A method of moments for the estimation of Weibull pdf parameters

VINCENZO NIOLA
Department of Mechanical Engineering for Energetics
University of Naples “Federico II”
Via Claudio 21, 80125, Napoli
ITALY

ROSARIO OLIVIERO
University of Naples “Federico II”
Via Pansini 5, 80131, Napoli
ITALY

GIUSEPPE QUAREMBA
University of Naples “Federico II”
Via Pansini 5, 80131, Napoli
ITALY

Abstract: - In order to estimate the parameters of a Weibull distribution, we study the performance of the method of moments. For that reason, we compare three methods for the estimation of the cumulative distribution. The first method is a classical commonly-used approximation, the second one is the “Monte Carlo corrected median rank”. Finally, the third one utilizes a “wavelet estimator” of the empirical distribution. In particular, by Monte Carlo procedure we have verified that in order to estimate the scale parameter, the wavelet estimator is more consistent with reference to all other ones. Finally, we argue that, the wavelet estimator is, in general, more robust than other ones.

Such a method could be utilized in mechanical applications (e.g., transmissions, rolling organs, kinematics couples, etc) or time series analysis.

Key-Words: - Wavelet analysis, Weibull distribution, parameter estimation, mechanical lifetime.

1 Introduction
The Weibull distribution is one of the frequently used in order to estimate the times-to-failure in constant stress tests on mechanical or electronics equipments. In parameter estimation, it is necessary to know the values that the cumulative distribution function (c.d.f.) assumes starting from the empirical data set.

In this paper we describe the calculation of three estimators of the empirical c.d.f.; therefore we utilize them in order to estimate the parameters by means of the method of moments [1]. In particular, we use the classical approximation, the so called Fothergill Monte Carlo median rank estimator [2][3], and a nonparametric wavelet estimator [4][5].

We show that, in order to estimate the Weibull parameters, the wavelet estimator is almost equally consistent as the Fothergill one. This is a logical consequence, in fact the Fothergill estimator has been derived by modifying the classical mean rank by a Monte Carlo procedure. On the other side, we argue that, respect to bilateral data censoring, the wavelet estimator is more robust with respect to the other ones.

2 The method of moments
Let \( X_1, X_2, \ldots, X_n \) be \( n \) observed random values, whose density is given by the two dimensional Weibull p.d.f., whose c.d.f. is

\[
F(t; \eta, \beta) = 1 - \exp\left(-\eta t^\beta\right),
\]

(1)
(with \( \beta > 0 \) and \( \eta > 0 \)), where \( \beta \) and \( \eta \) are the shape and scale parameter respectively. The moments method estimator, for the above parameters, has the form [1]
An alternative method, consists of estimating the density where

\[ \hat{f}(X_i) = \frac{i - 0.3}{n + 0.4} \]  

(2)

The estimator (2) is derived by using Monte Carlo simulations, in order to estimate the median values for the c.d.f. of the \( i^{th} \) of \( n \) i.i.d. samples. Note that, generally, the estimator (2) replaces the classical mean rank approximation [2][3]

\[ \hat{F}(X_i) = \frac{i}{n + 1}. \]  

(3)

Another approximation is given by the formula [2][3][4]

\[ \hat{F}(X_i) = \frac{i - 0.5}{n}. \]  

(4)

An alternative method, consists of estimating the density \( f \), by means of nonparametric estimator \( \hat{f} \): we can obtain \( \hat{F}(X) \) by integrating \( \hat{f} \) between 0 and \( X \). In our work, we propose a not parametric estimator. Preliminarily, for any \( i \in \{1, 2, ..., n\} \), set

\[ x_i = X_i / \sqrt{\sum_{h=1}^{n} X_h^2}. \]  

(5)

Observe that, by (3) we have simply made a change of scale in order to normalize \( X_1, X_2, ..., X_n \) between 0 and 1. We propose the density estimator given by [5][6]

\[ \hat{f}_{\hat{\beta}}(X_i) = \frac{1}{n} \sum_{i=1}^{n} \sum_{k=1}^{n} \varphi_{\hat{\beta}+1,k}(x_i)\varphi_{\hat{\beta}+1,k}(x_h), \]  

