TMA4275 LIFETIME ANALYSIS

Slides 9: Parametric inference for the Weibull model

Bo Lindqvist
Department of Mathematical Sciences
Norwegian University of Science and Technology
Trondheim

http://www.math.ntnu.no/∼bo/
bo@math.ntnu.no

NTNU, Spring 2015

In Slides 9 we consider parametric statistical inference for lifetime distributions with two (or more) parameters. The Weibull-distribution is the example throughout, but most of the theory is valid for any lifetime distribution with two or more parameters.

- Likelihood analysis for Weibull distribution
  - Likelihood function
  - Maximum likelihood estimation
  - Profile log likelihood
  - Observed information and standard error
  - MINITAB output
  - Confidence intervals (CI)
    - Standard intervals for θ and α
    - Profile-log likelihood interval for α (the 1.92-interval)
  - Testing for exponential vs. Weibull distribution
  - Probability plotting for Weibull and exponential distribution

A lifetime T is distributed as Weibull(θ, α) if

$$f(t;\theta,\alpha) = \frac{\alpha}{\theta^\alpha} t^{\alpha-1} e^{-(\frac{t}{\theta})^\alpha}, \quad R(t;\theta,\alpha) = e^{-(\frac{t}{\theta})^\alpha}$$

The likelihood and log-likelihood for right censored data are hence given by

$$\begin{aligned}
L(\theta,\alpha) &= \prod_{i:\delta_i=1} \frac{\alpha}{\theta^\alpha} y_i^{\alpha-1} e^{-(\frac{y_i}{\theta})^\alpha} \prod_{i:\delta_i=0} e^{-(\frac{y_i}{\theta})^\alpha} \\
&= \prod_{i:\delta_i=1} \frac{\alpha}{\theta^\alpha} y_i^{\alpha-1} \prod_{i=1}^{n} e^{-(\frac{y_i}{\theta})^\alpha} \\
&= \frac{\alpha^r}{\theta^{r\alpha}} \left(\prod_{i:\delta_i=1} y_i \right)^{\alpha-1} e^{-\frac{1}{\theta^\alpha} \sum_{i=1}^{n} y_i^\alpha} \\
\ell(\theta,\alpha) &= r \ln \alpha - \alpha r \ln \theta + (\alpha - 1) \sum_{i:\delta_i=1} \ln y_i - \frac{1}{\theta^\alpha} \sum_{i=1}^{n} y_i^\alpha,
\end{aligned}$$

where $r = \sum_{i=1}^{n} \delta_i$ is the number of failures.

The maximum likelihood estimators $\hat{\theta}, \hat{\alpha}$ are found by solving the likelihood equations (two equations in two unknowns):

1. $$\frac{\partial \ell(\theta,\alpha)}{\partial \theta} = -\frac{\alpha r}{\theta} + \frac{\alpha}{\theta^{\alpha+1}} \sum_{i=1}^{n} y_i^{\alpha} = 0$$

2. $$\frac{\partial \ell(\theta,\alpha)}{\partial \alpha} = \frac{r}{\alpha} - r \ln \theta + \sum_{i=1}^{n} \delta_i \ln y_i - \sum_{i=1}^{n} \left(\frac{y_i}{\theta}\right)^{\alpha} \ln\left(\frac{y_i}{\theta}\right) = 0$$

We may for example solve eq. 1 for θ as a function of α and substitute this into eq. 2. From eq. 1 we get

$$\begin{aligned}
\frac{\alpha r}{\theta} &= \frac{\alpha}{\theta^{\alpha+1}} \sum_{i=1}^{n} y_i^{\alpha} \\
\theta^{\alpha} &= \frac{\sum_{i=1}^{n} y_i^{\alpha}}{r} \\
\theta &= \left(\frac{\sum_{i=1}^{n} y_i^{\alpha}}{r}\right)^{1/\alpha}.
\end{aligned}$$

This is the maximum likelihood estimator of θ if the value of α is *known*. Denote it by

$$\hat{\theta}(\alpha) = \left(\frac{\sum_{i=1}^{n} y_i^{\alpha}}{r}\right)^{1/\alpha}.$$

If $\alpha = 1$, this is the MLE for the exponential model.

Recall from last slide that we can find $(\hat{\theta}, \hat{\alpha})$ by substituting $\hat{\theta}(\alpha)$ for θ into equation 2 and thereby get an equation with α as the only unknown.

The solution of this is $\hat{\alpha}$, while $\hat{\theta} = \hat{\theta}(\hat{\alpha})$ is finally the MLE of θ.

