Universidade Federal de Minas Gerais
Instituto de Ciências Exatas
Departamento de Estatística

Empirical Comparisons Between
Kaplan-Meier and Nelson-Aalen
Survival Function Estimators

E. A. Colosimo, F. F. Ferreira, M. D.
Oliveira e C. B. Sousa

Relatório Técnico
RTP-01/2000

Relatório Técnico
Série Pesquisa

Empirical Comparisons Between Kaplan-Meier and Nelson-Aalen Survival Function Estimators

Enrico A. Colosimo[a], Flávio F. Ferreira[a],
Maristela D. Oliveira[a,b] e Cleide B. Sousa[a]

[a]Departamento de Estatística - ICEx-UFMG,
Caixa Postal 702, 30161-970, Belo Horizonte – MG, Brazil

[b]Departamento de Estatística, IM-UFBA
Salvador – BA, Brazil

Abstract

The Kaplan-Meier is the most commonly used estimator of the survival function, while the Nelson-Aalen is an alternative estimator for the same function. There are many asymptotic results for these estimators in the literature. In particular, it is known that they are asymptotically equivalent. On the other hand empirical results comparing these estimators are difficult to obtain and they are necessary to guide applied statisticians. This paper addresses small-sample properties of these survival function estimators. Monte Carlo simulations are performed in order to compare both estimators. Percentile and survival fraction estimates of the survival function are used to attain this goal. The results show a slight superiority in favor of the Nelson-Aalen estimator in survival fraction estimation. However for percentile estimation the Kaplan-Meier estimator presents a better performance for decreasing failure rates while the Nelson-Aalen estimator provides better results for increasing failure rates.

Key Words: Mean square error; percentiles; product-limit estimator; step function; survival fractions; Weibull model.

1. INTRODUCTION

The first step in human lifetime studies, as well as in life-testing experiments in engineering, is usually the estimation of the underlying survival distribution. The standard estimator proposed by Kaplan and Meier (1958) is the most commonly used technique for this task. The Kaplan-Meier, or product-limit estimator (KME) is a step function which has been playing a central role in the analysis of most biomedical studies. However, an alternative estimator, suggested by Nelson (1972) and studied by Aalen (1978), is another choice which is becoming popular among applied statisticians. This estimator is referred to as the Nelson-Aalen estimator (NAE).

Both the Kaplan-Meier and the Nelson-Aalen estimators can be obtained using the theory of counting process, which allows the derivation of their asymptotic properties (Andersen et al., 1993; Fleming and Harrington, 1991). In particular, it has already been proved that they are asymptotically equivalent. Fleming and Harrington (1991, p.99) also proved that the KME has a nonnegative bias which converges to zero as the sample size increases.

Estimates of survival fractions at some specified time can be obtained using both KME and NAE. It is known that NAE is consistently larger than KME (Bohoris, 1994). Klein and Moeschberger (1997, p. 86) state that the Nelson-Aalen cumulative hazard estimator has a better small-sample performance than the estimator based on the KME. Some authors have proposed a linear interpolation, instead of using the step function directly, in order to obtain survival fraction estimates using both estimators (Lee, 1992). It seems that more empirical results are necessary to validate these statements for small-sample situations.

The aim of this paper is two-fold: (1) compare KME and NAE and (2) compare the direct and the linear interpolation methods to obtain estimates for survival fractions using KME and NAE. Monte Carlo simulations are used to attain these objectives by means of estimates of survival fractions and percentiles of the survival distribution. The paper is organized as follows: Section 2 reviews KME and NAE as well as the methods to obtain estimates using these estimators. In Section 3, Monte Carlo simulations are used in order to make the comparisons. The mean square error is the loss function used in the comparisons. Section 4 presents some general conclusions.

2. SURVIVAL FUNCTION ESTIMATORS

Let $t_1 < t_2 < \ldots < t_k$ represent the observed death times in a sample of n subjects from a homogeneous population with survival function $S(t)$. Consider $S(t)$ as a discrete function with probability mass at each $t_i; i =$

$1, \ldots, k$. Therefore it can be written that

$$S(t_i) = (1 - q_1)(1 - q_2)\ldots(1 - q_i) = \prod_{j=1}^{i}(1 - q_j), \tag{1}$$

where q_j is the probability of subject death in the interval $[t_{j-1}, t_j)$ conditional of being alive at t_{j-1}, that is, q_j can be written as

$$q_j = P(T \in [t_{j-1}, t_j)/T \geq t_{j-1}). \tag{2}$$

Suppose that d_i deaths occurs at t_i and there are n_i subjects at risk at t_i, KME is obtained from (1) and (2) as

$$\widehat{S}_{KM}(t) = \prod_{i/t_i<t}\left(\frac{n_i - d_i}{n_i}\right) = \prod_{i/t_i<t}\left(1 - \frac{d_i}{n_i}\right). \tag{3}$$

Another way of expressing the survival function is

$$S(t) = \exp(-H(t)), \tag{4}$$

where $H(t)$ is the cumulative hazard function. Expression (4) suggests that the estimation of $S(t)$ could also be based on $H(t)$. The Nelson-Aalen estimator of $H(t)$ is given by

$$\widehat{H}_{NA}(t) = \sum_{i/t_i<t}\left(\frac{d_i}{n_i}\right). \tag{5}$$

Therefore, the NAE of $S(t)$ is

$$\widehat{S}_{NA}(t) = exp(-\widehat{H}_{NA}(t)).$$

The Kaplan-Meier and Nelson-Aalen estimators are plotted as a step function since they remain constant within two consecutive observed survival times. These plots are useful to estimate survival fractions at time $t^*(\widehat{S}(t^*))$ and percentiles $\widehat{t}_p$ $(\widehat{S}(\widehat{t}_p) = 1 - p)$. Percentile estimates are not unique as the estimators are step functions. A practical solution advocated by some authors, including Lee (1992), is to make a linear interpolation between two steps. That is, points are connected by straight lines, and so the percentiles are obtained. However, there are no results that guarantee the adequacy of this procedure.

