Estimating the Three-Parameter Weibull Distribution by the Method of Probability-Weighted Moments with Application to Medical Survival Data

Alfred A. Bartolucci[1], Karan P. Singh[1], Anne D. Bartolucci[2], Mohammed A. El-Saidi[3], and J. Michael Hardin[1]

[1]Department of Biostatistics
University of Alabama at Birmingham
Birmingham, Alabama 35294, USA

[2]Department of Mathematics
Huntingdon College
Montgomery, Alabama 36106, USA

[3]Department of Economics and Applied Statistics
Ferris State University
Big Rapids, Michigan 49307, USA

Summary The method of probability-weighted moments is used to derive estimators of parameters and quantiles of the three-parameter Weibull distribution. The properties of these estimators are studied. The results obtained are compared with those obtained by using the method of maximum likelihood. The Weibull probability distribution has numerous applications in various areas: for example, breaking strength, life expectancy , survival analysis and animal bioassay. Because of its useful applications, its parameters need to be evaluated precisely, accurately and efficiently. There is a rich literature available on its maximum likelihood estimation method. However, there is no explicit solution for the estimates of the parameters or the best linear unbiased estimates. Further, the Weibull parameters cannot be expressed explicitly as a function of the conventional moments and iterative computational methods are needed. The maximum likelihood methodology is based on large-sample theory and the method might not work well when samples are small or moderate in size. Others have proposed a class of moments, called probability-weighted moments. This class seems to be interesting as a method for estimating parameters and quantiles of distributions which can be written in inverse form. Such distributions include the Gumbel, Weibull, logistic, Tukey's symmetric lambda, Thomas Wakeby, and Mielke's kappa. It has been illustrated that rather simple expressions for the parameters can be written in inverse form in terms of probability-weighted moments (PWM) for most of these distributions. In this paper we define the PWM estimators of the parameters for the three-parameter Weibull distribution. We investigate the properties of these estimators in a medical application setting. We also examine the added influence that censored data may have on the estimates.

1. Introduction

The Weibull probability distribution has numerous applications in various areas; for example, breaking strength, life expectancy, survival analysis, and animal bioassay (Weibull 1938, 1961; Henderson 1965). Because of its useful applications, its parameters need to be evaluated precisely, accurately and efficiently. There is a rich literature available on its maximum likelihood estimation. However, there is no explicit solution for the estimates of the parameters or the best linear unbiased estimates, and thus iterative computational methods are required. Further, the Weibull parameters cannot be expressed explicitly as a function of the conventional moments and iterative computational methods must be used.

The maximum likelihood methodology is based on large-sample theory, and hence may perform poorly for small or moderate sample sizes. Greenwood *et al* (1979) proposed a class of moments, called *probability weighted moments* (**PWM**). This class seems to be of interest as a method for estimating parameters and quantiles of distributions which can be written in inverse form. Such distributions include the Gumbel, Weibull, logistic, Tukey's symmetric lambda, Thomas Wakeby, and Mielke's kappa. Greenwood *et al* (1979), Landwehr and Matalas (1979), Hosking *et al* (1985), and Shoukri *et al* (1988) illustrate that rather simple expressions for the parameters can be written in inverse form in terms of probability weighted moments (PWM) for these distributions.

In this paper we derive the PWM estimators of the parameters and quantiles of the three-parameter Weibull distribution. We investigate the distributional properties of these estimators in large samples by using asymptotic theory, and in small and moderate samples by using Monte Carlo simulation.

