



Research Report  
Statistical Research Unit  
Department of Economics  
University of Gothenburg  
Sweden

---

**Relative Efficiency of a Quantile Method  
for Estimating Parameters in Censored  
Two-Parameter Weibull Distributions**

**Jonsson, R.**

**Research Report  
2010:3  
ISSN 0349-8034**

Mailing address:  
Statistical Research Unit  
P.O. Box 640  
SE 405 30 Göteborg  
Sweden

Fax  
Nat: 031-786 12 74  
Int: +46 31 786 12 74

Phone  
Nat: 031-786 00 00  
Int: +46 31 786 00 00

Home Page:  
<http://www.statistics.gu.se/>

**Relative Efficiency of a Quantile Method for Estimating Parameters  
in Censored Two-Parameter Weibull Distributions**

**Robert Jonsson**

## **Summary**

In simulation studies the computer time can be much reduced by using censoring. Here a simple method based on quantiles (Q method) is compared with the Maximum Likelihood (ML) method when estimating the parameters in censored two-parameter Weibull distributions. The ML estimates being obtained using the SAS procedure NLMIXED. It is demonstrated that the estimators obtained by the Q method are less efficient than the ML estimators, but this can be compensated for by increasing the sample size which nevertheless requires much less computer time than the ML method. The ML estimates can only be obtained by an iterative process and this opens the possibility for failures in the sense that reasonable estimates are presented as unreliable, or anomalous estimates are presented as reliable. Such anomalies were never obtained with the Q method.

Key words:

## 1. Introduction

Simulation studies with long realizations require computers of high capacity. A typical example is when one is estimating the expectation of ‘times-between-false alarms’ in a surveillance situation. The means of the latter times may vary substantially from one realization to another (cf. p. 14 in Jonsson, 2010) and it is apparent that extremely long realizations are needed in order to get stable results from the simulations. Alternatively one may estimate the expectation of ‘times to *first* false alarm’ (measured from an arbitrary time point) in a large number of realizations. In this way stable results are achieved with much fewer simulations. However, also in this case computer space may be insufficient. A radical way of reducing the computer time further is to simply cut off the distribution of times, so only observed times to first alarm below a certain limit are observed, while the rest of the observations are counted but not observed. Such data are called censored (Kruskal and Tanur, 1978).

This article was motivated by the author’s hardships when simulating ‘times to first false alarm’ for a modified CUSUM method that is presented in Jonsson, 2010. Due to limited computer resources it was decided that censoring should be used. In the present case it luckily turned out that the distribution could be well approximated by a two-parameter Weibull distribution. There is an extensive literature about how to estimate parameters in this case. The problem is that the latter are intended for specific situations with a relative small sample size. When dealing with huge sets of simulated data it was found that these methods sometimes failed and a great deal of time was spent to monitoring the results. Therefore a new simpler method was used that only utilizes information about quantiles in the censored samples and about the proportion of missing observations due to censoring. The expectation, or other summarizing measures, of times to first false alarm can then be estimated by inserting parameter estimates into the theoretical expressions. This method will be termed the Quantile (Q) method.

The Q method should give less precise estimates than more sophisticated methods such as Maximum Likelihood (ML). Therefore this paper focuses on comparisons of the efficiency of Q and ML estimates. The approach is entirely pragmatic and aims to supply researchers with a simple method that never seems to fail, and which can be used when computer resources are limited.

## 2. Quantile and Maximum Likelihood estimation of the parameters in the censored two-parameter Weibull distribution

The Weibull probability law has the following survival function  $S(y)$ , distribution function  $F(y)$  and density  $f(y)$

$$S(y) = \exp\left\{-\left(y/\lambda\right)^\theta\right\}, F(y) = 1 - S(y), f(y) = (\theta/\lambda) \cdot (y/\lambda)^{\theta-1} \exp\left\{-\left(y/\lambda\right)^\theta\right\} \quad (1)$$

Here, the parameters  $\theta > 0$  and  $\lambda > 0$  determine slope and scale, respectively. When  $\theta > 1$  the density has one single peak, while for  $0 < \theta \leq 1$  it is strictly decreasing. For  $\theta = 1$  one gets the exponential density with  $f(0) = 1/\lambda$ . For  $\theta < 1$  the major part of the mass of the density becomes more concentrated close to the  $x$  and  $y$  axis as  $\theta$  decreases and it is also seen that  $f(y) \rightarrow \infty$  as  $y \rightarrow 0$ . The median  $y_{.50}$  and expected value  $E$  are

$$y_{.50} = \lambda \cdot (\ln 2)^{1/\theta} \text{ and } E = \lambda \cdot \Gamma(1 + 1/\theta) \quad (2)$$

All these results presuppose that no observations are lost due to censoring.

Focus will be on the case  $\theta \leq 1$  since the times to first false alarm suggested a strictly decreasing density with a value of  $\theta$  much less than 1. The case  $\theta = 1$  is considered merely for curiosity.

Let  $y = C(p_{miss})$  be the point at which  $100 \cdot p_{miss}$  % of the observations are lost due to censoring (censoring point). When the value of  $p_{miss}$  is obvious the notation  $y = C$  will be used for simplicity. The relation between  $p_{miss}$  and  $C(p_{miss})$  is illustrated in Table 1 for  $p_{miss} = 0.1, \dots, 0.9$   $\theta = 0.5, 2/3, 1$  and  $y_{.50} = 20$ . The censoring points that are obtained for different values of  $\theta$  reflect the differences in shape of the densities. E.g. for  $\theta = 1$  a censoring point at 66 gives a missing proportion of 0.1, but for  $\theta = 0.5$  one has to choose a censoring point at 221 to get the same missing proportion.