3. SIMULATION STUDY

In this section, Monte Carlo simulations are performed to compare the behaviour of the estimators. Comparisons are made for survival fraction and percentile estimates. The simulation study is based on a Weibull distribution with shape parameter (δ) equal to $0.5, 1, 2$ corresponding to decreasing, constant and increasing failure rates, respectively. The scale parameter of the Weibull distribution is set equal to 1.

A set of independent random variables $\mathbf{T}^{'} = (T_1, ..., T_n)$ from a Weibull distribution $(1, \delta)$ is generated for each repetition and type II censoring mechanism is used in order to generate lifetimes. 1000 replications are run for each simulation. Simulations are performed for some sample sizes ($n = 10, 20, 50$) and the proportion of censoring is set equal to 30% in each sample. Simulation sample means aa well as the mean square error (MSE) of the percentile and survival fraction estimates are used to compare KME and NAE. Direct and linear interpolation methods are used to compare survival fraction estimates and they are referred to as direct and interpolation respectively in the simulation results. Linear interpolation is the only method used to estimate percentiles in order to guarantee the uniqueness of the estimates.

Codes were built in C^{++} language to run the simulations. Figures 1 to 12 display the simulation results. Figures 1 to 6 display the results for survival fraction estimates (bias and MSE), while Figures 7 to 12 display the results for percentile estimates (bias and MSE). C^{++} codes are available upon request from the second author.

4. DISCUSSION

Some conclusions can be drawn from the simulation studies presented in Section 3:

1. The estimators have neglible bias for $n = 50$ but a relatively large bias for $n = 10$ and 20 in both methods.

2. **Survival Fractions:** The interpolation method seems to be better than the direct one. In general, the NAE using the interpolation method achieved the smallest MSE. In other cases, the NAE also has smaller MSE than the KME for both methods.

3. **Percentile Estimates:** The KME seems to be better than the NAE when $\delta \leq 1$, since it presents smaller EQM and bias than the NAE for this situation. On the other hand, the NAE seems to be better than the KME for $\delta > 1$.

These results are a limited simulation study based on the Weibull distribution. A small-scale expansion of this simulation was also carried out, considering other values for the scale parameter of the Weibull $(0.5, 2)$ and for the proportion of censoring $(0, 50\%)$. The results are very similar to those reported in Section 3 although they are not presented in this paper. Other values for t were also used to calculate the survival fraction estimates. In general, the results are similar to those presented in Section 3, but for very small or very large survival fractions and a small sample size. In the latter situation, it is expected a great instability in the method since it uses very few observations to estimate in the tails of the distribution.

In general, it seems that the interpolation method works well to estimate survival fractions. As a matter of fact it works better than the direct method. It seems that the NAE is better than the KME to obtain survival fraction estimates, especially when using the interpolation method. However, the KME and the NAE have different behaviours to estimate percentiles of the survival distribution. The KME is better than the NAE for decreasing failure rates and the NAE is better than the KME for increasing failure rates.

ACKNOWLEDGEMENTS

We thank Prof. Frederico R. B. Cruz for motivating us in the construction of the C^{++} code.

References

Aalen, O. O. (1978) Nonparametric inference for a family of counting processes *Annals of Statistics*, **6**, 701-726.

Andersen, P. K., Borgan, O., Gill, R. D. and Keiding, N. (1993) *Statistical Models based on Counting Processes* Springer-Verlag, New York.

Bohoris, G. A. (1994) Comparison of the cumulative-hazard and Kaplan-Meier estimators of the survivor function *IEEE Transactions on Reliability*, **43**, 230-232.

Fleming, T. R. and Harrington, D. P. (1991). *Counting Processes and Survival Analysis.* John Wiley, New York.

Kaplan, E. L. and Meier, P. (1958). Nonparametric estimation from incomplete observations *Journ. of the American Statistical Association*, **53**, 457-481.

Klein, J. P. and Moeschberger, M. L. (1997). *Survival Analysis: Tecniques for Censored and Truncated Data* Springer-Verlag, New York.

Kalbfleisch, J. D. and Prentice, R. L. (1980). *The Statistical Analysis of Failure Time Data.* John Wiley, New York.

Lee, E. T. (1992). Statistical Methods for Survival Data Analysis. John Wiley, New York.

Nelson, W. (1972). Theory and application of hazard plotting for censored survival data. *Biometrics*, **14**, 945-966.