2. Probability-Weighted Moments

The (r,s,t) probability-weighted moment (PWM) of a random variable X with a cumulative distribution function (CDF), F(x), is defined by

$$M_{r,s,t} = E[\, X^r \,(\, F(\,X\,)^s\,)\,(\,1 - F(X)^t\,)\,] \quad (1)$$

where r, s, and t are real numbers. Let X(F) be the inverse distribution. Then expression (1) can be written as

$$M_{r,s,t} = \int_0^1 (\,X(F)^r\,)\,F^s\,(\,1 - F)^t\,dF \quad (2)$$

On the other hand, if we have a random sample of size (s + t + 1), then $M_{r,s,t}$ is equivalently the product of $(s+t)^{-1}B(s,t)$ and $E[X^r_{(s+1)}]$, where B(s,t) is the complete beta function and $X_{(s+1)}$ is the (s+1)the order statistic. However, the sample size, n, need not be equal to (s+t+1). For further details, the reader is referred to Greenwood *et al* (1979) and Hosking *et al* (1985).

Let X be the Weibull random variable with probability density function (Johnson and Kotz 1970, p. 250) given by

$$f(x) = \frac{\alpha}{\beta}\,\frac{(x-\alpha)^{\delta-1}}{\beta}\,e^{-\frac{(x-\alpha)}{\beta}}\;; \quad (3)$$

$$\alpha < x < \infty\,,\ \beta > 0\,,\ \textit{and}\ \delta > 0\,.$$

The corresponding distribution function is

$$F(x) = 1 - e^{-(\frac{x-\alpha}{\beta})^\delta}\;; \quad (4)$$

$$\alpha < x < \infty\,,\ \beta > 0\,,\ \delta > 0\,.$$

where α, β, and δ are the location, scale, and shape parameters, respectively. For properties of this distribution, Johnson and Kotz (1970, Chapter 20) is an excellent reference. The inverse distribution function of F(x) is given by

$$X(F) = \alpha + \beta[\,-\ln(1 - F)\,]^{\frac{1}{\delta}} \quad (5)$$

Let $M_r = M_{1,\alpha,,r}$. Then the following expression for M_r can easily be shown:

$$M_r = \frac{\alpha}{(1+r)} + \beta\Gamma(1 + \frac{1}{\delta})/(1+r)^{(1+\frac{1}{\delta})} \quad (6)$$

where $\Gamma(.)$ denotes the gamma function. A similar expression is also given in Greenwood *et al* (1979). As discussed in section 1, the parameters of the Weibull distribution cannot be expressed explicitly as functions of the conventional moments. However, it is interesting to note that Weibull parameters can be expressed as functions of probability weighted moments. Following Greenwood *et al* (1979) we consider two cases in this paper:

<u>Case 1:</u> The location parameter α is known and without loss of generality can be equal to be zero. We, therefore, need to estimate the β and δ only. The expression (6) reduces to

$$M_r = \beta\Gamma(1+\frac{1}{\delta})/(1+r)^{(1+\frac{1}{\delta})} \quad (7)$$

which yields

$$\hat{\beta}_0 = \frac{M_0}{\Gamma\left[\frac{\ln(M_0/M_1)}{\ln(2)}\right]} \quad (8)$$

and

$$\hat{\delta} = \frac{\ln(2)}{\ln(\frac{M_0}{2M_1})} \quad (9)$$

<u>Case 2:</u> The location parameter α is unknown and $\alpha \neq 0$. Then the parameters α, β, and δ are estimated by

$$\hat{\alpha} = \frac{4(M_0M_3-M_1^2)}{(M_0+4M_3-4M_1)} \quad (10)$$

$$\hat{\beta} = \frac{(M_0-\hat{\alpha})}{\Gamma\left[\ln\left(\frac{M_0-2M_1}{M_1-2M_3}\right)/\ln(2)\right]} \quad (11)$$

and

$$\hat{\delta} = \frac{\ln(2)}{\ln\left[\frac{M_0-2M_1}{2M_1-4M_3}\right]} \quad (12)$$

An unbiased estimator of M, based on the order sample $X_{(1)} < X_{(2)} < ... < X_{(n)}$ of a random sample of size n from a universal distribution has been briefly discussed by Landwehr et al (1979), Hosking et al (1985) and others. For the Weibull distribution it is defined by

$$\hat{M}_r = n^{-1} \sum_{j=1}^{n} \frac{(j-1)(j-2)...(j-r)}{(n-1)(n-2)...(n-r)} X_{(j)} \tag{13}$$