**Table 1** Relation between  $p_{miss}$  (proportion censored observations) and  $C(p_{miss})$  (censoring point) for three values of  $\theta$ . In all cases the median  $y_{.50}$  is 20.

$p_{miss}$	$C(p_{miss})$		
	$\theta = 0.5$	$\theta = 2/3$	$\theta = 1$
0.1	221	121	66
0.2	108	71	46
0.3	60	46	35
0.4	35	30	26
0.5	20	20	20
0.6	11	13	15
0.7	5.3	7.4	10
0.8	2.1	3.7	6.4
0.9	0.5	1.2	3.0

### 2.1 Estimation from censored data based on quantiles (*Q method*)

The 100 $p$  % percentile when censoring at  $y = C(p_{miss}) = C$  is denoted  $y_p(C)$ . Then obviously

$$p = P(Y < y_p(C) | Y < C) = \frac{F(y_p(C))}{1 - p_{miss}}, \text{ or } S(y_p(C)) = 1 - p(1 - p_{miss}).$$

Since  $S(y_p(C)) = \exp\left\{-\left(y_p(C)/\lambda\right)^\theta\right\}$  one finally gets

$$\ln(y_p(C)) = \ln \lambda + \frac{1}{\theta} \cdot \ln\left[-\ln(1 - p(1 - p_{miss}))\right], \text{ or } z = \alpha + \beta \cdot x \quad (3)$$

Eq. (3) is a linear relation between the independent variable  $x$  and the dependent variable  $z$ . In each sample  $p$  takes a number of predetermined fixed values,  $p_{miss}$  takes one fixed value that varies from sample to sample but with small variance, and  $z$  is a random response. This suggests that one may use ordinary least square method for estimating  $\alpha$  and  $\beta$ , which in turn give estimates of  $\theta$  and  $\lambda$ . In Appendix A a computer program is presented that first simulates 5 censored samples each with 2000 Weibull distributed random numbers, and then prints the results from the estimation when  $p$  takes the values 0.25, 0.50 and 0.75. The latter values of  $p$  and the small number of replicates are only chosen for illustration.

It is far from obvious how the values of  $p$  in Eq. (3) are to be chosen in order to get estimates that are optimal in terms of bias and variance. Intuitively one might suspect that a large number of spread values of  $p$  should be chosen, since then  $SS_x$  in the expression  $V(\hat{\beta}) = \sigma^2 / SS_x$  is increased. But, in the latter case also the residual variance in the numerator is increased. To find a solution of the problem, simulations were performed using a large number of various sets of values of  $p$ . A typical result is illustrated in Table 2 which shows the standard deviations (stds) of the  $\theta$  and  $\lambda$  estimates for various values of  $p_{miss}$ .

**Table 2** Standard deviations of quantile estimates of  $\theta$  and  $\lambda$  obtained in various settings. Each reported figure is based on 4000 simulated samples each with  $n = 500$ .

$C(p_{miss})$	$p_{miss}$	Values of $p$ that are used for obtaining quantile estimates							
		.25, .50, .75		.10, .50, .90		.10, .25, .50, .75, .90		.01, .05, .10, .25, .50, .75, .90, .95, .99	
		Std( $\hat{c}$ )	Std( $\hat{\lambda}$ )	Std( $\hat{c}$ )	Std( $\hat{\lambda}$ )	Std( $\hat{c}$ )	Std( $\hat{\lambda}$ )	Std( $\hat{c}$ )	Std( $\hat{\lambda}$ )
220.7	0.1	.031	4.6	.027	4.3	.013	4.2	.034	4.1
60.34	0.3	.037	6.3	.034	5.3	.016	5.5	.042	5.9
20.0	0.5	.046	9.8	.042	8.5	.020	8.4	.051	11
5.30	0.7	.062	22	.056	18	.026	17	.069	27
0.46	0.9	.119	1018	.105	657	.049	232	.124	3752

As expected the stds increased with increasing  $p_{miss}$ . Perhaps more interesting is that the stds decreased as the number of  $p$ -values in each set increased, but only up to a certain point. In fact, a set of  $p$ -values with just one or two components resulted in very large stds (not shown in the table). By gradually increasing the number of  $p$ -components up to five it was possible to reduce the stds, but for a larger number of  $p$ -components the stds started to increase again. The set consisting of five values of  $p$  (0.10, 0.25, 0.50, 0.75, 0.90) turned out empirically to give the smallest std among all sets that were considered.

Estimates of  $\theta$  and  $\lambda$  obtained by the Q method are in the sequel denoted  $\hat{\theta}_Q$  and  $\hat{\lambda}_Q$ , respectively. For each set of parameters, 2000 replicates were generated (this number turned out

to be un-necessary large). For the largest sample size of  $n = 3000$  the total run-time at the computer was about 20 minutes.

## 2.2 ML estimation from censored data

Let  $m$  be the number of observations that are lost due to censoring and put  $n^* = n - m$ , the ‘efficient sample size’. The Likelihood is (cf. Cohen, 1965)

$$L = const \cdot \prod_{i=1}^{n^*} (\theta / \lambda) (y_i / \lambda)^{\theta-1} \cdot \exp \left\{ - (y_i / \lambda)^\theta \right\} \cdot \left( \exp \left\{ - (C / \lambda)^\theta \right\} \right)^m \quad (4)$$

On differentiating  $\ln L$  with respect to  $\theta$  and  $\lambda$  and equating to zero one gets the estimating equations

$$\frac{\sum_{i=1}^{n^*} y_i^\theta \ln y_i}{\sum_{i=1}^{n^*} y_i^\theta} - \frac{1}{\theta} = \frac{\sum_{i=1}^{n^*} \ln y_i}{n^*}, \quad (5)$$

$$\lambda = \left( \frac{\sum_{i=1}^{n^*} y_i^\theta}{n^*} \right)^{1/\theta}$$

where

$$\sum_{i=1}^{n^*} y_i^\theta = \sum_{i=1}^{n^*} y_i^\theta + m \cdot C^\theta, \quad \sum_{i=1}^{n^*} y_i^\theta \ln y_i = \sum_{i=1}^{n^*} y_i^\theta \ln y_i + m \cdot C^\theta \ln C$$

Here the challenge is to get a solution of  $\theta$  from the first equation in Eq. (5). For a specific data set the solution of  $\theta$  can be obtained by a variety of methods that use graphical tools (see e.g. Cohen, 1965), but in simulation studies where a large number of solutions are required one has to look for other alternatives. Mao and Li (2007) prove some properties of the structure in Eq. (5) that guaranties feasible solutions and propose some general algorithms, but at present these results seem not yet to have been implemented in available computer programs.

As an alternative one may try to maximize  $L$  directly by using standard procedures. One such is the SAS-procedure *NLMIXED* which can be used to find the ML solutions in a general class of



distributions. An example of this procedure is given in Appendix B. In the example given in the Appendix the procedure succeeded to find the estimates iteratively. But sometimes the procedure may fail in the sense that in the print-out it is stated that either (a) a ML solution can not be reached although the solution is acceptable, or (b) a ML solution has been reached although it is far from correct. Examples of (a) and (b) are given in Appendix B.

ML estimates are denoted  $\hat{\theta}_{ML}$  and  $\hat{\lambda}_{ML}$ . The latter were obtained by first running *NLMIXED* with 100 replicates and then calculating sequential averages of estimated means and stds until stability was reached. In a few occasions a further set of 100 replicates was needed to reach stability. Anomalous solutions such as (a) and (b) mentioned above were omitted. The run-time for generating the ML estimates from 100 replicates with  $n = 3000$  was about 5 minutes. However, to this one has to add the time for editing in the print-out and checking for anomalies. The total run time was in fact longer than for the Q method.

### 3. Results

#### 3.1 Bias of estimators

The bias of the Q and ML estimators were roughly the same. It was largest for  $p_{miss} = 0.90$ , but for  $p_{miss} = 0.60$  and smaller the bias fluctuated around zero. No substantial reduction in bias was noticed when  $n$  increased from 500 to 3000. The magnitude of the biases was constantly only 5-10 % of the stds, so very little was changed by measuring precision by Mean Squared Error instead of variance of the estimators. For this reason only variances were considered when precision was compared.

### 3.2 Checking that variances of estimators are proportional to $n^{-1}$

The relative efficiency (*RE*) of an unbiased estimator  $T_1$  based on the sample size  $n_1$  relative to another  $T_2$  based on the sample size  $n_2$  is defined as  $RE(n_1, n_2) = V(T_1)/V(T_2)$ . If both variances are proportional to  $n_i^{-1}$ ,  $RE(n_1, n_2) = C \cdot (n_2 / n_1)$  and the value of  $C$  can be used to find the sample sizes needed to get  $V(T_1) = V(T_2)$ . For instance  $RE = 1/2$  implies that  $n_2 = 2n_1$  will give equal variances. The ML estimators obtained from Eq. (5) are proportional to  $n^{-1}$  in large samples, but this only holds asymptotically (cf. e.g. p. 244 in Zacks, 1971). It is far from obvious that  $n$  in the range 500-3000 is large enough. Also the quantile estimators that are obtained from the regression estimators in Eq. (3) may have variances that are not proportional to  $n^{-1}$ .

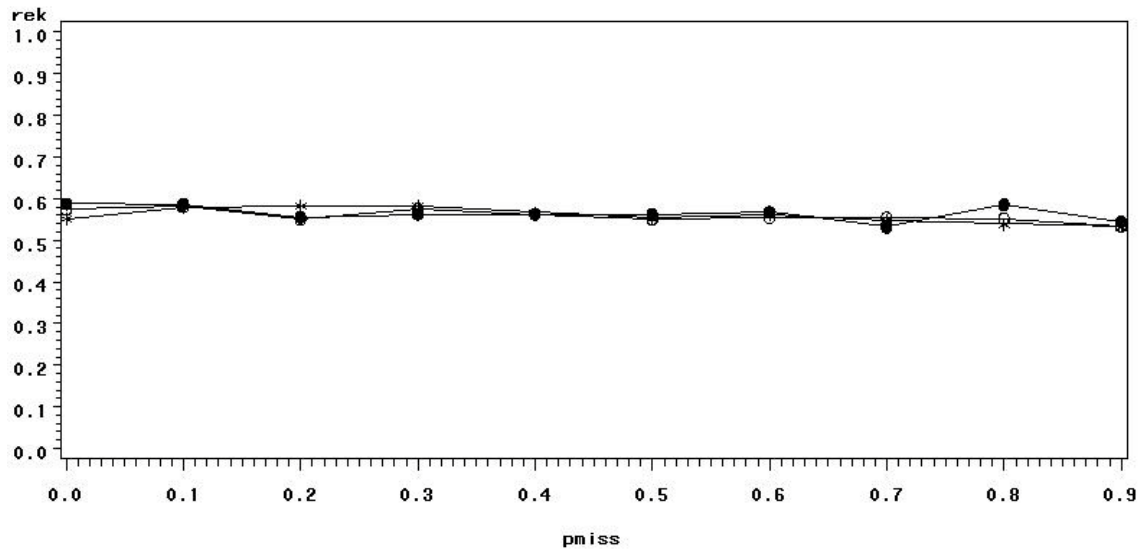
To check the proportionality to  $n^{-1}$  of  $V(T)$  one may study whether  $nV(T)$  as a function of  $n$  remains constant, for example by fitting a straight line to the data points and estimating the slope. This was done for  $T = \hat{\theta}_Q, \hat{\theta}_{ML}, \hat{\lambda}_Q, \hat{\lambda}_{ML}$ ,  $\theta = 0.5, 2/3, 1$ ,  $q_{.50} = 20$  and  $p_{miss} = 0, 0.1, \dots, 0.9$ . For each of the four estimators there were thus 30 fitted lines to the points  $nV(T)$  and  $n = 500, 1000, 2000, 3000$ .

The hypothesis of zero slope was on the average rejected in about 1 case out of 30 with p-values being somewhat smaller than 0.05. Since so many hypotheses were tested simultaneously this is not enough to reject proportionality to  $n^{-1}$  (Wright, 1992). However, when looking closer at the slope estimates it was found that most slopes were negative. This was most striking for the Q estimators. For  $nV(\hat{k}_Q)$  plotted against  $n$  there were 22 slopes out of 30 that were negative, while for  $nV(\hat{\lambda}_Q)$  plotted against  $n$  there were 25 negative slopes. The probabilities of obtaining these outcomes and more extreme ones assuming a 50/50 chance (two-sided p-values) are 0.016 and 0.0004, respectively. Since  $nV(T) = \beta_0 - \beta_1 \cdot n$  implies that  $V(T) = \beta_0 \cdot n^{-1} - \beta_1$  this indicates that a negative constant should be added to the variance of the Q estimators. However, the magnitude of  $\beta_1$  was too small to have any practical impact (largest value being 0.00017). For the two ML estimators 18 slopes out of 30 were negative giving the two-sided p-values 0.36, so in this case there seems to be less reason to reject the proportionality of the variance to  $n^{-1}$ .

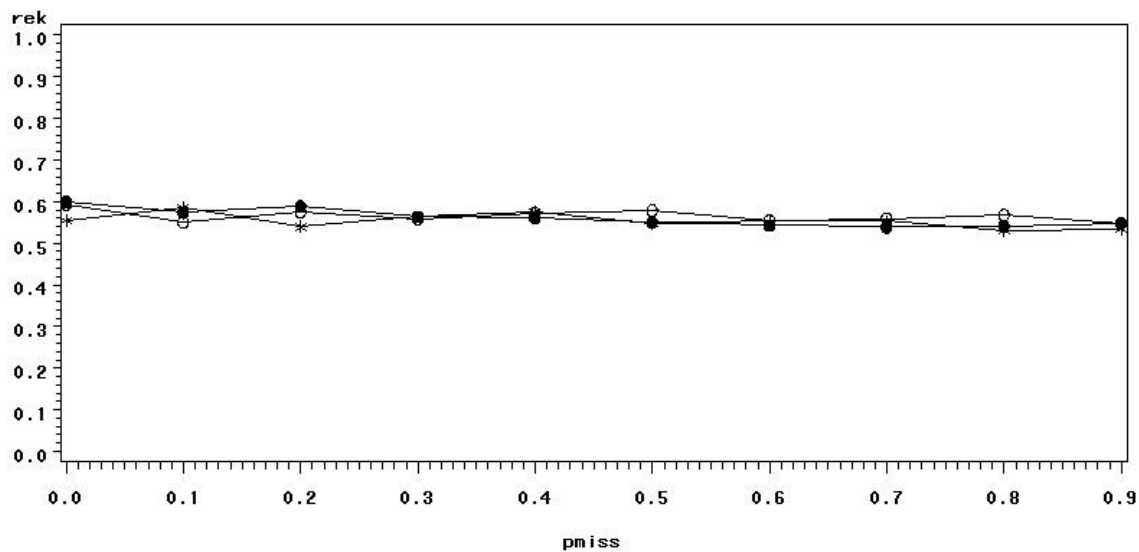
### 3.3 Relative efficiency for estimators of $\theta$

Let  $RE_{\hat{\theta}} = V(\hat{\theta}_{ML})/V(\hat{\theta}_Q)$  be the relative efficiency for the two estimators of  $\theta$  based on the same sample sizes. This is presented in the table in Appendix C when  $q_{.50} = 20$ . Here some interesting features can be noticed. First, for given  $\theta$  and  $p_{miss}$  there is no trend in  $RE_{\hat{\theta}}$  with increasing  $n$ . In fact, when  $RE_{\hat{\theta}}$  was considered as a function of  $n$  the 30 estimated linear slopes were obtained within the range  $\pm 0.00003$  and none of the slopes differed significantly from zero. Therefore the means in the last column may be used to represent the  $RE$ s. Furthermore, for given  $\theta$  and  $n$   $RE_{\hat{\theta}}$  remains constant with increasing  $p_{miss}$  and this also holds for given  $p_{miss}$  and  $n$  with increasing  $\theta$ . The results are summarized in Figure 1.

Figure 2 illustrates the same relations as Figure 2, but now when  $q_{.50} = 200$  (i.e.  $\lambda = 416$ ). The patterns in the two figures are hard to separate. Attempts to study the relative efficiency for larger values of  $q_{.50}$  were made but in this case the iterations in *NLMIXED* often failed to converge and in fact, for  $q_{.50} \geq 2000$  a ML solution could never be obtained. The conclusion from this section is that  $RE_{\hat{\theta}}$  is somewhere between 0.5 and 0.6 irrespective of the values of  $p_{miss}$  and  $n$ , provided that  $\theta$  and  $\lambda$  are in the ranges 0.5-1 and 20-200, respectively.