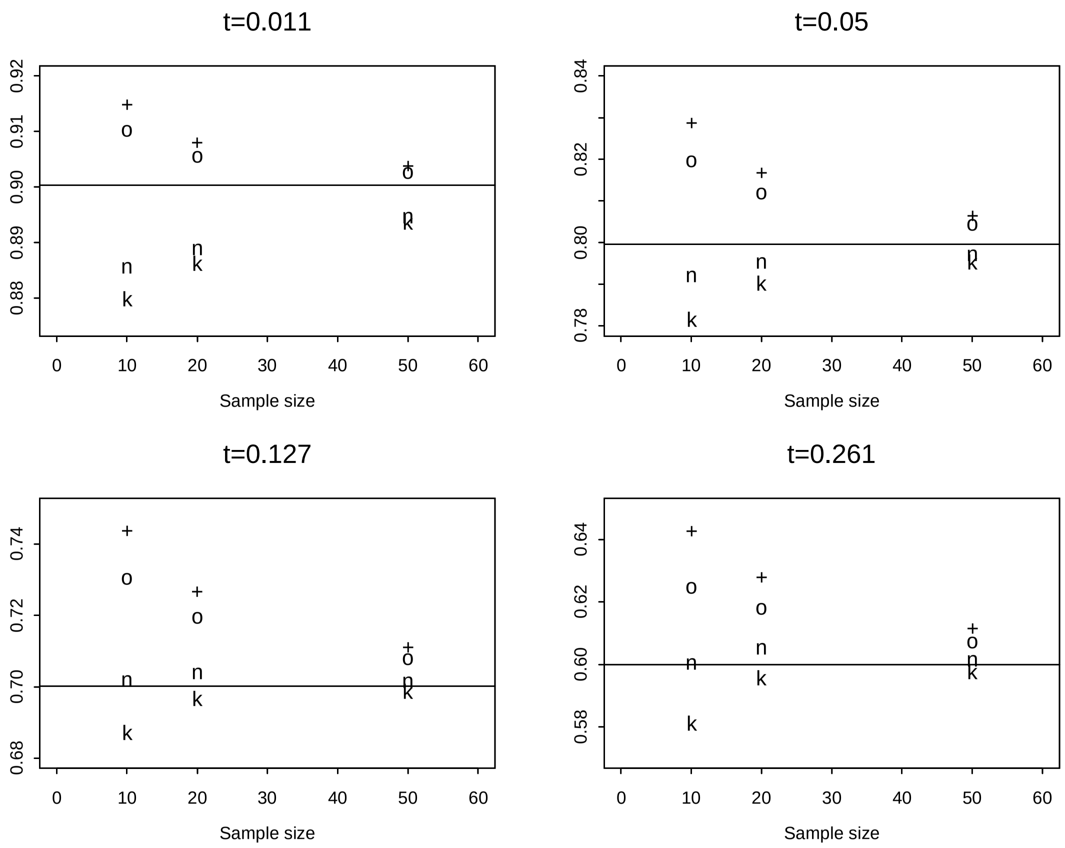

Figure 1: Survival fraction estimates at *t* for a Weibull distribution with δ=0.5, comparing NAE-direct method (+), KME-direct method (o), NAE-interpolation method (n), KME-interpolation method (k) and the true value (—).

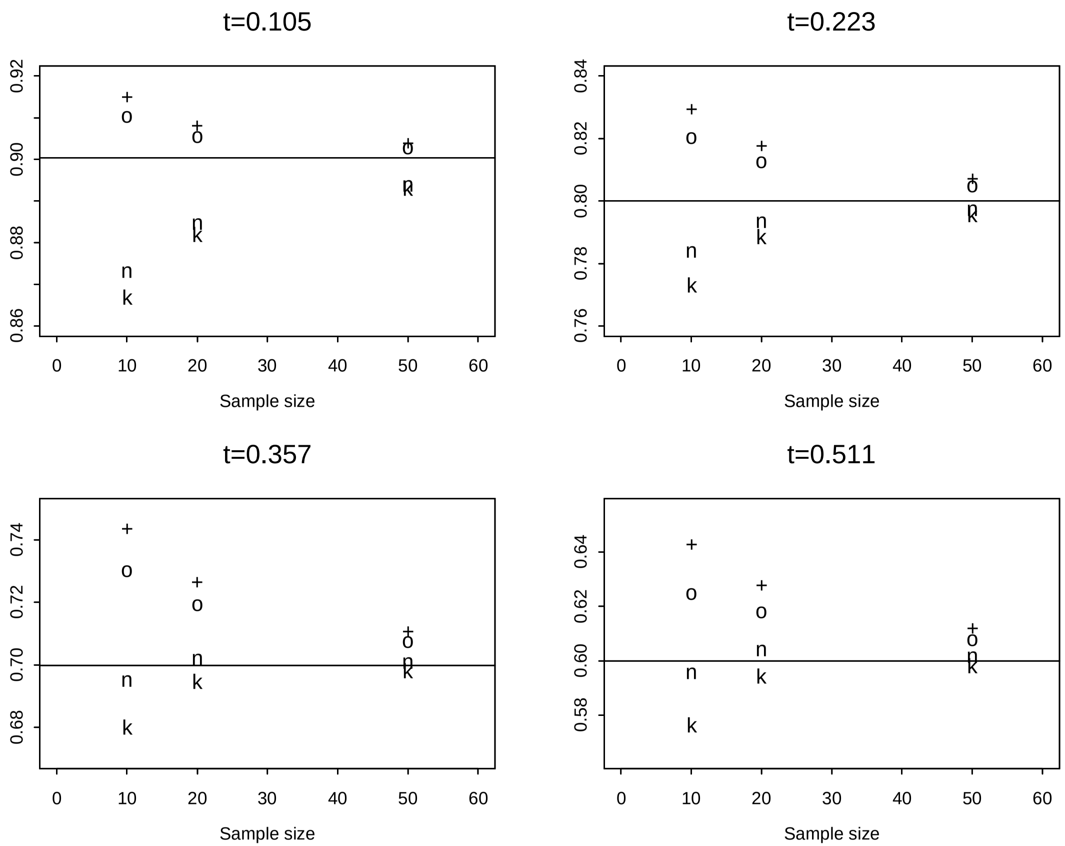

Figure 2: Survival fraction estimates at *t* for a Weibull distribution with δ=1.0, comparing NAE-direct method (+), KME-direct method (o), NAE-interpolation method (n), KME-interpolation method (k) and the true value (—).