Data (* means right censored observation):

0.35, 0.50*, 0.75*, 1.00, 1.30, 1.80, 3.00*, 3.15*, 4.85*, 5.50, 5.50*, 6.25*

```
Estimation Method:  Maximum Likelihood Distribution:  Weibull


Parameter Estimates
                             Standard          95,0% Normal CI
Parameter    Estimate           Error        Lower        Upper
Shape          0,9780          0,3694       0,4665       2,0504
Scale
                6,880           3,517        2,526       18,740


Log-Likelihood = -14,576
```

From this we estimate the mean time to failure by

$$\widehat{\text{MTTF}} = \hat{\theta}\,\Gamma(\frac{1}{\hat{\alpha}} + 1) = 6.88\,\Gamma(\frac{1}{0.978} + 1) = 6.88\Gamma(2.0225) = 6.9469$$

We may also use $\hat{\theta}(\alpha)$ in to obtain the so called *profile log likelihood function* for α.

We then substitute $\hat{\theta}(\alpha)$ into the log likelihood function $\ell(\theta, \alpha)$ to get

$$\tilde{\ell}(\alpha) = \ell(\hat{\theta}(\alpha), \alpha) \equiv \max_\theta \ell(\theta, \alpha).$$

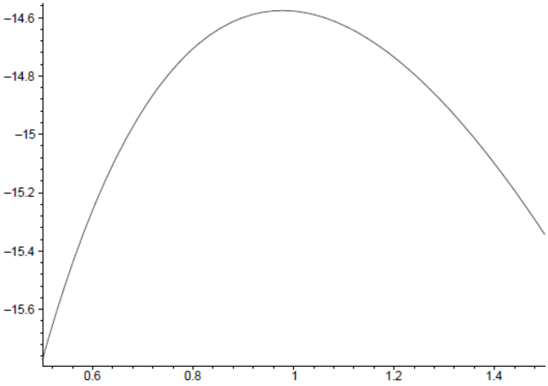
Note that $\tilde{\ell}(1) = \ell(\hat{\theta}(1), 1)$ equals the maximum value of the log likelihood $\ell(\theta)$ for the exponential distribution.

For the Weibull case we compute $\tilde{\ell}(\alpha)$ as follows,

$$\begin{aligned}
\tilde{\ell}(\alpha) &= r \ln \alpha - \alpha r \ln \left[\left(\frac{\sum y_i^\alpha}{r} \right)^{1/\alpha} \right] + (\alpha - 1) \sum \delta_i \ln y_i - \frac{1}{\left(\frac{\sum y_i^\alpha}{r} \right)} \sum_{i=1}^{n} y_i^\alpha \\
&= r \ln \alpha - \alpha r \frac{1}{\alpha} \ln(\sum y_i^\alpha) + \alpha r \frac{1}{\alpha} \ln r + (\alpha - 1) \sum \delta_i \ln y_i - r \\
&= r \ln \alpha - r \ln(\sum y_i^\alpha) + r \ln r + (\alpha - 1) \sum \delta_i \ln y_i - r.
\end{aligned}$$

Data:

0.35, 0.50*, 0.75*, 1.00, 1.30, 1.80, 3.00*, 3.15*, 4.85*, 5.50, 5.50*, 6.25*



Graph of $\tilde{\ell}(\alpha)$. Maximum is at $\hat{\alpha} = 0.978$.
From this, $\hat{\theta}$ is given as $\hat{\theta}(\hat{\alpha}) = 6.880$.

Recall that for the case of a single parameter θ we defined the observed information $I(\hat{\theta}) = -\ell''(\hat{\theta})$ and obtained the estimator $\widehat{Var(\hat{\theta})} = 1/I(\hat{\theta})$. In the case of two parameters (θ, α) we define the *observed information* **matrix** (also called *Hessian* matrix)to be

$$I(\hat{\theta}, \hat{\alpha}) =_{def} \begin{bmatrix} -\frac{\partial^2 \ell(\theta,\alpha)}{\partial \theta^2} & -\frac{\partial^2 \ell(\theta,\alpha)}{\partial \theta \partial \alpha} \\ -\frac{\partial^2 \ell(\theta,\alpha)}{\partial \alpha \partial \theta} & -\frac{\partial^2 \ell(\theta,\alpha)}{\partial \alpha^2} \end{bmatrix}_{\theta=\hat{\theta}, \alpha=\hat{\alpha}}$$

By the theory of maximum likelihood,

$$\left[I(\hat{\theta}, \hat{\alpha})\right]^{-1} = \begin{bmatrix} \widehat{Var(\hat{\theta})} & \widehat{Cov(\hat{\theta}, \hat{\alpha})} \\ \widehat{Cov(\hat{\alpha}, \hat{\theta})} & \widehat{Var(\hat{\alpha})} \end{bmatrix}$$

Thus, by inverting the observed information matrix we get a matrix with the estimated variances of the parameters on the diagonal.

Standard errors, $\widehat{\mathrm{SD}(\hat{\theta})}$ and $\widehat{\mathrm{SD}(\hat{\alpha})}$, are computed by taking square roots.

The standard errors $\widehat{\mathrm{SD}(\hat{\theta})}$ and $\widehat{\mathrm{SD}(\hat{\alpha})}$ can be used to compute either *standard confidence intervals* or *standard confidence intervals for positive parameters*. The latter are given as

$$\hat{\theta} e^{\pm 1.96 \frac{\widehat{\mathrm{SD}(\hat{\theta})}}{\hat{\theta}}}, \qquad \hat{\alpha} e^{\pm 1.96 \frac{\widehat{\mathrm{SD}(\hat{\alpha})}}{\hat{\alpha}}}$$

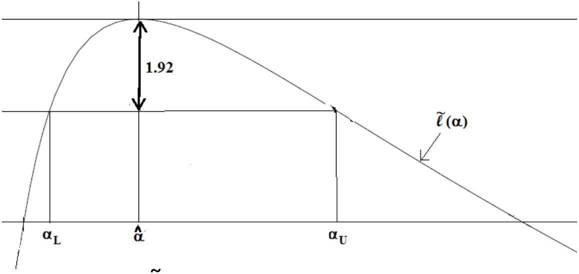
The inverse of the observed information matrix can also be used to compute standard errors and confidence intervals for quantities which involve both θ and α, e.g., MTTF, median and percentiles t_p. For this we need the covariance of $(\hat{\theta}, \hat{\alpha})$.

We will not do this in more detail, but MINITAB does the computations!

Exercise: Show that for Weibull(θ, α) we have

$$t_p = \theta(-\ln(1-p))^{1/\alpha}$$

```
Distribution:   Weibull


Parameter Estimates

                        Standard    95,0% Normal CI
Parameter   Estimate       Error      Lower     Upper
Shape       0,977997    0,369395   0,466481   2,05041
Scale        6,88032     3,51735    2,52615   18,7395

Log-Likelihood = -14,576

Goodness-of-Fit
Anderson-Darling (adjusted) = 30,049


Characteristics of Distribution

                                         Standard    95,0% Normal CI
                             Estimate       Error      Lower     Upper
Mean(MTTF)                    6,94720     4,20887    2,11895   22,7772
Standard Deviation            7,10402     6,40851    1,21238   41,6265
Median                        4,72991     2,20169    1,89948   11,7780
First Quartile(Q1)            1,92463     1,00544   0,691314   5,35822
Third Quartile(Q3)            9,60850     5,56069    3,09060   29,8723
Interquartile Range(IQR)      7,68386     5,24523    2,01616   29,2843
```



The profile log likelihood $\tilde{\ell}(\alpha)$ can be used in the same manner as the log likelihood for a single parameter to obtain a confidence interval for α. This is because likelihood theory tells us that if α is the true parameter value,

$$W(\alpha) = 2\big(\tilde{\ell}(\hat{\alpha}) - \tilde{\ell}(\alpha)\big) = 2\big(\ell(\hat{\theta}, \hat{\alpha}) - \ell(\hat{\theta}(\alpha), \alpha)\big) \approx \chi_1^2$$

Here $\tilde{\ell}(\hat{\alpha}) = \ell(\hat{\theta}, \hat{\alpha})$ can be read off directly from the MINITAB output.

An approximate 95% confidence interval for α can be found in the example as the set of α for which $\tilde{\ell}(\alpha) \geq -14.58 - 1.92 = -16.50$ (which unfortunately is not covered by the graph on the earlier slide).