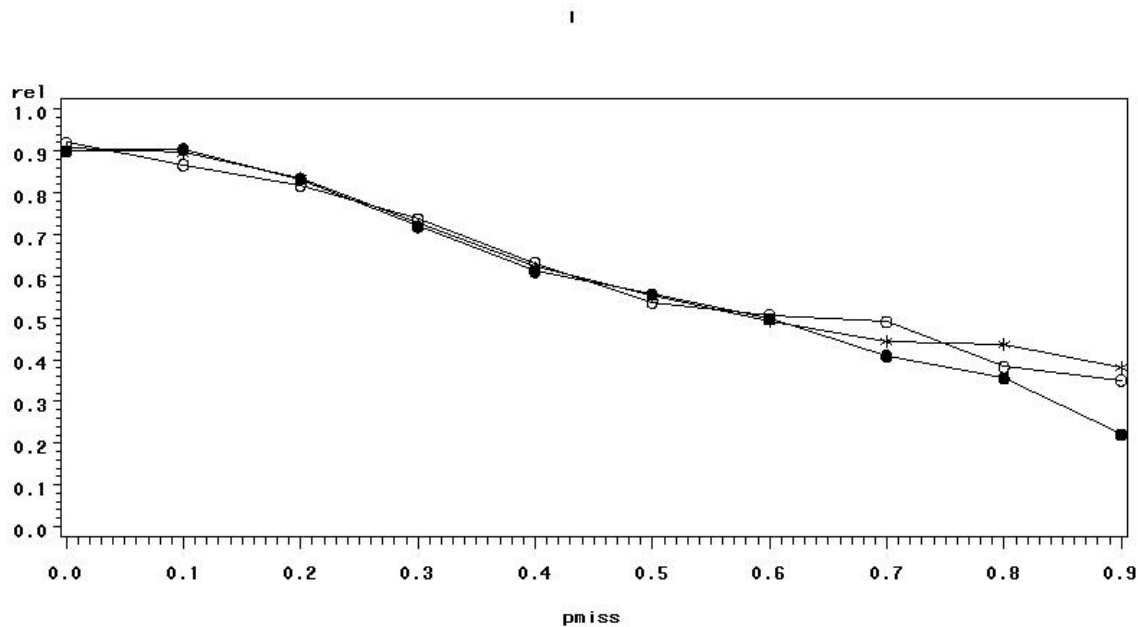
**Figure 1** Relative efficiency,  $V(\hat{\theta}_{ML})/V(\hat{\theta}_Q)$ , when  $q_{.50} = 20$  plotted against the censoring proportion (pmiss) for  $\theta = 0.5$  (filled circles),  $\theta = 2/3$  (circles) and  $\theta = 1$  (stars). Each point represents the mean over  $n$  in the table in Appendix C.



**Figure 2** The same relations as in Figure 1, but now with  $q_{.50} = 200$

### 3.4 Relative efficiency for estimators of $\lambda$

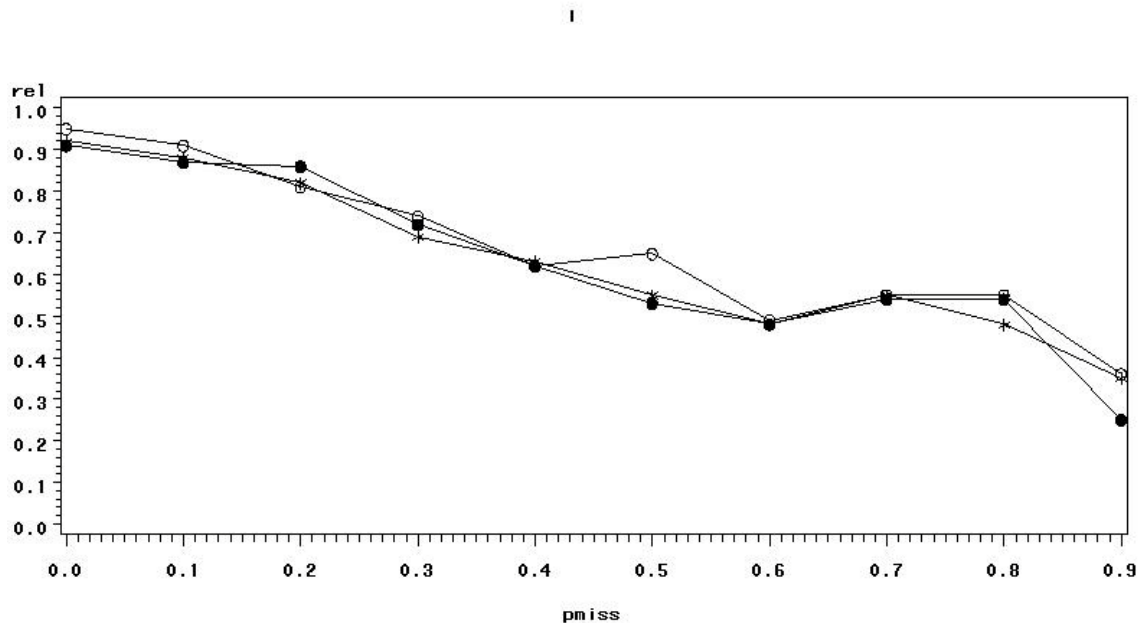
Now, consider  $RE_{\hat{\lambda}} = V(\hat{\lambda}_{ML})/V(\hat{\lambda}_Q)$ . Values of the latter are presented in the table in Appendix D when  $q_{.50} = 20$ . Also here the relative efficiency remains constant over  $n$ , so the means in the final column can be used to represent the relative efficiency. Furthermore, for given  $p_{miss}$  the relative efficiencies do not vary much with  $\theta$ . However, in contrast to the pattern in Figures 1 and 2 the efficiency now decreases with increasing censoring proportion. These findings are summarized in Figure 3.



**Figure 3** Relative efficiency  $V(\hat{\lambda}_{ML})/V(\hat{\lambda}_Q)$  when  $q_{.50} = 20$  plotted against the censoring proportion for  $\theta = 0.5$  (filled circles),  $\theta = 2/3$  (circles) and  $\theta = 1$  (stars). Each point represent the mean over  $n$  in the table in Appendix D.

Since  $\lambda$  is a location parameter the variance of single measurements will increase with increasing  $\lambda$ . This will in turn lead to that the precision of the Q and ML methods are lowered. But which method is most capable to resist against higher variability in data? Figure 4 illustrates the relation between relative efficiency and censoring proportion when  $q_{.50} = 200$  ( $\lambda = 416$ ). A comparison with Figure 3 suggests that no method is better in this respect, possibly with

exception for large values of the censoring proportion where the Q method seems to loose relatively less in precision. From figures 3 and 4 it is concluded that  $RE_{\hat{\lambda}}$  is above 0.5 for  $p_{miss}$  smaller than 0.6.



**Figure 4** The same relations as in Figure 3, but now with  $q_{.50} = 200$ .

#### 4. Discussion and concluding remarks

A simple procedure based on quantiles (Q method) has been suggested for estimating the two parameters in censored Weibull distributions. From simulations using a large number of alternative quantiles it was found empirically that the quantiles corresponding to the set  $p = 0.10, 0.25, 0.50, 0.75, 0.90$  gave estimates with smallest variance. There may be other sets of quantiles that give estimates with smaller variance, but to find the latter a theoretical study is required that is beyond the scope of this paper, which is entirely pragmatic.

The Q method was found to have less precision than the ML method, with relative efficiencies for both parameters being at least 50 % when less than 60 % of the observations were censored.

However, the Q method has the advantage that it is not based on iterative solutions and always works, in contrast to the ML method which may fail, especially for large values of the scale parameter. It should be pointed out that this drawback has been observed when the SAS-procedure *NLMIXED* is used and it might be the case that other procedures works better in this respect.

The Q method requires much less computer time than *NLMIXED*. E.g. with 1000 replicates, each with  $n = 3000$ , the ML solutions were found after about 21 minutes. To this one should add time for working through the print-outs and checking for anomalies. For the Q method the distribution of estimates together with basic statistics was printed out after 3 minutes, using the same number of replicates and  $n$ . Since the relative efficiency of the Q method is at least 50 %, samples with  $n = 6000$  should give at least the same precision for the Q method as for the ML method. The computer time for the Q method was in the latter case 6 minutes.

## References

Cohen, A.C. (1965) Maximum Likelihood Estimation in the Weibull Distribution Based on Complete and on Censored Samples, *Technometrics* **7**, 579-588.

Jonsson, R. (2010) A CUSUM procedure for detecting outbreaks in Poisson distributed medical Events, *Research Report 2010:4*, Statistical Research Unit, Department of Economics, University of Gothenburg, Sweden.

Kruskal, W.H. and Tannur, J.M. (1978) International Encyclopedia of Statistics, vol. 2, McMillan, New York.

Mao, DT. and Li, W. (2007) A Bounded Derivation Method for the Maximum Likelihood Estimation on the Parameters of Weibull Distribution, *Journal of Latex Class Files* **6**, 1-5.

Wright, S.P. (1992) Adjusted P-Values for Simultaneous Inference, *Biometrics* **48**, 1005-1013.

Zacks, S. (1971) The Theory of Statistical Inference, Wiley, New York.



## APPENDIX

### A. SAS code for quantile estimation in the censored two-parameter Weibull distribution.

The parameter  $\theta$  in Eq. (1) is denoted  $c$  ( $=0.5$ ) and  $\lambda$  in Eq. (1) is denoted  $\text{lambda}$  ( $=41.6$ )

```
options ps=12000;
data a;
/* chose the censoring point tup=0.46, giving 90 % censored observations */
tup=0.46;
/* chose the parameters c=0.5 and a lambda giving the median 20 */
c=0.5; m=20; lamb=m/(log(2))**(1/c);
/* generate 5 simulations, each with 2000 Weibull distributed random numbers */
do rep=1 to 5;
do i=1 to 2000;
u=uniform(0);
x=lamb*(-log(1-u))**(1/c); if x>=tup then x='.';
output; end; end;
/* calculate yp (=25 %, 50 % and 75 % percentiles), and pmiss (=proportion missing
observations) */
proc sort; by rep;
proc univariate noprint; var x;
output out=sas1 n=n nmiss=nmiss q1=m25 median=m50 q3=m75;
by rep;
data b; set sas1;
pmiss=nmiss/(n+nmiss);
p=.25; yp=m25; output; p=.50; yp=m50; output; p=.75; yp=m75; output;
/* print the result */
data c; set b;
proc print; var rep yp pmiss p;
/* calculate the least squares estimates of the parameters in equation (3) */
data d; set b;
if yp>0 then lyp=log(yp);
if yp=0 then lyp='.';
z=log(-log(1-(1-pmiss)*p));
z2=z*z; zlyp=z*lyp; lyp2=lyp*lyp;
proc sort; by rep;
proc means noprint n sum; var z lyp z2 zlyp lyp2;
output out=sas1 sum=sz slyp sz2 szlyp slyp2 n=n; by rep;
data b; set sas1;
beta=(n*szlyp-sz*slyp)/(n*sz2-sz*sz);
alfa=slyp/n-beta*sz/n;
c=1/beta; lamb=exp(alfa);
/* average over the 5 samples */
proc means n mean std; var c lamb;
run;
```

The program gives the following print-outs. (When thousands of replicates are made the proc print procedure should be omitted.)