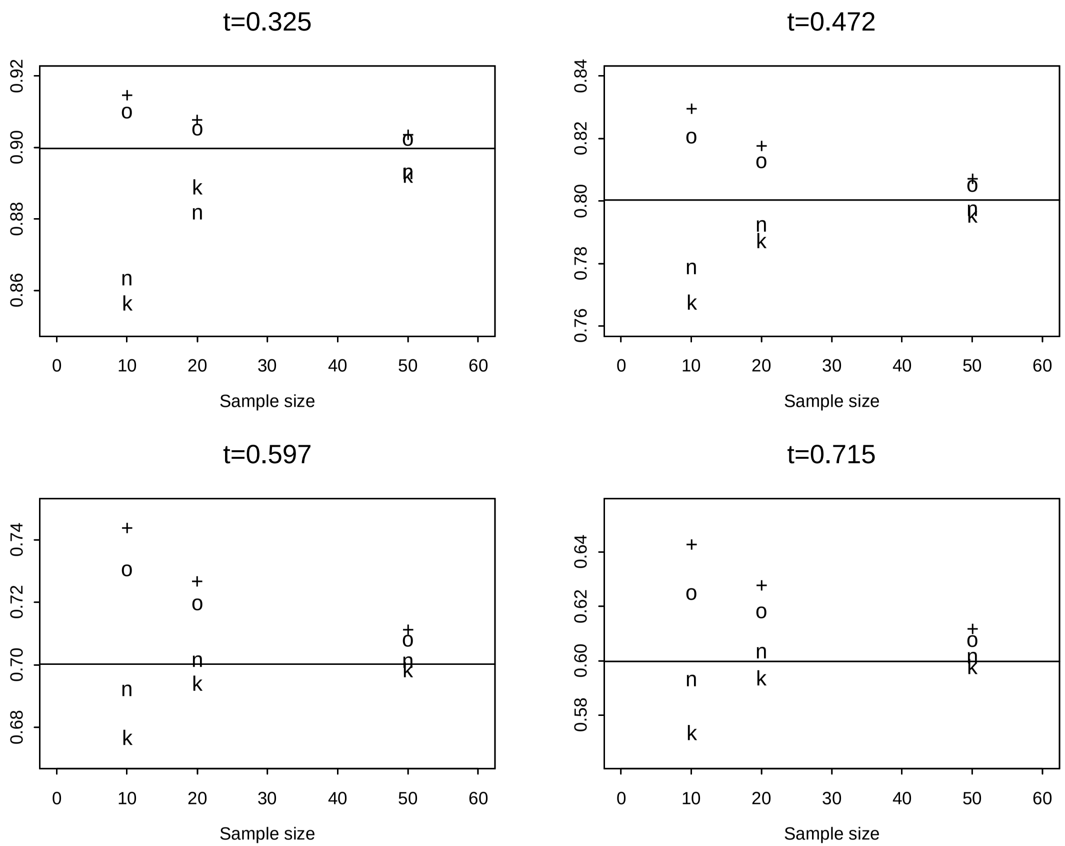

Figure 3: Survival fraction estimates at *t* for a Weibull distribution with δ=2.0, comparing NAE-direct method (+), KME-direct method (o), NAE-interpolation method (n), KME-interpolation method (k) and the true value (—).

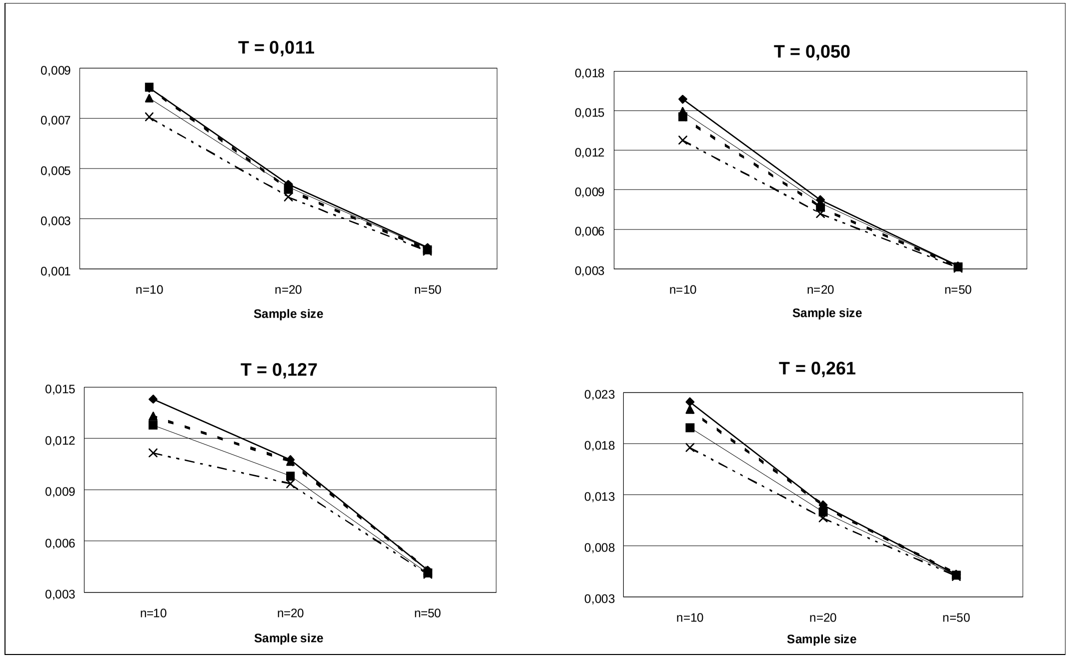

Figure 4: Mean Square Error for Weibull Survival with δ = 0.5, comparing KME-direct(—◆—), NAE-direct(- - ▲- - -), KME-interp(—■—) and NAE-interp(---✕---)