(6)

where

\[ \varphi_{\hat{\beta}}(x) = 2^{1/2} \varphi(2^{1/2} x - k), \quad k \in Z, \]

\[ \varphi(x) = \begin{cases} 1, & x \in (0,1] \\ 0, & x \notin (0,1] \end{cases} \]

and

\[ j_1 = \text{Int} \left[ \log_2 \left( \frac{1}{\text{median}(X_i - x_{i-1})} \right) + \right. \]

\[ - \ln \left( \log_2 \left( \frac{1}{\text{median}(X_i - x_{i-1})} \right) \right) \]  

See [5]. In general, given a function \( \psi(x) \), whose first \( h \) \( (h \in N) \) moments are zero (mother wavelet), \( \varphi(x) \) is chosen in order to be (father wavelet) orthonormal to \( \psi(x) \), according to the \( L^2 \) norm. In our work we have chosen

\[ \psi(x) = \begin{cases} -1, & x \in \left[0, \frac{1}{2}\right] \\ 1, & x \in \left\{ \frac{1}{2}, 1 \right\} \end{cases} \]

The choice of a similar interpolating curves is done because the order \( N \), of the empirical data set, can be relatively small. It is easy to prove that, if \( \psi \) is a mother (father) wavelet, then also \( \varphi_{\hat{\beta}} \) is a mother (father) wavelet.

### 3 Monte Carlo calculation of consistency estimator

First of all observe that, for any random variable, the c.d.f. follows a uniform distribution between 0 and 1. Therefore, in order to investigate the consistency of the method of moments we adopt the following procedure: we perform \( N \) simulated samples; each of them is composed of \( n \) random numbers \( F_i^{(k)} \) \( (i \in \{1, 2, ..., n\}, \ k \in \{1, 2, ..., N\}) \) showing a uniform distribution between 0 and 1. They represent \( n \) possible values for the c.d.f. of Weibull distributions with given parameters \( \beta \) and \( \eta \). Therefore, for any \( k \in N, \) set

\[ X_i^{(k)} = \left[ \frac{\ln(1 - F_i^{(k)})}{\eta} \right]^2 \quad (i \in \{1, 2, ..., n\}). \]
each of these $N$ samples represents $n$ possible values for a Weibull random variable with parameters $\beta$ and $\eta$. Finally, we calculate $\hat{\beta}$ and $\hat{\eta}$, by utilizing the method of moments. 

Now, fix $k^* \in \{1, 2, \ldots, N\};$ let

$$X_i^{(k^*)} = \left[ \frac{-\ln(1 - F_i^{(k^*)})}{\eta} \right]^{\frac{1}{\beta}} \quad (\eta \in \mathbb{R})$$

and

$$Y_i^{(k^*)} = \left[ \frac{-\ln(1 - F_i^{(k^*)})}{\eta'} \right]^{\frac{1}{\beta}} = \frac{\eta}{\eta'} X_i^{(k^*)},$$

with $i \in \{1, 2, \ldots, n\}$ and $\eta' \in \mathbb{R}$.

This last sample represents $n$ possible values for a Weibull random variable with parameters $\beta$ and $\eta'$. In this case, from (2), (3) and (4), given that $F(X_i)$ depends only on $i$ and not on $\beta$ and $\eta$, we deduce that, for any $i \in \{1, 2, \ldots, n\}$,

$$\hat{F}(X_i^{(k^*)}) = \hat{F}(Y_i^{(k^*)})$$

for any classical estimator of $F$. On the other side, note that

$$\ln(X_n^{(k^*)}) - \ln(X_1^{(k^*)}) = \ln(Y_n^{(k^*)}) - \ln(Y_1^{(k^*)}).$$

This equality follows from the statement that, for any $a, x, y \in \mathbb{R}$,

$$\ln(ax) - \ln(ay) = \ln(x) - \ln(y).$$

Finally, let us apply the estimator of moments to the two samples $X_1^{(k^*)}, X_2^{(k^*)}, \ldots, X_n^{(k^*)}$ and $Y_1^{(k^*)}, Y_2^{(k^*)}, \ldots, Y_n^{(k^*)}$. It follows that $\hat{\beta}$ is independent by the value of the scale parameter $\eta$ assigned \textit{a priori}, whatever classical method we use in order to estimate $F$. We can therefore suppose that $\eta = 1$. Note that, on the other side, $\hat{\eta}$ may depend on the value of $\beta$ assigned \textit{a priori}. Table 1 shows that, in general, for the estimator

$$\hat{f}_j(X_i) = \frac{1}{n} \sum_{i=1}^{n} \sum_{k} \varphi_{j+1,k}(x_i) \varphi_{j,k}(x_i), \quad (5)$$