Obs	rep	yp	pmiss	p
1	1	0.02023	0.9005	0.25
2	1	0.09622	0.9005	0.50
3	1	0.25729	0.9005	0.75
4	2	0.02369	0.8940	0.25
5	2	0.12134	0.8940	0.50
6	2	0.24112	0.8940	0.75
7	3	0.01731	0.8935	0.25
8	3	0.09657	0.8935	0.50
9	3	0.22707	0.8935	0.75
10	4	0.03729	0.8855	0.25
11	4	0.11547	0.8855	0.50
12	4	0.27298	0.8855	0.75
13	5	0.02056	0.8905	0.25
14	5	0.11938	0.8905	0.50
15	5	0.29741	0.8905	0.75

The MEANS Procedure

Variable	N	Mean	Std Dev
c	5	0.4699702	0.0614640
lamb	5	64.9270537	34.1831539

## B. SAS code for ML estimation in the censored two-parameter Weibull distribution

In the example below just 1 sample is generated. If e.g. 1000 samples are required, then the statement 'do rep=1 to 1' is changed to 'do rep=1 to 1000'.

```
data test;
/* chose the censoring point tup=0.46, giving 90 % censored observations */
tup=0.46;
/* chose the parameters c=0.5 and a lambda giving the median 20 */
c=0.5; m=20; lamb=m/(log(2))**(1/c);
/* generate 1 sample with 2000 Weibull distributed random numbers */
do rep=1 to 1;
do i=1 to 2000;
u=uniform(0);
y=lamb*((-log(u))**(1/c));
if y>tup then do;
y=tup; cens=1;
end; else cens=0; output; end; end;
keep rep y cens i;
run;
data b; set test;
proc sort; by rep;
/* invoke the procedure NLMIXED to estimate the parameters */
proc nlmixed;
/* ensure that shape and scale parameters are positive */
c=exp(log_c); lamb=exp(log_lamb);
if cens=0 then ll=log(c)-log(lamb)+(c-1)*(log(y)-log(lamb))-(y/lamb)**c;
else ll=-(y/lamb)**c;
model y~general(ll); estimate "c" exp(log_c); estimate "lamb" exp(log_lamb);
by rep;
run;
```

The program gives the following print-out (Information criteria such as AIC and BIC and further details have been omitted since focus is on the estimates.)

The NLMIXED Procedure

----- rep=1 -----

The NLMIXED Procedure

Dimensions

Observations Used	2000
Observations Not Used	0
Total Observations	2000
Parameters	2

Parameters

log_c	log_lamb	NegLogLike
1	1	1259.92777

Iteration History

Iter	Calls	NegLogLike	Diff	MaxGrad	Slope
1	2	892.969302	366.9585	533.2836	-31564.1
2	3	548.922993	344.0463	491.9649	-634.126
3	4	405.232366	143.6906	201.7811	-359.52
4	5	389.933964	15.2984	37.35354	-38.23
5	7	386.317527	3.616437	45.48596	-2.90184
6	8	380.712102	5.605425	6.557471	-2.92069
7	10	379.868079	0.844024	16.67495	-1.19448
8	12	379.753913	0.114166	0.467872	-0.18927
9	14	379.75134	0.002572	0.117899	-0.00396
10	16	379.751329	0.000011	0.00024	-0.00002
11	18	379.751329	5.3E-10	7.298E-7	-1.04E-9

NOTE: GCONV convergence criterion satisfied.

Parameter Estimates

Parameter	Estimate	Standard Error	DF	t Value	Pr >  t	Alpha	Lower	Upper	log_c
-0.8080	0.07414	2000	-10.90	<.0001	0.05	-0.9534	-0.6626	log_lamb	4.5489
0.4326	2000	10.51	<.0001	0.05	3.7005	5.3974			

Additional Estimates

Label	Estimate	Standard Error	DF	t Value	Pr >  t	Alpha	Lower	Upper
c	0.4457	0.03305	2000	13.49	<.0001	0.05	0.3809	
lamb	94.5297	40.8963	2000	2.31	0.0209	0.05	14.3259	

### C. Examples of anomalies with the SAS-procedure *NLMIXED*

a) An example where a correct ML solution is obtained although it is stated that the solution can not be reached

Iteration History						
Iter	Calls	NegLogLike	Diff	MaxGrad	Slope	
1	12	17851.6582	2308928	52967.74	-3.89E12	
2	15	11041.0457	6810.613	10837.43	-40440.1	
3	16	3366.86774	7674.178	2864.557	-23151.3	
4	18	3019.56554	347.3022	1851.341	-1571.62	
5	19	3017.62873	1.936814	1831.711	-3.17491	
6	21	3017.27656	0.352168	1822.819	-0.36901	
7	25	2699.54193	317.7346	1383.342	-0.34861	
8	27	1991.31525	708.2267	357.0605	-439.544	
9	28	1910.08271	81.23254	177.3794	-23.1151	
10	29	1774.83981	135.2429	109.4697	-46.1012	
11	31	1738.88315	35.95667	208.6956	-36.2012	
12	32	1725.77375	13.1094	115.5781	-103.136	
13	33	1716.66791	9.105842	62.42045	-19.4089	
14	35	1714.09113	2.576778	4.716303	-5.77166	
15	37	1714.06761	0.023514	0.561202	-0.04209	
16	39	1714.06543	0.002184	0.135825	-0.0017	
17	41	1714.06542	0.000013	0.000096	-0.00003	
17	70	1714.06542	0	0.000096	-1.05E-6	

ERROR: Optimization cannot be completed.

Parameter Estimates		
Parameter	Estimate	Gradient
log_c	-0.7700	-0.0001
log_lamb	6.0715	-0.00004

The 'label parameters'  $c$  and  $lamb$  (i.e.  $\theta$  and  $\lambda$ ) are suppressed and can not be read from the printout. Notice however that  $c = \exp(-0.7700) = 0.46$  and  $lamb = \exp(6.0715) = 433$  are not far from the true values 0.50 and 416, respectively.

b) An example where an incorrect ML solution is obtained although it is stated that it has been reached

Iteration History

Iter	Calls	NegLogLike	Diff	MaxGrad	Slope
1	18	114573115	6.985E10	1.6925E9	-2.41E22
2	22	86092.5492	1.1449E8	84466.6	-1.17E12
3	26	13552.2532	72540.3	11443.88	-447.257
4	30	10839.0783	2713.175	12517.35	-9462.84
5	32	10825.3823	13.69595	11678.37	-89.0995
6	34	10825.2982	0.084119	11731.81	-0.15578
7	36	10825.2981	0.000115	11733.86	-0.00019
8	42	10822.4294	2.868657	12067.43	-0.00004

NOTE: GCONV convergence criterion satisfied.

Parameter	Estimate
log_c	1.5578
log_lamb	8.3752

Label	Estimate
c	4.7482
lamb	4338.20

*c* and *lamb* being very far from the true values 0.50 and 416, respectively.

### D. Relative efficiency of estimators for $\theta$

$k$	$P_{miss}$ (%)	$n$				$Mean$
		500	1000	2000	3000	
0.5	0	.58	.65	.56	.56	.588
"	10	.55	.60	.56	.64	.587
"	20	.54	.54	.57	.58	.556
"	30	.50	.56	.58	.60	.563
"	40	.57	.56	.53	.60	.563
"	50	.54	.57	.55	.59	.563
"	60	.59	.56	.54	.58	.568
"	70	.50	.50	.54	.58	.532
"	80	.62	.59	.56	.57	.586
"	90	.54	.57	.53	.54	.545
2/3	0	.58	.57	.56	.59	.574
"	10	.57	.59	.58	.58	.581
"	20	.57	.52	.54	.57	.550
"	30	.55	.60	.58	.58	.576
"	40	.62	.57	.50	.56	.563
"	50	.56	.57	.55	.53	.551
"	60	.56	.56	.54	.56	.555
"	70	.55	.54	.55	.58	.554
"	80	.56	.56	.56	.52	.552
"	90	.53	.53	.52	.57	.535
1	0	.55	.56	.53	.57	.551
"	10	.58	.53	.58	.62	.579
"	20	.57	.57	.62	.57	.583
"	30	.60	.56	.57	.60	.582
"	40	.57	.58	.53	.58	.567
"	50	.53	.56	.58	.55	.553
"	60	.58	.59	.55	.54	.563
"	70	.53	.56	.55	.56	.548
"	80	.53	.54	.53	.56	.539
"	90	.50	.53	.57	.54	.533

### E. Relative efficiency for estimators of $\lambda$

$k$	$P_{miss}$ (%)	$n$				$Mean$
		500	1000	2000	3000	
0.5	0	.87	.95	.90	.90	.900
"	10	.86	.93	.93	.90	.904
"	20	.88	.78	.84	.84	.833
"	30	.72	.75	.72	.68	.719
"	40	.61	.64	.58	.62	.612
"	50	.58	.59	.53	.53	.557
"	60	.46	.50	.47	.56	.498
"	70	.40	.41	.41	.41	.409
"	80	.17	.39	.47	.39	.356
"	90	.02	.11	.29	.46	.221
2/3	0	.94	.92	.90	.92	.920
"	10	.86	.84	.86	.91	.867
"	20	.82	.78	.83	.84	.818
"	30	.73	.70	.76	.76	.737
"	40	.62	.60	.65	.66	.632
"	50	.53	.59	.50	.54	.538
"	60	.44	.54	.53	.53	.507
"	70	.53	.46	.51	.47	.491
"	80	.23	.42	.44	.44	.383
"	90	.13	.52	.37	.38	.351
1	0	.93	.94	.89	.87	.907
"	10	.88	.91	.89	.90	.896
"	20	.83	.80	.83	.88	.835
"	30	.73	.73	.75	.69	.726
"	40	.60	.67	.59	.63	.623
"	50	.52	.54	.56	.59	.553
"	60	.45	.49	.52	.52	.493
"	70	.37	.44	.48	.49	.444
"	80	.38	.46	.50	.41	.437
"	90	.25	.51	.40	.37	.381



## Research Report

- 2007:12 Frisé, M., Andersson, E. & Schiöler, L. Robust outbreak surveillance of epidemics in Sweden.
- 2007:13 Frisé, M., Andersson, E. & Pettersson, K. Semiparametric estimation of outbreak regression.
- 2007:14 Pettersson, K. Unimodal regression in the two-parameter exponential family with constant or known dispersion parameter.
- 2007:15 Pettersson, K. On curve estimation under order restrictions.
- 2008:1 Frisé, M. Introduction to financial surveillance.
- 2008:2 Jonsson, R. When does Heckman's two-step procedure for censored data work and when does it not?
- 2008:3 Andersson, E. Hotelling's T2 Method in Multivariate On-Line Surveillance. On the Delay of an Alarm.
- 2008:4 Schiöler, L. & Frisé, M. On statistical surveillance of the performance of fund managers.
- 2008:5 Schiöler, L. Explorative analysis of spatial patterns of influenza incidences in Sweden 1999 – 2008.
- 2008:6 Schiöler, L. Aspects of Surveillance of Outbreaks.
- 2008:7 Andersson, E & Frisé, M. Statistiska varningssystem för hälsorisker
- 2009:1 Frisé, M., Andersson, E. & Schiöler, L. Evaluation of Multivariate Surveillance
- 2009:2 Frisé, M., Andersson, E. & Schiöler, L. Sufficient Reduction in Multivariate Surveillance
- 2010:1 Schiöler, L. Modelling the spatial patterns of influenza incidence in Sweden
- 2010:2 Schiöler, L. & Frisé, M. Multivariate outbreak detection