failure distribution for estimators which have been discussed previously (semi-parametric estimators (SPH & SPE), maximum likelihood estimator (ML), and ordinary least square estimator (OLS)). These estimators will be compared by computing the mean square error (MSE) and the length of the 95% bootstrap confidence interval of the percentiles ($r = 0.5$) of the time-to-failure distribution.

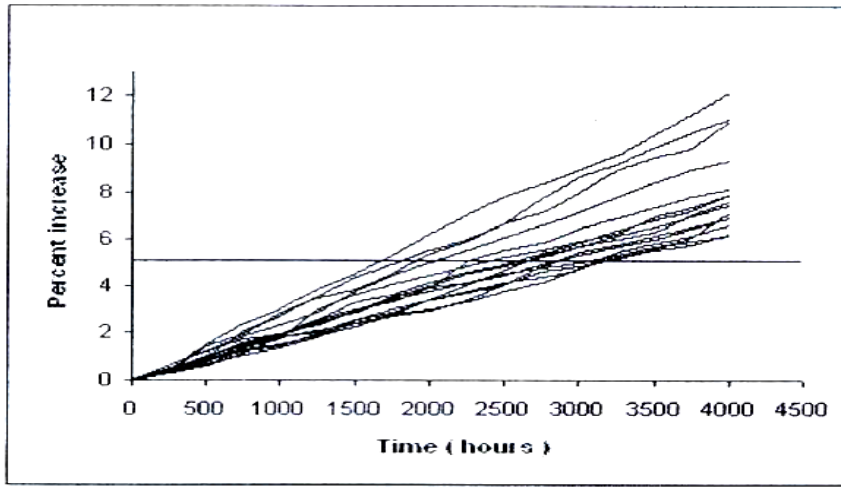


Figure 1. Percent increase in laser operating current for GaAs lasers tested at 80°C

Table 9. Failure time and slope for each unit i

Unit i	t_i	$\hat{\beta}_i$
1	7.1479	0.694
2	8.0192	0.599
3	11.6707	0.445
4	12.0667	0.418
5	10.6066	0.460
6	7.5732	0.692
7	12.5926	0.407
8	12.6000	0.390
9	10.2093	0.493
10	6.6667	0.756
11	10.5000	0.475
12	9.4884	0.505
13	8.9730	0.523
14	11.2931	0.430
15	12.4933	0.408

To compute the MSE, the true value of t_r and the values of $\beta_1, \beta_2, \dots, \beta_n$, the slopes of the linear model (3), must be known. From the laser degradation data, scale the times of the degradation measurements by dividing each time by 250. To get the failure time, use linear interpolation and, by fitting simple linear regression

between t_i and y_{ij} , obtain the estimation of $\beta_1, \beta_2, \dots, \beta_n$. Table 9 contains the time-to-failure t_i and the values of the slope estimate $\hat{\beta}_i$.

Estimating the 50th Percentile of the Time-to-Failure Distribution

Under the assumption that the random effect parameter is distributed as half normal(σ^2) or exp(α), we estimate the 50th percentile of the time-to-failure distribution, $t_{0.5}$, for the semi-parametric estimators using formulas (7) and (8), OLS estimator, and ML estimator as follows:

To estimate $t_{0.5}$ using the semi-parametric estimators:

1. Take a random sample, with replacement, of size 15 from the slopes in Table 9.
2. Depending on this sample, obtain $\hat{t}_{0.5\text{-SPH}}$ and $\hat{t}_{0.5\text{-SPE}}$ by solving $\hat{F}_{T\text{-SPH}}(\hat{t}_{0.5\text{-SPH}}) = 0.5$ and $\hat{F}_{T\text{-SPE}}(\hat{t}_{0.5\text{-SPE}}) = 0.5$, respectively.

To estimate $t_{0.5}$ using the ML estimator:

1. Take a random sample, with replacement, of size 15 from the failure times t_i in Table 9.
2. Depending on this sample, obtain $\hat{t}_{0.5\text{-ML}}$ using (12) or (16) according to the assumed distribution.

Table 10. The results of the real data under the assumption $\beta_i \sim$ half normal(σ^2)

Estimator	Bias	MSE	Length of 95% bootstrap CI
Semi parametric half normal	-0.1345	0.4562	2.6621
Semi parametric exponential	-0.1003	0.4302	2.6232
ML	3.6277	13.8605	3.1407
OLS	1.5333	7.7073	7.3704

Table 11. The results of the real data under the assumption $\beta_i \sim \exp(\sigma^2)$

Estimator	Bias	MSE	Length of 95% bootstrap CI
Semi parametric half normal	-0.1345	0.4562	2.6621
Semi parametric exponential	-0.1003	0.4302	2.6232
ML	3.5381	13.1561	3.0727
OLS	4.2186	25.4938	8.9887

To estimate $t_{0.5}$ using the OLS estimator:

1. Take a random sample, with replacement, of size 15 of vectors from the data, where each vector consists of y_{ij} and time t_j for $i = 1, 2, \dots, 15, j = 1, 2, \dots, 16$.
2. Depending on this sample, obtain $\hat{t}_{0.5-OLS}$ using (20) or (22) according to the assumed distribution.

From Tables 10 and 11, the following may be concluded:

- The estimation of the median of the time-to-failure distribution using the semi-parametric exponential estimator has the smallest MSE value and smallest 95% confidence interval length.
- The performance of the semi-parametric half-normal estimator and semi-parametric exponential estimator are comparable.
- The ML and OLS estimators perform poorly compared to the semi-parametric estimators.

Conclusions

When the distribution of the random effect is assumed to be known, the ML and OLS estimators of t_r perform better than the semi-parametric estimators in terms of the MSE and the length of the 95% bootstrap confidence interval. Otherwise, the semi-parametric estimators perform best.

References

Al-Haj Ebrahim, M., Eidous, O., & Kmail, G. (2009). Estimating percentiles of time-to-failure distribution obtained from a linear degradation

SEMI-PARAMETRIC ESTIMATION FOR LINEAR DEGRADATION MODEL

model using the kernel density method. *Communication in Statistics – Simulation and Computation*, 38(9), 1811-1822. doi: 10.1080/03610910903145130

Amster, S. J., & Hooper, J. H. (1983, August). *Accelerated life tests with measured degradation data and growth models*. Paper presented at the American Statistical Association Annual Meeting, Toronto, Canada.

Efron, B., & Tibshirani, R. (1993). *An introduction to the bootstrap*. New York, NY: Chapman and Hall.

Gertsbakh, I. B., & Kordonskiĭ, Kh. B. (1969). *Models of failure* (Scripta Technica Inc., Trans.). New York, NY: Springer-Verlag. (Original work published 1966)

Lu, C. J., Meeker, W. Q., & Escobar, L. A. (1996). A comparison of distribution and failure time analysis methods for estimating a time-to-failure distribution. *Statistica Sinica*, 6(3), 531-546. Available from <http://www.jstor.org/stable/24305606>

Meeker, Q. W., & Escobar, L. A. (1998). *Statistical method for reliability data*. New York, NY: John Wiley and Sons, Inc.

Racine, J. S. & MacKinnon, J. G. (2007). Inference via kernel smoothing of bootstrap *P* values. *Computational Statistics & Data Analysis*, 51(12), 5949-5957. doi: 10.1016/j.csda.2006.11.013

Silverman, B. W. (1986). *Density estimation for statistics and data analysis*. London, UK: Chapman and Hall.