Hoeffding (1948) proved that $\hat{M}_r$ are closely related to U - statistics. The properties of U - statistics are given by Locke and Spurrier (1976), Fraser (1957, Chap. 4), Randles and Wolfe (1979, Chap. 3) and others. The U - statistics have desirable properties including robustness to outliers in the sample, high efficiency and asymptotic normality. In fact, Hosking et al extended some of these properties to the probability weighted moment estimator $\hat{M}_r$ and other quantities calculated from them. However, Hosking et al used

$$\hat{M}_r(P_{j,n}) = n^{-1} \sum_{j=1}^{n} (P^r_{j,n}) X_{(j)} \tag{14}$$

where $P_{j,n} = (j-a)/n$, $0 < a < 1$, or $P_{j,n} = (j-a)/(n+1-2a)$, $1/2 < a < 1/2$, it can be shown that estimators $\hat{M}_r(P_{j,n})$ are asymptotically equivalent to $\hat{M}_r$ and therefore, consistent estimators of M_r. In this paper, we use estimators of M_r proposed by Hosking, et al.

3. Moment Estimates

As discussed in section 1, the Weibull parameters can not be expressed explicitly as a function of the conventional moments. Let $\mu^1_\kappa = E[X^\kappa]$. Then in general the first, second and third central moments are given as functions of the Weibull parameters α, β and δ and can easily be derived:

$$\mu_1 = \alpha + \beta \ \Gamma(1 + \frac{1}{\delta}) \tag{15}$$

$$\mu_2 = \beta^2[\Gamma(1 + \frac{2}{\delta} - \Gamma^2(1 + \frac{1}{\delta}] \tag{16}$$

$$\mu_3 \quad \beta^3[\Gamma(1 + \frac{3}{\delta}) - 3\Gamma(+ \frac{1}{\delta})\Gamma(1 + \frac{2}{\delta}) + 2\Gamma^3(1 + \frac{1}{\delta})] \tag{17}$$

It is interesting to note that β_1 which is the coefficient of skewness, will be a function of δ alone. Let $\beta_1 = \phi(\delta)$. Then

$$\phi(\delta) = \frac{\left[\Gamma\left(1+\frac{3}{\delta}\right) - 3\Gamma\left(1+\frac{1}{\delta}\right)\Gamma\left(1+\frac{2}{\delta}\right) + 2\Gamma^3\left(1+\frac{1}{\delta}\right)\right]^2}{\left[\Gamma\left(1+\frac{2}{\delta}\right) - \Gamma^2\left(1+\frac{1}{\delta}\right)\right]^3} \tag{18}$$

Similar expressions for the moment estimators of the three parameter Weibull distribution are given in Sinha (1986, p. 71). As suggested by Sinha, one may tabulate β_1 for different values of δ. From a given sample we can compute the sample value of β_1 since the first three sample moments are $\mu_1^1 = \bar{x}$, $\mu^1_2 = \Sigma x_i^2/n$ and $\mu^1_3 = \Sigma x_i^3/n$. Once we have the sample value of β_1 we can use the tabulated table of $(\beta, \varphi(\delta))$ to compute δ. One may need interpolation. By using the estimated δ in the following expressions one can estimate α and β.

$$S^2 = \beta^2\left[\Gamma(1 + \frac{2}{\delta}) - \Gamma^2(1 + \frac{1}{\delta})\right] \tag{19}$$

$$\bar{X} = \alpha + \beta\Gamma(1 - \frac{1}{\delta}) \tag{20}$$

where

$$S^2 = \Sigma(x_i - \bar{x})^2/(n - 1)$$

From the moment procedure it seems that the moment estimates are easier to compute than their ML - Counterparts. But efficiency - wise they are of limited use. However, these may be used as good initials for the maximum likelihood method.

