# *Estimation of Scale Parameter of Weibull distribution That has Exponential Family using T.O.M*

***By: Lecturer***

***Jubran A.A.K.***

***Department of Statistics and Information***

***College of Computer Science and Mathematics***

***University of Al-Qadisiya***

***E-mail: jubalaa98@yahoo.com***

**Summary**:

In this research we will estimate scale parameter of Weibull distribution that has exponential family using T.O.M (Term Omission Method) and compare it with other methods by MSE (Mean Square Error) with simulation.

**Keywords**: Term Omission Method (T.O.M), Exponential Family, Lease Square Method (LSM), Maximum Likelihood Estimator (MLE), Method of Moment (MOM)

## 1. Introduction

Estimators that based on T.O.M[5] deals with many distributions with discrete or continuous random variable that have exponential family, in our case we will take continues random variable with Weibull distribution.

## 2. Exponential Family

We note that there are many definitions of such type of that representation of exponential families. In this research, we will recall the regular type of exponential class.

## 3. Definitions

**Def. (1):** A family of discrete (continuous) random variables is called an exponential family if the probability density functions (probability mass functions) can be expressed in the form

$$f_X(x/\theta) = h(y)c(\theta)\exp\left(\sum_{i=1}^{k}\theta_i t_i(x)\right) \;,\; x = 0,1,2,... \qquad \dots (1)$$

for x in the common domain of the $f_X(x/\theta)$, $\theta \in R^k$.

Obviously h and c are non-negative functions. The $t_i(x)$ are real-valued functions of the observations.[3]

**Def. (2)**: Let $\vartheta$ be an interval on the real line. Let $\{f(x;\theta): \theta \in \vartheta\}$ be a family of pdf's (or pmf's). Assume that the set $\{\underline{x} : f(\underline{x};\theta) > 0\}$ is independent of θ, where $\underline{x} = (x_1, x_2, ..., x_n)$. We say that the family {f(x;θ): θ∈ϑ} is a one-parameter exponential family if there exist real-valued functions Q(θ) and D(θ) on ϑ and Borel-measurable functions T( $\underline{X}$ ) and S( $\underline{X}$ ) on $R^n$ such that

$$f(\underline{x};\theta) = \exp\left(Q(\theta)T(\underline{x}) + D(\theta) + S(\underline{x})\right) \quad \ldots (2)$$

if we write $f(\underline{x};\theta)$ as

$$f(\underline{x};\eta) = h(\underline{x})c(\eta)\exp\left(\eta T(\underline{x})\right) \quad \ldots (3)$$

where $h(\underline{x}) = \exp\left(S(\underline{x})\right)$, $\eta = Q(\theta)$, and $c(\eta) = \exp\left(D\left(Q^{-1}(\eta)\right)\right)$, then we call this the exponential family in canonical form for a natural parameter η.

**Def. (3)**: Let $\Theta \subseteq R^k$ be a k-dimensional interval. Let {f(x; $\boldsymbol{\theta}$): $\boldsymbol{\theta} \in \Theta$ } be a family of pdf's (or pmf's). Assume that the set $\{\underline{x} : f(\underline{x};\boldsymbol{\theta}) > 0\}$ is independent of $\boldsymbol{\theta}$, where $\underline{x} = (x_1, x_2, \ldots, x_n)$. We say that the family {f(x; $\boldsymbol{\theta}$): $\boldsymbol{\theta} \in \Theta$ } is a k-parameter exponential family if there exist real-valued functions $Q_1(\boldsymbol{\theta}), \ldots, Q_k(\boldsymbol{\theta})$ and $D(\boldsymbol{\theta})$ on $\Theta$ and Borel-measurable functions $T_1(\underline{X}), \ldots, T_k(\underline{X})$ and S( $\underline{X}$ ) on $R^n$ such that:[4]

$$f(\underline{x};\boldsymbol{\theta}) = \exp\left(\sum_{i=1}^{k} Q_i(\boldsymbol{\theta})T_i(\underline{x}) + D(\boldsymbol{\theta}) + S(\underline{x})\right) \quad \ldots (4)$$

**Def. (4)**: Exponential family is a class of distributions that all share the following form:

$$P(y/\eta) = h(y)\exp\{\eta^T T(y) - A(\eta)\} \quad \ldots (5)$$

* η is the natural parameter. For a given distribution η specifies all the parameters needed for that distribution.
* *T(y)* is the sufficient statistic of the data (in many cases T(y) = y, in which case the distribution is said to be in canonical form and η is referred to as the canonical parameter).
* A(η) is the log-partition function which ensures that p(y/η) remains a probability distribution.
* h(y) is the non-negative base measure (in many cases it is equal to 1).