The result

$$W(\alpha) = 2\big(\tilde{\ell}(\hat{\alpha}) - \tilde{\ell}(\alpha)\big) \approx \chi_1^2$$

can be used to test whether the data come from an exponential distribution. This is done by testing

$$H_0 : \alpha = 1 \quad \text{vs.} \quad H_1 : \alpha \neq 1,$$

and using the following test statistic:

$$W(1) = 2\big(\tilde{\ell}(\hat{\alpha}) - \tilde{\ell}(1)\big)$$

which is approximately χ_1^2 if the null hypothesis is true. Hence we reject the null hypothesis at 5% significance level if $W(1) \geq 3.84$. Note that

$$W(1) = 2\Big[\underbrace{\ell(\hat{\theta}, \hat{\alpha})}_{\text{max value in Weibull model}} - \underbrace{\ell(\hat{\theta}(1), 1)}_{\text{max value in exponential model}}\Big]$$

where we can find the two maxima in the Minitab output for, respectively, Weibull-distribution and exponential distribution. In the present case we get $\ell(\hat{\theta}, \hat{\alpha}) = -14.576$, and $\ell(\hat{\theta}(1), 1) = -14.577$. Thus, $W(1) = 2(-14.576 - (-14.577)) = 0.002$ and *we do not reject* H_0.

Suppose $T \sim \text{Weibull}(\theta, \alpha)$, so that $R(t) = e^{-\left(\frac{t}{\theta}\right)^{\alpha}}$. Then

$$
\begin{aligned}
-\ln R(t) &= \left(\frac{t}{\theta}\right)^{\alpha} \\
\ln(-\ln R(t)) &= \alpha \ln t - \alpha \ln \theta.
\end{aligned}
$$

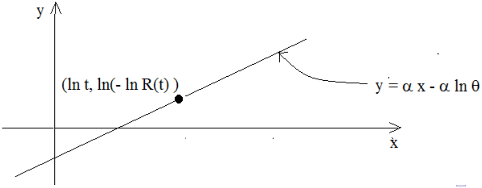
Thus, for any t, the point

$$\big(\ln t, \ln(-\ln R(t))\big)$$

is on the line

$$y = \alpha x - \alpha \ln \theta$$

in the (x, y) coordinate system. Thus α is the slope of the line, while $-\alpha \ln \theta$ is the intercept.



Suppose we have right-censored data $(y_1, \delta_1), \ldots, (y_n, \delta_n)$, where $t_{(1)} < \ldots < t_{(k)}$ are the observed *failure* times (i.e. uncensored times). Then we can compute the Kaplan-Meier estimator $\hat{R}(t)$ for $R(t)$.

The *Weibull probability plot* is a plot of the points

$$\left(\ln t_{(i)}, \ln(-\ln \hat{R}(t_{(i)})) \right)$$

in the (x, y) coordinate system.

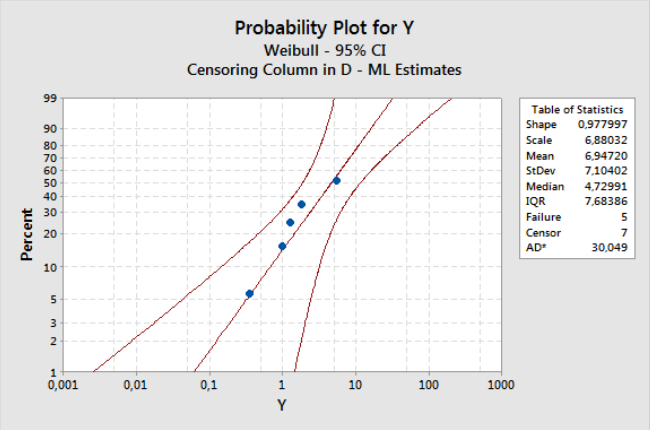
If the Weibull model is the correct one, then these points will tend to be close to a straight line since $\hat{R}(t)$ then is supposed to be close to the underlying Weibull survival function.

In practice is often used a modified KM-estimator,

$$\hat{\hat{R}}(t_{(i)}) = \frac{\hat{R}(t_{(i)}) + \hat{R}(t_{(i-1)})}{2}.$$

This gives a somewhat more "smooth" KM-curve, and avoids $\ln 0$.

Minitab plots the points $\left(\ln t_{(i)}, \ln(-\ln \hat{\hat{R}}(t_{(i)}))\right)$ together with the straight line $y = \hat{\alpha}x - \hat{\alpha}\ln\hat{\theta}$. Note that the scales on the axis are displayed on log-scale (x-axis) and log(-log)-scale (y-axis). The KM-estimate used by MINITAB is close to the estimator on the previous slide. MINTAB also computes 95% confidence bounds for the line.



Minitab here plots exactly the same points as for the Weibull plot, but since $\alpha = 1$ is known, the line to compare with is $y = x - \ln\hat{\theta}$, where $\hat{\theta}$ is the estimate of θ obtained from the exponential model.

The 95% confidence bounds are also different, since α is known.