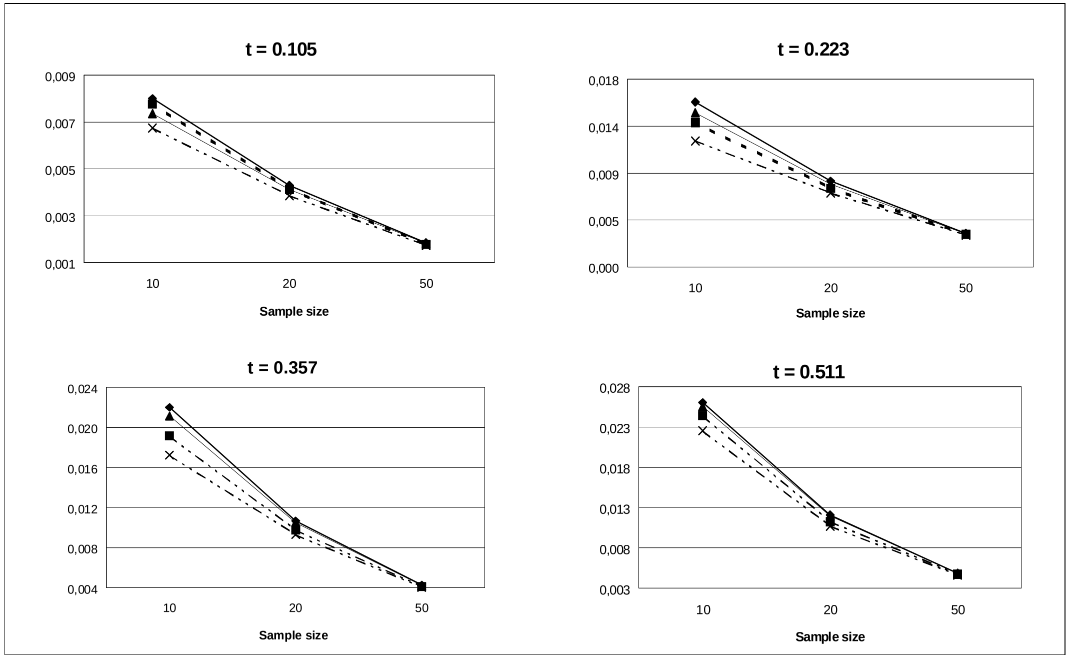

Figure 5: Mean Square Error for Weibull Survival with δ = 1, comparing KME-direct(—◆—), NAE-direct(- – -▲- – -), KME-interp(—■—) and NAE-interp(---✕---)

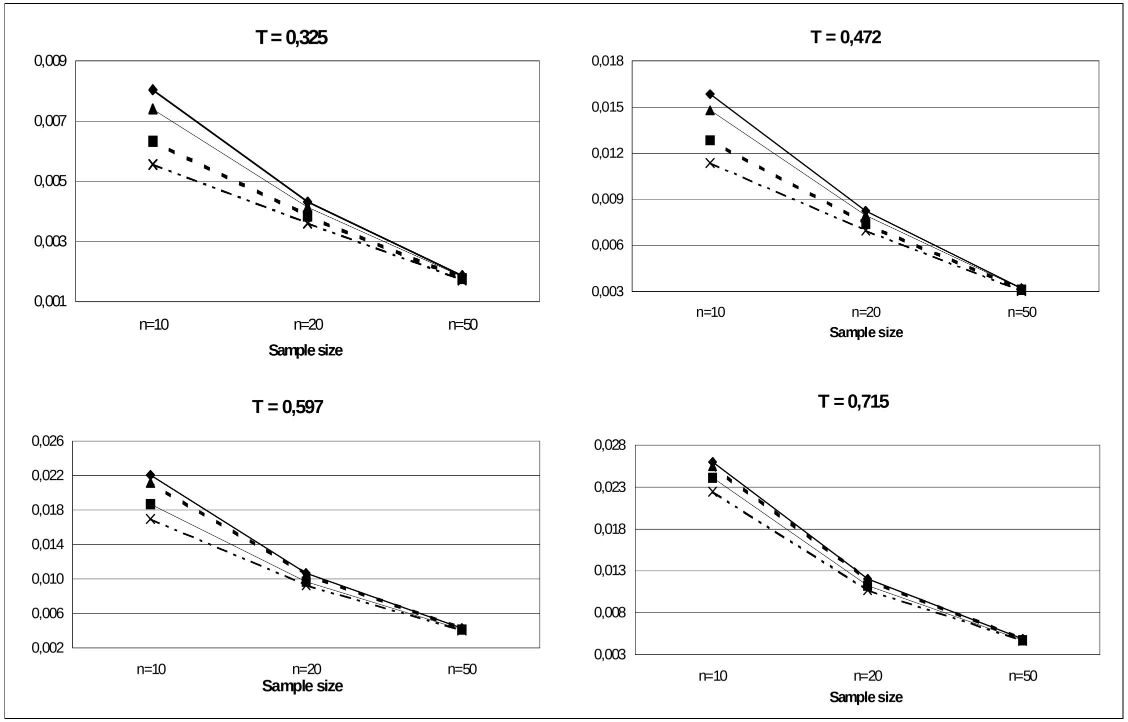

Figure 6: Mean Square Error for Weibull Survival with δ = 2, comparing KME-direct(—◆—), NAE-direct(- - ▲- - -), KME-interp(—■—) and NAE-interp(---×---)

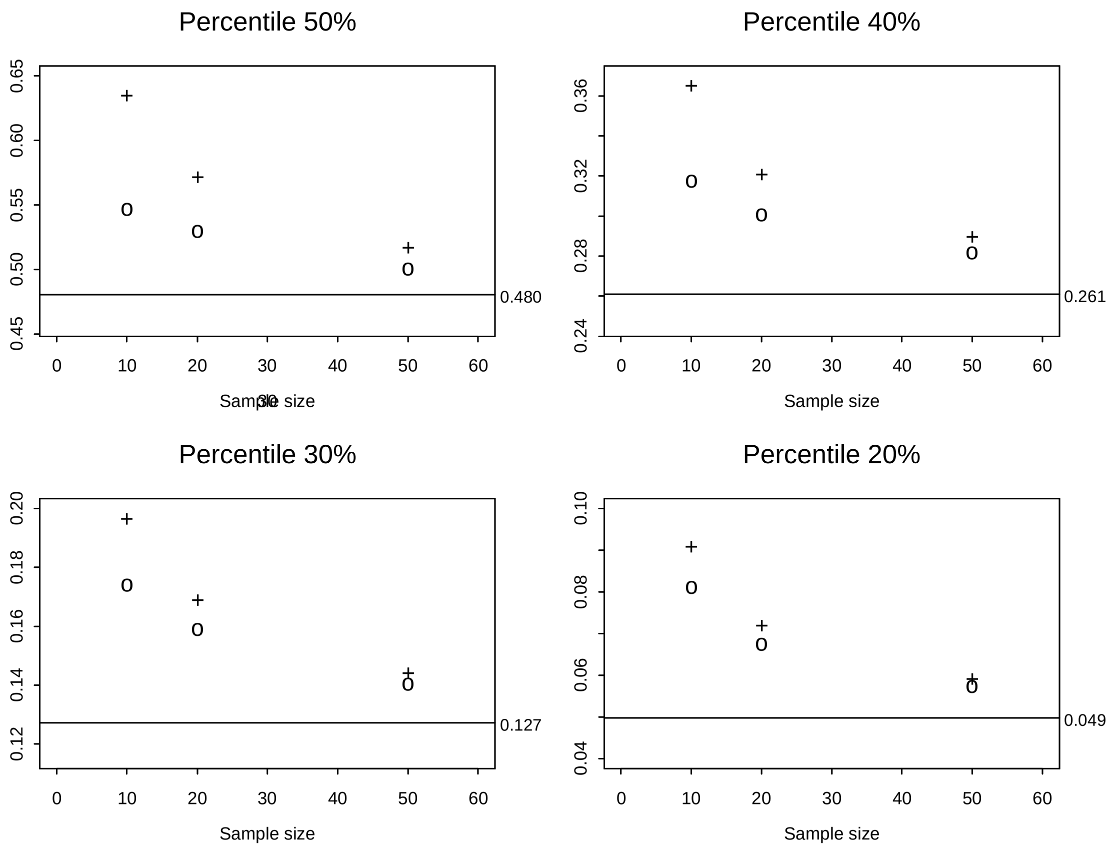

Figure 7: Percentile estimates for a Weibull distribution with δ=0.5, comparing NAE (+), KME (o), and the true value (—).

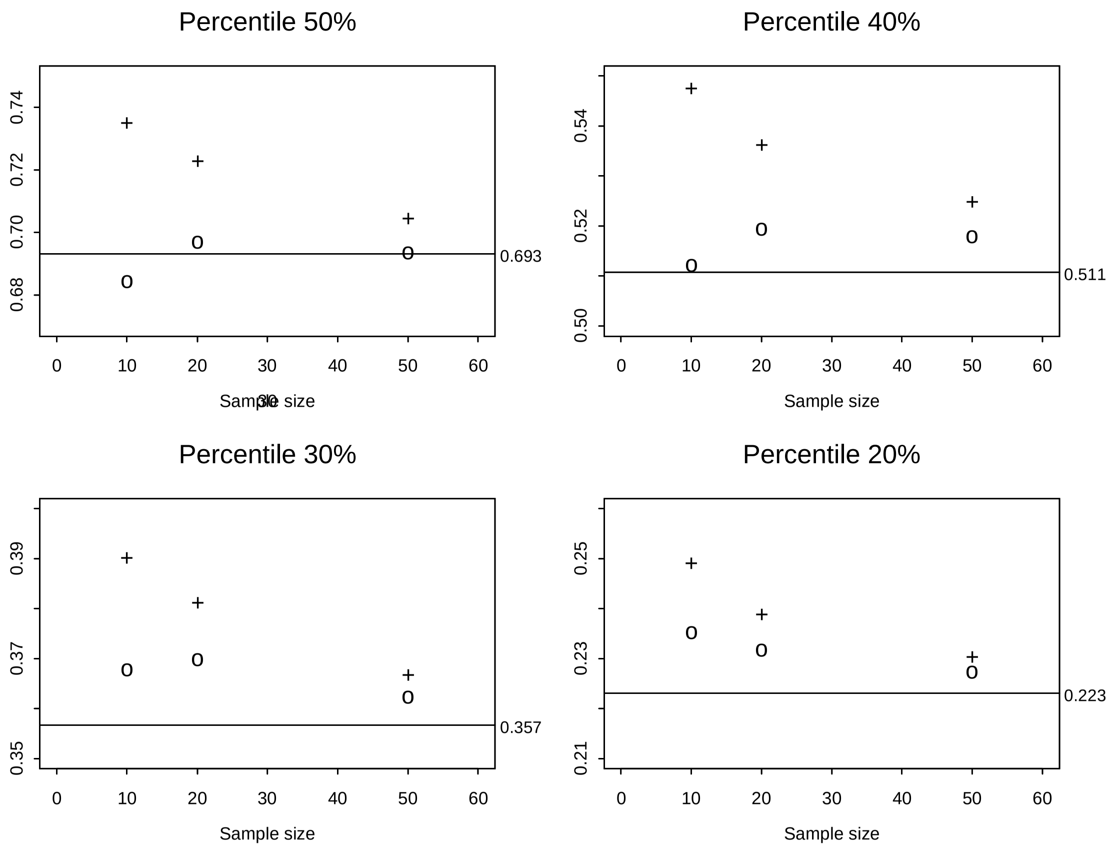

Figure 8: Percentile estimates for a Weibull distribution with δ=1.0, comparing NAE (+), KME (o), and the true value (—).