4. Maximum Likelihood Estimates

The log - likelihood function for a random sample of size n with d failures and n - d censored observations from the Weibull distribution is given by

$$LnL = d\ Ln\delta - d\delta\ Ln\ \beta + (\delta - 1)\ \sum_d Ln(x_i - \alpha) - \sum_n \frac{(x_i - \alpha)^\delta}{\beta} \tag{21}$$

By taking the partial derivatives with respect to α, β and δ respectively and equating them to zero yields:

$$-(\delta - 1)\beta^{\delta} \sum_{d} \frac{1}{(x_i - \alpha)} + \delta\sum_{n}(x_i - \alpha)^{\delta-1} = 0 \quad (22)$$

$$-d\beta^{\delta} + \sum_{n}(x_i - \alpha)^{\delta} = 0 \quad (23)$$

and

$$n - n\delta Ln\beta + \delta\Sigma Ln(x_i - \alpha) - \delta\Sigma\left(\frac{x_i - \alpha}{\beta}\right)^{\delta} * [Ln(x_i - \alpha) - Ln\beta] = 0. \quad (24)$$

Solving (23) for β gives

$$\beta = \left[\Sigma\frac{(x_i - \alpha)^{\delta}}{d}\right]^{\frac{1}{\delta}} \quad (25)$$

Solving (22), (23) and (24) iteratively yields the m.l.e.'s $\hat{\alpha}$, $\hat{\beta}$ and $\hat{\delta}$. Allowing $\hat{\alpha}$ = first order statistic simplifies the calculations considerably.

5. Illustration

We examined a small set of data dealing with time to failure on an adjuvant breast cancer therapy trial. Twenty-one subjects received one therapy call it B. There were eight censored data points. Times ranged from 3.5 months to 137.4 months. The second group of women, Group C, totaled 20 with 12 failures and ranged from 1.5 months to 127 months. A Weibull fit to either set of data was appropriate.

In Table 1 we have the PWM for Group B computed according to (14) assuming d=21 failures and d=13 failures or 8 censored observations. Our choice of P_{jn} was (j-a) /n for a = 0.5, j = 1,... n, and in this case n = 21. In Table 2 we have the parameter estimates under the full and censored model derived from the PWM in equations (8) to (12). Also, we have chosen the location, α, to be equal to 0.0 or the first order statistic, 3.44. In order for the estimate of the shape parameter to be positive we restricted the denominators in (9) and (12) to be positive by adding 1 to the denominator.. This is reflected by resulting parentheses in Tables 2 and 5. When one compares the estimates in Table 2 to the m.l.e. estimates in Table 3 we see that for our example the PWM derived parameter estimates are not very accurate. The estimates of β by PWM, however, are within two standard errors of the m.l.e. estimates for the non censored model of d = 21 deaths. The estimate of β for the censored case by PWM for d = 0.0 can be transformed by a factor of $(n/d)^{1/\delta}$ to convert to approximate the m.l.e. of the censored case. Taking this transformation on 54.917 in Table 2 yields a value of 231.491 which is not within two standard errors of the estimate, 264.793, in Table 3. Thus dealing with censored data by PWM is not very promising for estimating β. Also, the estimates of the shape parameter, δ, by PWM are not accurate as seen in Tables 2 and 3. However, when all values in Table 2 are used as initial estimates for the m.l.e. procedure, conversion to the appropriate solutions were rapid in the calculation for equations (22), (23) and (24) resulting in Table 3.

Tables 4 to 6 repeat the results of Tables 1 to 3 for a different set of values which we label as Group C. Here we have 20 observations and 12 deaths. As before for the non-censored case the estimates of β in Table 5 using the PWM's of Table 4 are within two standard errors of the results of the m.l.e.'s in Table 6. The censored case as before for β was not as accurate. The δ values by PWM were not very reasonable, but closer to the m.l.e.'s in Table 6 than the previous Group B data. Also as above, when the values by PWM were used as initial estimates for solving for the m.l.e.'s the conversion to solutions as in Table 6 were rapid.