Note that since η contains all the parameters needed for a particular distribution in its original form, we can express it with respect to the mean parameter θ:[2]

$$P(y/\theta) = h(y)\exp\{\eta(\theta)T(y) - A(\eta(\theta))\} \quad \ldots (6)$$

**Def. (5)**: (Regular Exponential Family): Consider a one-parameter family $\{f(x;\theta) : \theta \in \Omega)$ of probability density functions, where Ω is the interval set $\Omega = \{\theta : \gamma < \theta < \delta\}$, where γ and δ are known constants, and [7]

$$f(x;\theta) = \begin{cases} e^{p(\theta)k(x)+s(x)+q(\theta)} & a < x < b \\ 0 & o.w \end{cases} \quad \ldots (7)$$

The form (7) is said to be a member of the exponential class of probability density functions of the continuous type, if the following conditions are satisfied:

1) Neither (a) nor (b) depends upon θ. increasingly
2) p(θ) is a nontrivial continuous function of θ.
3) Each of k′(x)≠0 and s(x) is a continuous function of x.

and the following conditions with discrete random variable $X_i$:

1) The set $\{x : x = a_1, a_2, ...\}$ does not depend upon θ.
2) p(θ) is a nontrivial continuous function of θ.
3) k(x) is a nontrivial function of x.

## 4. Weibull Distribution that belong to Exponential Family

With continuous random variables that distributed Weibull, we can write the pdf as an exponential class form as follows:

The probability density function of our distribution of two parameters can be written as follows

$$f(x) = \frac{\alpha}{\beta} x^{\alpha-1} e^{-\left(\frac{x}{\beta}\right)^{\alpha}} \qquad x > 0$$

where β is a scale parameter and α is a shape parameter.
We can write the pdf of Weibull distribution of two parameters as follows

$$f(x) = \exp\left( \ln(\alpha) - \alpha \ln(\beta) + (\alpha - 1)\ln(x) - \left(\frac{x}{\beta}\right)^{\alpha} \right)$$

So the pdf of Weibull distribution of one parameter with known shape parameter α=a is given by:

$$f(x) = \exp\left( \ln(a) - a \ln(\beta) + (a - 1)\ln(x) - \left(\frac{x}{\beta}\right)^{a} \right)$$

$$f(x;\theta) = \begin{cases} e^{p(\theta)k(x) + s(x) + q(\theta)} & \text{a < x < b} \\ 0 & \text{o.w} \end{cases}$$

where $k(x) = x^{a}$ , $p(\beta) = -\beta^{-a}$ , $s(x) = (a-1)\ln(x)$ , $q(\beta) = \ln(a) - a\ln(\beta)$

## 5. T.O.M with Exponential Family

T.O.M[6] can be used to estimate the parameter, below we will derive the parameter by T.O.M of distributions (with one parameter) that have exponential family as follows:

For sample with size n having the p.d.f (p.m.f) f(x;θ), and for any two values $x_i$ and $x_j$, where $1 \le i < j \le n$,

$x_i$ $\qquad y_i = e^{k(x_i)p(\theta) + q(\theta) + s(x_i)}$

$x_j$ $\qquad y_j = e^{k(x_j)p(\theta) + q(\theta) + s(x_j)}$

by taking the natural logarithm to $y_i$ , $y_j$ we have

$x_i$ $\qquad k(x_i)p(\theta) + q(\theta) + s(x_i)$

$x_j$ $\qquad k(x_j)p(\theta) + q(\theta) + s(x_j)$

and by subtract the last result $y_j$ from $y_i$ we obtain

$$p(\theta)\left[k(x_j) - k(x_i)\right] + \left[s(x_j) - s(x_i)\right] = \ln y_j - \ln y_i$$

and again by subtract $\lfloor s(x_j) - s(x_i) \rfloor$ from both sides we get

$$p(\theta)\lfloor k(x_j) - k(x_i) \rfloor = \lfloor \ln y_j - \ln y_i \rfloor - \lfloor s(x_j) - s(x_i) \rfloor$$

Finally, by divided this amount over $\lfloor k(x_j) - k(x_i) \rfloor$ we have the function of θ, p(θ).

Therefore, we can define the $p^i(\theta)$ as follows:

$$p^i(\theta) = \frac{\lfloor \ln(y_j) - \ln(y_i) \rfloor - \lfloor s(x_j) - s(x_i) \rfloor}{\lfloor k(x_j) - k(x_i) \rfloor} \qquad \dots (8)$$

where $p^i(\theta)$ represents the values that we have from previous steps of T.O.M, ∀i=1,2,…,n-1. Thus from eq. (8) we can educe values of $\theta^i$ from $p^i(\theta)$.

Therefore the estimation of θ can found now by the least square error with the following equation:

$$\hat{\theta} = Min\left( \sum_{m=1}^{n} \left[ f(x_m, \theta^i) - y_m \right] \right)^2 \qquad \text{i} = 1,2,...,\text{n} - 1$$

where $f(x_m, \theta^i)$ is the value of function $f(x_m)$ on $\theta^i$ , and $y_m = y(x_m)$ is the observed value on $x_m$.

## 6. T.O.M with Exponential Family of Weibull Distribution

Recalling the exponential family of Weibull distribution with scale parameter β and known α=a in section (4)

$$f(x) = \exp\left(\ln(a) - a\ln(\beta) + (a-1)\ln(x) - \beta^{-a} x^a\right) \qquad \dots (9)$$

where $k(x) = x^{\text{a}}$ , $\text{p}(\beta) = -\beta^{-a}$ , $\text{s(x)} = (a-1)\ln(x)$ , $\text{q}(\beta) = \ln(a) - a\ln(\beta)$

therefore by using (8) we can write

$$p^i(\beta) = \frac{[\ln(y_j) - \ln(y_i)] - (a-1)[\ln(x_j) - \ln(x_i)]}{x_j^a - x_i^a}$$

or simply

$$p^i(\beta) = \frac{\lfloor \ln(y_j) - (a-1)x_j \rfloor - \left[ \ln(y_i) - (a-1)x_i \right]}{\left[ x_j^a - x_i^a \right]} \qquad \dots (10)$$

If we represent the right term of the eq. (10) by $k_i$, for 1≤ i ≤n-1, then we can here educe the $\beta^i$ , 1≤ i ≤n-1 from $p^i(\beta)$ to have the estimating of α, thus

$$p^i(\beta) = k_i \Rightarrow -\beta^{-a} = k_i \Rightarrow \beta^i = -k_i^{-a} \;,\; 1 \le i \le n-1$$

## 7. Results

The quantitative comparison of different estimators, Mean Square Error (MSE0 was used to test estimators of the four methods by the 1000 frequency (LSM, MLE, MOM [1], T.O.M)

The following table represents the comparison between the methods (LSM, MLE, MOM [1], T.O.M) for Weibull distribution and choose the best method according to MSE (Mean Square Error) with equation below:

$$MSE = \frac{\sum_{i=1}^{n} \left[ f(x_i; \theta) - f(x_i; \hat{\theta}) \right]^2}{n}$$

where $f(x_i; \theta), f(x_i; \hat{\theta})$ represent to the probability function of variable $x_i$, $\theta$ , $\hat{\theta}$ parameter and estimating parameter respectively.

**Table**. The estimate values of parameter that founding by LSM, MLE, MOM and T.O.M for Weibull distribution with β scale parameter and shape parameter α=1.

| β | n | T.O.M<br>MSE | LSM<br>MSE | MLE<br>MSE | MOM<br>MSE |
|---|---|---|---|---|---|
| 10 | 20 | 8.787 | 12.5 | 10.697 | 10.697 |
| | | 0.32449851 | 0.33121188 | 0.32801263 | 0.32801263 |
| | 40 | 10.001214 | 10.5 | 10.113 | 10.113 |
| | | 9.58E-11 | 2.1403E-06 | 1.3132E-07 | 1.3132E-07 |
| | 60 | 10.011619 | 9.84 | 9.4378 | 9.4378 |
| | | 8.28E-10 | 1.068E-07 | 1.5062E-06 | 1.5062E-06 |
| | 80 | 10.000142 | 8.44 | 9.3112 | 9.3112 |
| | | 4.74E-13 | 1.1703E-05 | 1.9778E-06 | 1.9778E-06 |
| | 100 | 10.000173 | 8.92 | 10.645 | 10.645 |
| | | 4.41E-13 | 4.8879E-06 | 1.2997E-06 | 1.2997E-06 |
| 145 | 20 | 147.69688 | 156 | 131.49 | 131.49 |
| | | 1.68E-08 | 6.2998E-08 | 4.9226E-08 | 4.9226E-08 |
| | 40 | 145.28894 | 183 | 154.89 | 154.89 |
| | | 6.89E-11 | 8.8589E-08 | 7.7989E-09 | 7.7989E-09 |
| | 60 | 145.76915 | 178 | 182.37 | 182.37 |
| | | 1.34E-10 | 8.4642E-08 | 1.0426E-07 | 1.0426E-07 |
| | 80 | 145.04504 | 134 | 135.63 | 135.63 |
| | | 4.71E-13 | 7.024E-09 | 4.6052E-09 | 4.6052E-09 |
| | 100 | 145.22579 | 164 | 158.65 | 158.65 |
| | | 4.65E-12 | 2.0437E-08 | 1.1436E-08 | 1.1436E-08 |
| 357 | 20 | 373.59027 | 455 | 466.91 | 466.91 |
| | | 1.45E-08 | 7.3112E-08 | 8.4619E-08 | 8.4619E-08 |
| | 40 | 360.02868 | 460 | 408.45 | 408.45 |
| | | 1.01E-10 | 2.1877E-08 | 6.779E-09 | 6.779E-09 |
| | 60 | 358.52417 | 362 | 406.22 | 406.22 |
| | | 6.76E-11 | 2.2969E-10 | 1.0095E-08 | 1.0095E-08 |
| | 80 | 357.44195 | 330 | 329.57 | 329.57 |
| | | 1.16E-12 | 2.0025E-09 | 2.0894E-09 | 2.0894E-09 |
| | 100 | 359.89547 | 335 | 370.34 | 370.34 |
| | | 1.61E-11 | 9.2952E-10 | 2.9692E-10 | 2.9692E-10 |
| 872 | 20 | 974.45364 | 1292 | 1144.6 | 1144.6 |
| | | 1.43E-08 | 3.9394E-08 | 2.6806E-08 | 2.6806E-08 |
| | 40 | 886.98981 | 1305 | 1219.5 | 1219.5 |
| | | 1.02E-10 | 1.1476E-08 | 8.3459E-09 | 8.3459E-09 |
| | 60 | 873.31129 | 1003 | 864.62 | 864.62 |
| | | 7.67E-13 | 1.3192E-09 | 3.772E-12 | 3.772E-12 |
| | 80 | 873.99156 | 758 | 770.18 | 770.18 |
| | | 6.93E-13 | 9.2891E-1 | 7.1553E-10 | 7.1553E-10 |
| | 100 | 873.31561 | 724 | 796.85 | 796.85 |
| | | 5.92E-13 | 2.5595E-09 | 5.5785E-10 | 5.5785E-10 |
| 1270 | 20 | 1289.6028 | 1187 | 1137.7 | 1137.7 |
| | | 1.83E-10 | 3.4705E-10 | 1.0659E-09 | 1.0659E-09 |
| | 40 | 1335.6016 | 1001 | 1084 | 1084 |
| | | 3.19E-10 | 3.5389E-09 | 1.3805E-09 | 1.3805E-09 |
| | 60 | 1273.4793 | 1208 | 1041.5 | 1041.5 |
| | | 8.67E-13 | 6.6701E-11 | 1.2037E-09 | 1.2037E-09 |
| | 80 | 1271.8136 | 1246 | 1304.9 | 1304.9 |
| | | 4.64E-13 | 1.0114E-11 | 2.7383E-11 | 2.7383E-11 |
| | 100 | 1271.8341 | 1592 | 1514.1 | 1514.1 |
| | | 3.04E-13 | 7.6538E-10 | 4.7841E-10 | 4.7841E-10 |

## 8. Conclusion

In this paper, we have presented analytical methods for estimating scale parameter of Weibull distribution (T.O.M., LSM, MLE, and MOM). It has been shown from the computational results that T.O.M is the best estimate method for scale parameter. Moreover, we can have an exact estimation when data sample exact fitting the distribution.

## References:


[1] Al-Fawzan, Mohammad A., 2000. “Methods for Estimating the Parameters of the Weibull Distribution”, King Abdulaziz City for Science and Technology, p. 4-10.

[2] Andreas Vlachos, 2010, “Notes on exponential family distributions and generalized linear models”.

[3] Clark, David R. and Thayer, Charles A. 2004. “A Primer on the Exponential Family Distributions”, p. 117-121.

[4] Jurgen, S., 1999. “Mathematical Statistics I”, Utah State University, p: 120.

[5] Labban, J.A., 2005. “Modified and Simplified Method for Finite Difference (Divided)”. Al-Qadisiya Journal for Science, Vol. 10, No. 2, pp. 244-251.

[6] Labban, J.A., 2009. “Approximation of Scale Parameter of Inverted Gamma Distribution by TOM Modified”. Accepted for Publishing

[7] Robert V. Hogg and Allen T. Craig, 1978. “Introduction to Mathematical Statistics”, 4th Edition, p: 357.