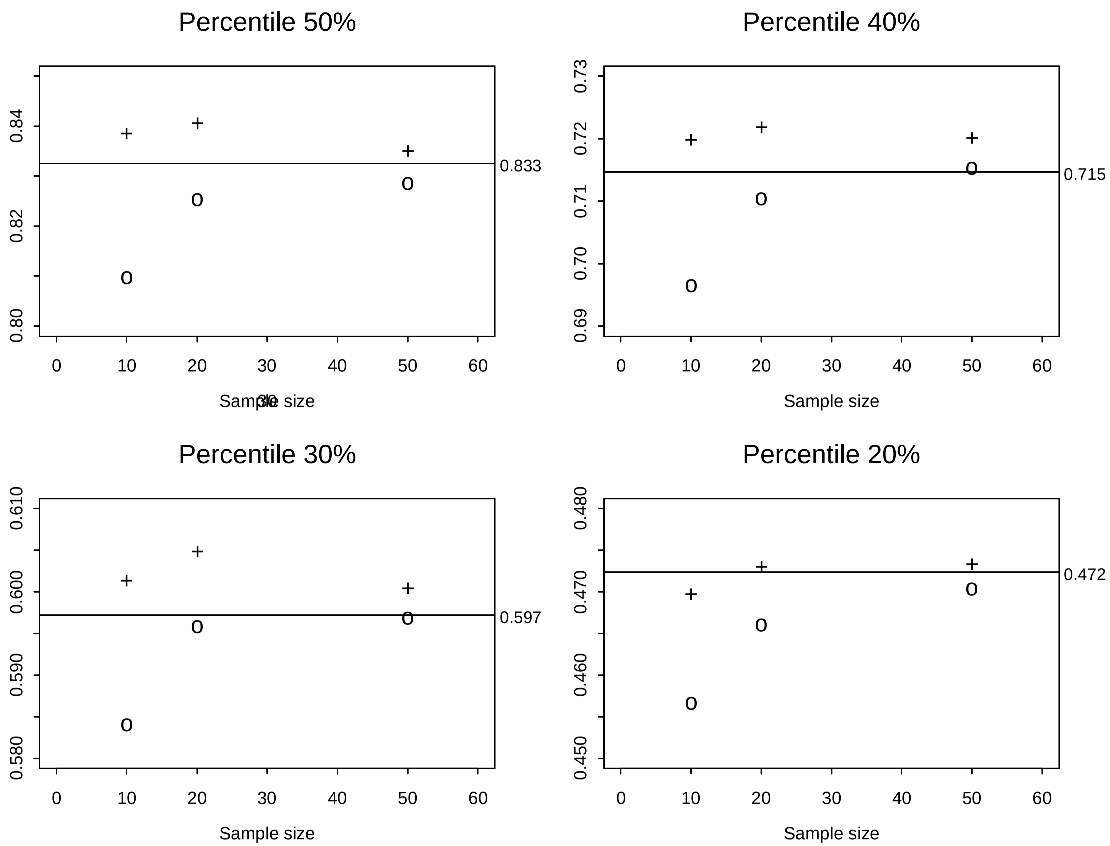

Figure 9: Percentile estimates for a Weibull distribution with δ=2.0, comparing NAE (+), KME (o), and the true value (—).

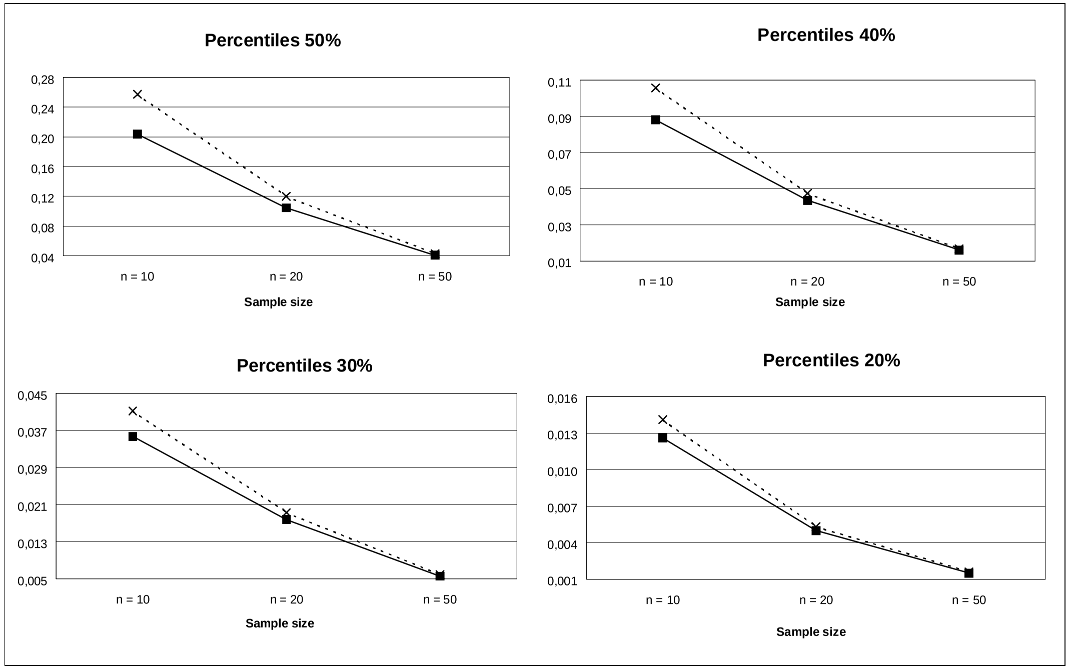

Figure 10: Mean Square Error for Weibull Survival with $\delta = 0.5$, comparing KME-interp(—■—) and NAE-interp(---X---)

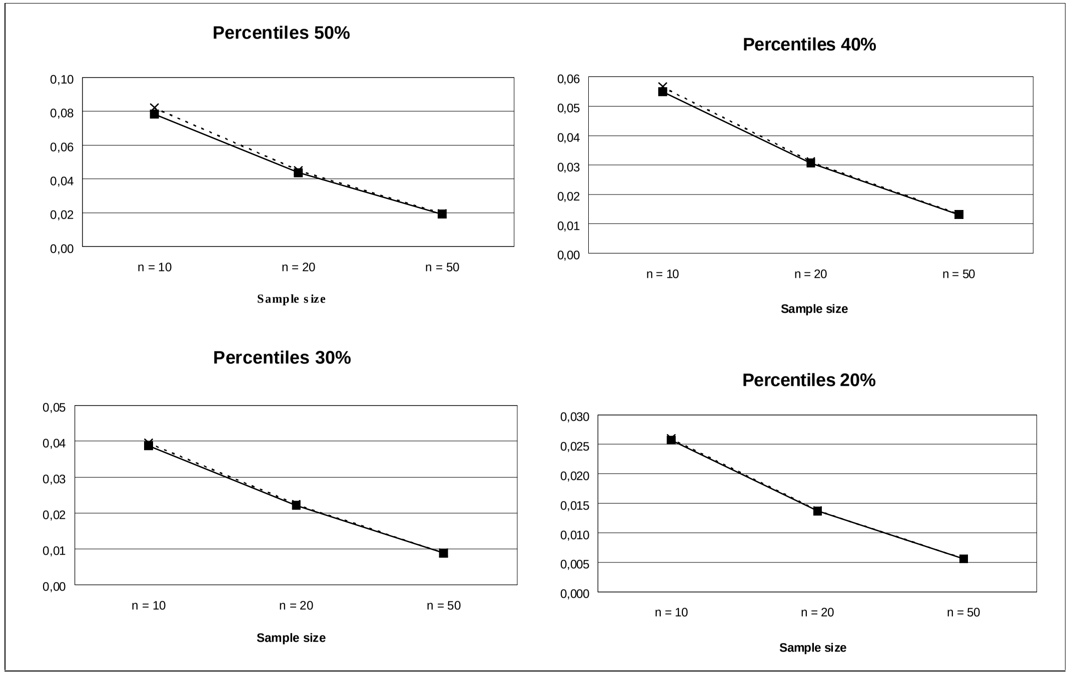

Figure 11: Mean Square Error for Weibull Survival with δ = 1, comparing NAE-interp (—■—) and KME-interp (---X---)

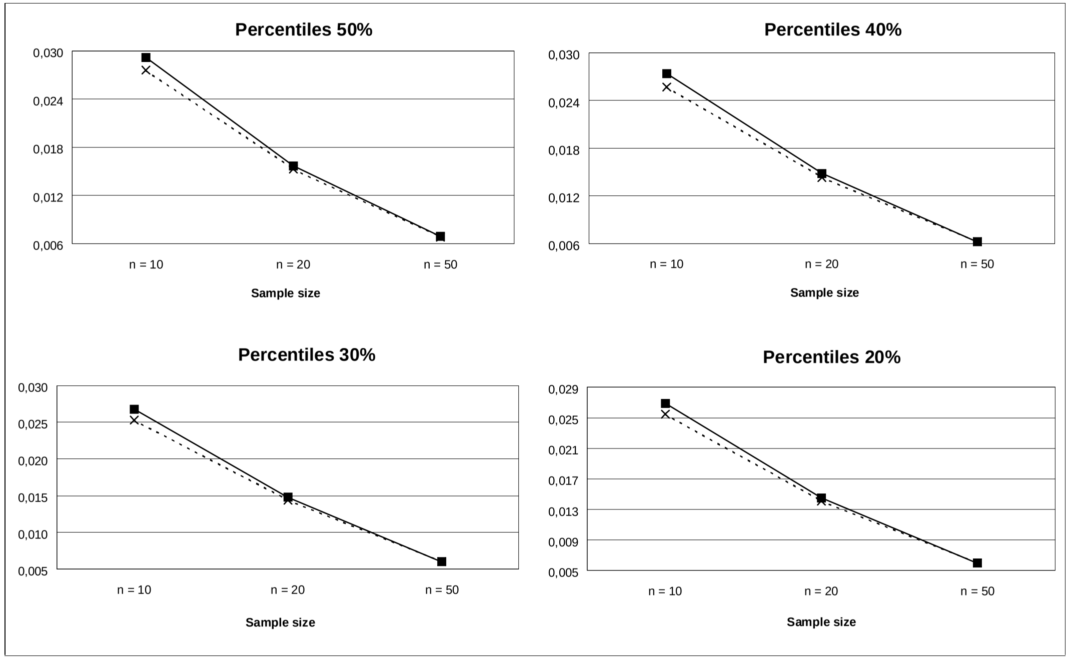

Figure 12: Mean Square Error for Weibull Survival with δ = 2, comparing KME-interp(—■—) and NAE-interp(---X---)