$(j \in \mathbb{N})$, an optimal choice is $j = j_1$, where

$$j_1 = \text{int} \left[ \frac{\log_2 \frac{1}{\text{mean} \left( x_i - x_{i-1} \right)} + }{-\ln \left( \frac{1}{\text{mean} \left( x_i - x_{i-1} \right)} \right)} \right].$$

By Tables 2-3, note that the Fothergill estimator, of the c.d.f., appears almost as stable as the wavelet one. This is natural, given that the first estimator is indeed derived also by a Monte Carlo procedure.

### 4 Robustness respect to bilateral data censoring

Suppose $X_1, X_2, \ldots, X_n$ ordered in ascending way. In this case, by (1) the moment estimator for the shape parameter may be written in a simplified form as

$$\hat{\beta} = \frac{\ln \left( \frac{1 - \hat{F}(X_n)}{1 - \hat{F}(X_1)} \right)}{\ln(X_n) - \ln(X_1)}.$$

From this, it follows that, if $F$ is given by mean or median rank approximations, then the moment shape parameter estimator does not depend on the relative position of the intermediate values.

In order to testing the robustness of our proposed estimator, we adopt the following procedure. For any $k \in \{1, 2, \ldots, N\}$, we simulate the sample

<table>
<thead>
<tr>
<th>$j$</th>
<th>$\hat{\beta}$</th>
<th>$\hat{\eta}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$j_1 - 2$</td>
<td>1.3120</td>
<td>0.8735</td>
</tr>
<tr>
<td>$j_1 - 1$</td>
<td>1.5610</td>
<td>0.9671</td>
</tr>
<tr>
<td>$j_1$</td>
<td>1.9869</td>
<td>1.0547</td>
</tr>
<tr>
<td>$j_1 + 1$</td>
<td>2.2223</td>
<td>1.0911</td>
</tr>
<tr>
<td>$j_1 + 2$</td>
<td>2.2994</td>
<td>1.1446</td>
</tr>
</tbody>
</table>

Table 1. Median values of the parameters estimates for various values of $j$, for different Monte Carlo simulations ($N = 10000$ for any value of $j$). $\beta = 2$, $n = 6$ and $\eta = 1$. 


Table 2. Median values of the shape parameter estimates, for different values of $\beta$, for different Monte Carlo simulations ($N = 10000$ for any value of $\beta$), $n = 6$ and $\eta = 1$.

In the Fig.1 are depicted the comparison of results obtained by the showed methods.

Table 3. Median values of the scale parameter estimates, for different values of $\beta$, for different Monte Carlo simulations ($N = 10000$), $n = 6$ and $\eta = 1$.

In the Fig.2 below, are compared the results of methods in the case of bilateral data censoring.

Table 4. Median values of the shape parameter estimates, for different values of $\beta$, for different Monte Carlo simulations, with bilateral data censoring, ($N = 10000$ for any value of $\beta$), $n = 10$ and $\eta = 1$.

Table 5. Median values of the scale parameter estimates, for different values of $\beta$, for different Monte Carlo simulations, with bilateral data censoring, ($N = 10000$), $n = 10$ and $\eta = 1$. 
of the Weibull random variable, as before defined. Therefore, after having considered the sample to be ordered in ascending order, we consider the sample

\[ (x_2^{(i)}, x_3^{(i)}, \ldots, x_{i-1}^{(i)}, x_{i+1}^{(i)}, \ldots, x_{n-1}^{(i)}, x_n^{(i)}) \]

\((i \in \{1, 2, \ldots, n\})\). It is obtained by censoring its extreme values. Finally, for any so modified sample, we estimate the parameters by the two given methods. Note that, as shown by Tables 4-5, our proposed method appears to be more robust.

5 Conclusions
We have proposed a method of moments to calculate the parameters of a Weibull distribution. The proposed method belongs to the family of nonparametric methodologies. It is based on the nonparametric determination of the values of the cumulative distribution, given the assigned sample. We have compared the estimates, for the shape and scale parameters, with reference for the classical method of moments. The parameter estimates, performed by the proposed methodology, appear to be robust enough, if compared to the classical method of moments. For this reason, the proposed methodology can be indicated for analysing the real working conditions of mechanical systems.

References: