Maximum Likelihood Estimates

Weibull Distribution

© Dennis Bricker
Dept of Mechanical & Industrial Engineering
The University of Iowa

Weibull Distribution:

pdf: $$f(t) = \frac{k}{u}\left(\frac{t}{u}\right)^{k-1} \exp\left\{-\left(\frac{t}{u}\right)^{k}\right\}$$

Suppose t_1, t_2, ... t_n are *times to failure* of a group of n mechanisms.

The *likelihood function* is

$$L(t;k,u) = \prod_{i=1}^{n} \frac{k}{u}\left(\frac{t_i}{u}\right)^{k-1} \exp\left\{-\left(\frac{t_i}{u}\right)^{k}\right\}$$

$$= \frac{k^n}{u^{nk}}\left[\prod_{i=1}^{n} t_i^{k-1}\right] \exp\left\{-u^{-k}\sum_{i=1}^{n} t_i^{k}\right\}$$

$$L(t;k,u) = \frac{k^n}{u^{nk}} \left[\prod_{i=1}^{n} t_i^{k-1} \right] \exp\left\{ -u^{-k} \sum_{i=1}^{n} t_i^k \right\}$$

We wish to choose values of u & k which *maximize* L (or equivalently, the logarithm of L), i.e., which make the observed values of t as large as possible!

The log-likelihood function is

$$\ln L(t;k,u) = n \ln k - nk \ln u + (k-1) \sum_{i=1}^{n} \ln t_i - u^{-k} \sum_{i=1}^{n} t_i^k$$

The *optimality conditions* for the maximum of the log-likelihood function are

$$\begin{cases} \dfrac{\partial}{\partial u} \ln L(t;u,k) = 0 \\ \dfrac{\partial}{\partial k} \ln L(t;u,k) = 0 \end{cases}$$

This gives us a pair of nonlinear equations in two unknowns (u & k):

$$\begin{cases} -\dfrac{n\hat{k}}{\hat{u}} + \hat{k}\hat{u}^{-\hat{k}-1} \displaystyle\sum_{i=1}^{n} t_i^{\hat{k}} = 0 \\ \dfrac{n}{\hat{k}} - n \ln \hat{u} + \displaystyle\sum \left(\dfrac{t_i}{\hat{u}}\right)^{\hat{k}} \ln\left(\dfrac{t_i}{\hat{u}}\right) = 0 \end{cases}$$

But the left side of the first equation can be factored:

$$-\frac{n\hat{k}}{\hat{u}}+\hat{k}\hat{u}^{-\hat{k}-1}\sum_{i=1}^{n}t_i^{\hat{k}}=0\Rightarrow\hat{k}\hat{u}^{-1}\left[-n+\hat{u}^{-\hat{k}}\sum_{i=1}^{n}t_i^{\hat{k}}\right]=0$$

Since the first factor *cannot* be zero, we set the second

factor equal to zero and solve for $\hat{u}$ in terms of $\hat{k}$:

$$\hat{u}=\left(\frac{1}{n}\sum_{i=1}^{n}t_i^{\hat{k}}\right)^{1/\hat{k}}$$

Eliminating $\hat{u}$ in the *second* equation by substituting the first, we get the following nonlinear equation in $\hat{k}$ alone:

$$\frac{1}{\hat{k}}-\frac{\sum_{i=1}^{n}t_i^{\hat{k}}\ln t_i}{\sum_{i=1}^{n}t_i^{\hat{k}}}+\frac{1}{n}\sum_{i=1}^{n}\ln t_i=0$$

This can now be solved by, for example, the *secant method*.