6. Conclusion

As can be seen the ultimate advantage for the PWM are their use as initial estimates for solving the maximum likelihood equations. Our two data sets did yield reasonable estimates by PWM for the parameter β in the non-censored cases, either with known or unknown location, α. Also in cases in which it is known that the shape parameter should yield an increasing hazard function one has to be aware of the restriction that the first PWM i.e. M_0 be restricted to greater than twice the next PWM i.e. M_1. Our challenge is to investigate the usage of these techniques in future applications and determine what further restrictions may be required to yield sensible solutions to our investigations.

7. References

Fraser, D.A.S. (1957). Nonparametric statistics. *New York: John Wiley.*

Greenwood, J. A., Landwehr, J. M., and Matalas, N. C.(1979). Probability weighted moments: definition and relation of parameters of several distributions expressible in inverse form. *Water Resources Research,* 15, 1049-1054.

Henderson, A. J. (1965). The Weibull distribution, unpublished M.S. Thesis, *Iowa State University*, Ames, Iowa.

Hosking, J. R. M., Wallis, J. R., and Wood, E. F. (1985). Estimation of the generalized extreme value distribution by the method of probability-weighted moment. *Technometrics*, 27, 251 - 261.

Johnson, M. L., and Kotz, S. (1970). Continuous univariate distributions-2. *Houghton Mifflin Company*, Boston.

Landwehr, J. M. and Matalas, N. C. (1979). Probability-weighted moments compared with some traditional techniques in estimating Gumbel parameters and quantiles. *Water Resources Research*. 1055 - 1064.

Locke, C. and Spurrier, J. D. (1976). The use of U- statistics for testing normality against non-symmetric alternatives. *Biometrika*. 63, 143 - 147.

Randles, R. H. and Wolfe, D.A. (1979). Introduction to the Theory of nonparamatric statistics. *New York: John Wiley*.

Shouki, M. M., Mian, I.U.H. and Tracy, D. S. (1988). Sampling properties of estimators of the log- logistic distribution with application to Canadian precipitation data. *The Canadian Journal of Statistics*. 16, 223 - 236.

Sinha, S. K. (1986). Reliability and life testing. *Wiley Eastern Limited, New Deihi.*

Weibull, W. (1938). Investigations into strength properties of brittle materials. *Proc. The Royal. Swedish Institute for Engr. #149.*

Weibull, W. (1981). Fatique testing and analysis of results. *New York, New York: Pergaman Press.*

Group	n	d	r	M_r
B	21	21	0 1 3	56.537 39.058 24.572
	21	13	0 1 3	91.328 63.093 39.693

Table 1: Probability Weighted Moments (PWM) For Group B.

n	d	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\delta}$
21	21	0.00 3.44	34.53 52.927	(1.024) 10.249
21	13	0.00 3.44	54.917 95.832	(0.899) 10.282

Table 2: Parameter Estimates For Group B Derived From The PWM

m	d	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\delta}$
21	21	0.00 3.44	45.936 38.652	0.297 0.271
21	13	0.00 3.44	264.793 281.695	0.272 0.238

Table 3: Maximum Likelihood Estimates For Group B Data

Group	n	d	r	M_r
C	20	20	0 1 3	40.809 30.069 20.165
	20	12	0 1 3	68.016 50.115 33.608

Table 4: Probability Weighted Moments (PWM) For Group C

n	d	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\delta}$
20	20	0.00 1.49	20.302 37.163	(.981) (.738)
20	12	0.00 1.49	64.531 62.879	(1.131) (0.737)

Table 5: Parameter Estimates For Group C Derived From The PWM

m	d	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\delta}$
20	20	0.00 1.49	29.424 25.416	0.286 0.263
20	12	0.00 1.49	187.936 171.560	0.275 0.268

Table 6: Maximum Likelihood Estimates For Group C Data