Maximum Likelihood Estimation

with "censored" data

Suppose that an experiment was terminated at time τ after only r of the n units in a lifetest had failed. This is accounted for by defining the likelihood as

$$L(t,\theta) = \left[1 - F(\tau;\theta)\right]^{n-r} \times \prod_{i=1}^{r} f(t_i;\theta)$$

The log-likelihood function is therefore

$$\ln L(t;\theta) = (n-r)\ln\left[1 - F(t;\theta)\right] + \sum_{i=1}^{r} \ln f(t_i;\theta)$$

Example: MLE of Weibull parameters, given censored data

The CDF of the Weibull distribution is

$$F(t;k,u) = 1 - \exp\left\{-\left(\frac{t}{u}\right)^k\right\}$$

and so the likelihood function is

$$L(t;k,u) = \left[\exp\left\{-\left(\frac{\tau}{u}\right)^k\right\}\right]^{n-r} \times \prod_{i=1}^{r} \frac{k}{u}\left(\frac{t_i}{u}\right)^{k-1} \exp\left\{-\left(\frac{t_i}{u}\right)^k\right\}$$

$$= \frac{k^r}{u^{nk}}\left[\prod_{i=1}^{r} t_i^{k-1}\right] \exp\left\{-u^{-k}\left[\sum_{i=1}^{r} t_i^k + (n-r)\tau^k\right]\right\}$$

The log-likelihood function is

$$\ln L(t;k,u) = r\ln k - nk\ln u + (k-1)\sum_{i=1}^{r}\ln t_i - u^{-k}\left[\sum_{i=1}^{r} t_i^k + (n-r)\tau^k\right]$$

The ***optimality*** conditions for a maximum of the log-likelihood at $(\hat{k},\hat{u})$ are

$$\begin{cases} \dfrac{\partial}{\partial u}\ln L(t;\hat{u},\hat{k}) = 0 \\ \dfrac{\partial}{\partial k}\ln L(t;\hat{u},\hat{k}) = 0 \end{cases}$$

A result similar to the uncensored case can be derived:

$$\hat{u} = \left(\frac{\sum_{i=1}^{r} t_i^{\hat{k}} + (n-r)\tau^{\hat{k}}}{n} \right)^{1/\hat{k}}$$

and

$$\frac{1}{\hat{k}} - \frac{\sum_{i=1}^{r} t_i^{\hat{k}} \ln t_i + (n-r)\tau^{\hat{k}} \ln \tau}{\sum_{i=1}^{r} t_i^{\hat{k}} + (n-r)\tau^{\hat{k}}} + \frac{1}{r}\sum_{i=1}^{r} \ln t_i = 0$$

This *second* equation can be solved for $\hat{k}$ by the **secant** method, and

then $\hat{k}$ can be used to calculate $\hat{u}$ by the *first* equation.

EXAMPLE: Twenty devices are tested simultaneously until 500 days have passed, at which time the following failure times (in days) have been recorded:

```
 31.5
 74.0
 87.5
100.1
103.3
181.9
279.9
297.1
462.5
465.4
```

Estimate the lifetime for which the device is 90% reliable.

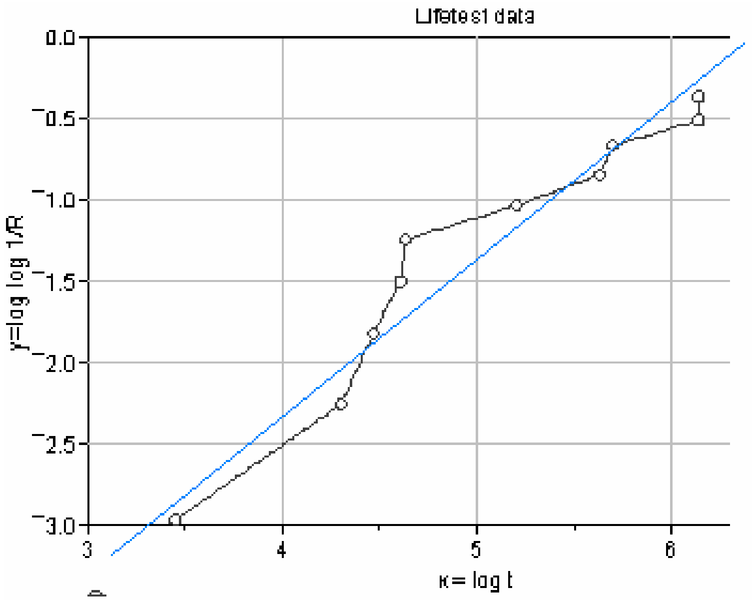
A plot of Y vs X, obtained by the transformations:

$Y = \log\log\frac{1}{R(t)}$ where *R(t)* is the observed fraction of the devices which have survived until time *t*, and

$X = \log t$

should be a line if the Weibull model were to fit the data perfectly.

LEAST SQUARES REGRESSION RESULTS:



u (scale parameter) = 653.504

k (shape parameter) = 0.908313

so that

mean = 630.396

standard deviation = 754.336

Note: this is determined by minimizing the sum of the squared errors in the linearized version of $F(t) = 1 - e^{-\left(t/u\right)^k}$,

namely $y = kx - k\ln u$ *where* $x = \ln t$ & $y = \ln\ln\frac{1}{R(t)}$,

rather than in the original equation!

If we use these parameters found by linear regression, the reliability function would have the values:

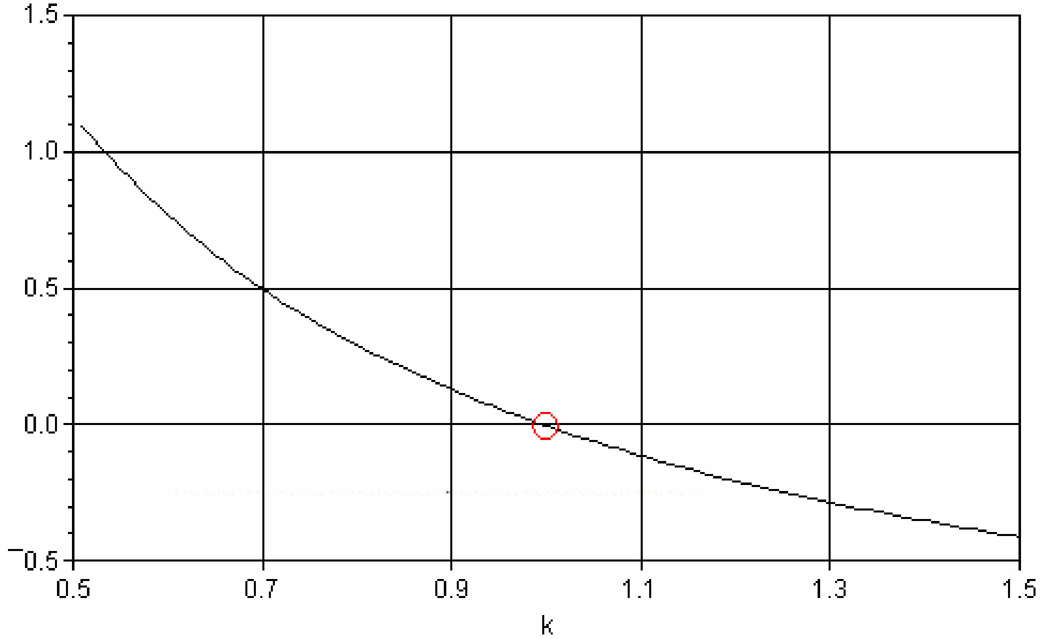
t	F(t)	1-F(t)
4.12824	0.01	0.99
8.90435	0.02	0.98
13.993	0.03	0.97
19.3163	0.04	0.96
24.837	0.05	0.95
30.5337	0.06	0.94
36.3925	0.07	0.93
42.4042	0.08	0.92
48.5623	0.09	0.91
54.8622	0.10	0.90

Hence, according to this model, 90% of the devices should be operating at 54.8 (approximately **55**) days.

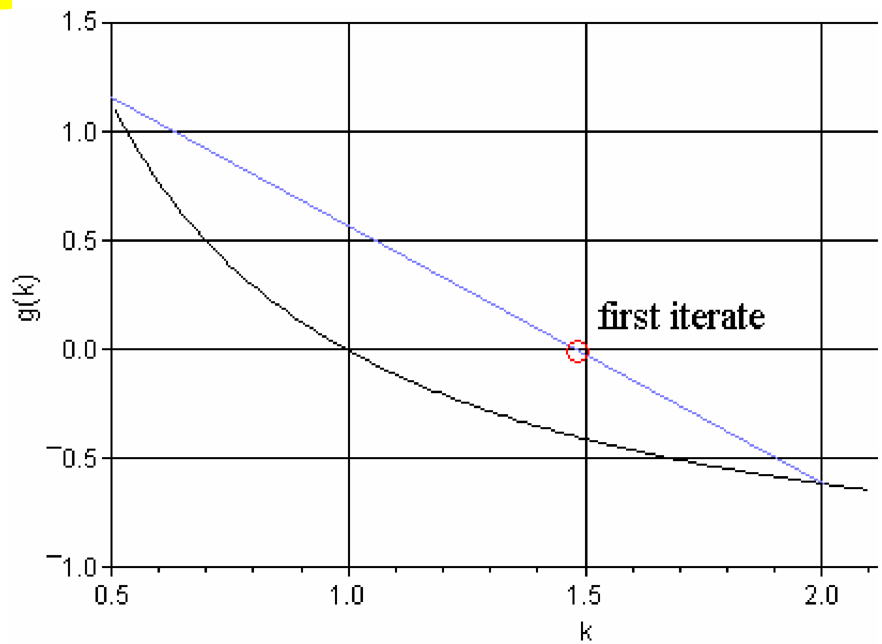
Maximum Likelihood result:

Solving the nonlinear equation for k:

$$g(k) = \frac{1}{\hat{k}} - \frac{\sum_{i=1}^{r} t_i^{\hat{k}} \ln t_i + (n-r)\tau^{\hat{k}} \ln \tau}{\sum_{i=1}^{r} t_i^{\hat{k}} + (n-r)\tau^{\hat{k}}} + \frac{1}{r}\sum_{i=1}^{r} \ln t_i = 0$$



SECANT METHOD



If our first two "guesses" at the value of k are 0.5 and 2.0, then we determine that

$$g(0.5) = 1.13739 \ \&\ \text{and}\ g(2.0) = -0.618085.$$

The secant joining the two points on the graph of g cross the k axis at 1.47187.

We then repeat, with the 2 improved “guesses” k=0.5 and k=1.47187.

SECANT METHOD RESULTS:

k	error
0.5	1.13739
2.0	¯0.618085
1.47187	¯0.397478
0.5203	1.05148
1.21083	¯0.217582
1.09244	¯0.108608
0.974445	0.025006
0.996528	¯0.00227829
0.994684	¯0.0000438242
0.994648	7.83302E¯8
0.994648	¯2.68896E¯12

Once we determine the value of $\hat{k}$ which maximizes the likelihood function, then the corresponding value of the parameter $\hat{u}$ is found by

$$\hat{u} = \left(\frac{\sum_{i=1}^{r} t_i^{\hat{k}} + (n-r)\tau^{\hat{k}}}{n} \right)^{1/\hat{k}}$$

MAXIMUM LIKELIHOOD RESULT:

u (scale parameter) = 710.339,
k (shape parameter) = 0.994648

t	F(t)	1-F(t)
6.9646	0.01	0.99
14.0526	0.02	0.98
21.2337	0.03	0.97
28.5026	0.04	0.96
35.8579	0.05	0.95
43.2993	0.06	0.94
50.8272	0.07	0.93
58.4427	0.08	0.92
66.1467	0.09	0.91
73.9408	0.10	0.90

According to this model, then, 90% of the devices should be operating at 73.94 (approximately **74**) days.