TMA4275 LIFETIME ANALYSIS

Slides 3: Parametric families of lifetime distributions

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

MEAN TIME TO FAILURE (MTTF); EXPECTED LIFETIME

For a lifetime T we define

$$MTTF = E(T) = \int_0^\infty tf(t)dt = \int_0^\infty R(t)dt$$

(The last equality is proven by partial integration, noting that $R'(t) = -f(t)$. Do it!)

$$\begin{aligned} Var(T) &= \int_0^\infty (t - E(T))^2 f(t)dt \\ &= \int_0^\infty t^2 f(t)dt - (E(T))^2 \\ &= E(T^2) - (E(T))^2 \end{aligned}$$

$$\mathrm{SD}(T) = (Var(t))^{1/2}$$

EXAMPLE: EXPONENTIAL DISTRIBUTION

Let T be exponentially distributed with density $f(t) = \lambda e^{-\lambda t}$. Then you may check the following computations:

$$\begin{aligned} E(T) &= \int_0^\infty t\lambda e^{-\lambda t} dt = \int_0^\infty e^{-\lambda t} dt = \frac{1}{\lambda} \\ \text{Var}(T) &= E(T^2) - (E(T))^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2} \\ \text{SD}(T) &= \frac{1}{\lambda} \end{aligned}$$

Thus: For a component with exponentially distributed lifetime,

$$\text{MTTF} = 1/\text{failure rate}$$

NOTE: We will mainly use the parameterization $f(t) = \frac{1}{\theta} e^{-t/\theta}$, so that

$$R(t) = e^{-t/\theta}, \;\; E(T) = \theta, \;\; \text{SD} = \theta$$

The lifetime T is Weibull-distributed with *shape* parameter $\alpha > 0$ and *scale* parameter $\theta > 0$, written $T \sim \text{Weib}(\alpha, \beta)$, if

$$R(t) = e^{-(\frac{t}{\theta})^{\alpha}}$$

From this we can derive:

$$\begin{aligned} Z(t) &= \left(\frac{t}{\theta}\right)^{\alpha} \\ z(t) &= \frac{\alpha}{\theta}\left(\frac{t}{\theta}\right)^{\alpha-1} \\ f(t) &= z(t)e^{-Z(t)} = \frac{\alpha}{\theta}\left(\frac{t}{\theta}\right)^{\alpha-1} e^{-(\frac{t}{\theta})^{\alpha}} \end{aligned}$$

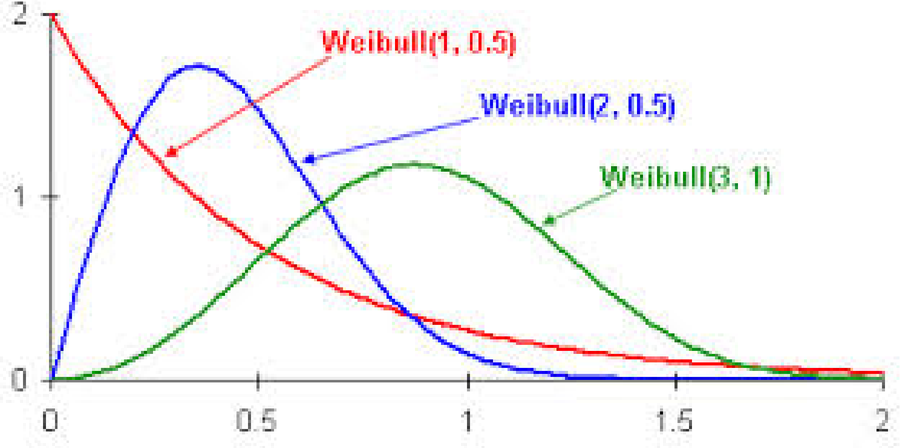
$\alpha = 1$ *corresponds to the exponential distribution;*
$\alpha < 1$ *gives a decreasing failure rate (DFR);*
$\alpha > 1$ *gives an increasing failure rate (IFR).*

$$E(T) = \int_0^\infty R(t)dt = \int_0^\infty e^{-(\frac{t}{\theta})^\alpha} dt = \cdots = \theta \cdot \Gamma\left(\frac{1}{\alpha} + 1\right)$$

where $\Gamma(\cdot)$ is the gamma-function defined by $\Gamma(a) = \int_0^\infty u^{a-1} e^{-u} du$.

$$\begin{aligned}
\text{Var}(T) &= \theta^2 \left(\Gamma\left(\frac{2}{\alpha}+1\right) - \Gamma^2\left(\frac{1}{\alpha}+1\right)\right) \\
SD(T) &= \theta \left(\Gamma\left(\frac{2}{\alpha}+1\right) - \Gamma^2\left(\frac{1}{\alpha}+1\right)\right)^{1/2}
\end{aligned}$$

WEIBULL DISTRIBUTION (CONT.)



Standard normal distribution, $Z \sim N(0,1)$:

$$
\begin{aligned}
f_Z(z) &= \phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \\
F_Z(z) &= \Phi(z) = \int_{-\infty}^{z} \phi(w)dw
\end{aligned}
$$

Now Let $Y \sim N(\mu, \sigma)$. Then it is well known that

$$
\begin{aligned}
F_Y(y) &= P(Y \leq y) = \Phi\left(\frac{y-\mu}{\sigma}\right) \\
M_Y(t) &= E(e^{tY}) = e^{\mu t + \frac{1}{2}\sigma^2 t^2} \quad \textit{(moment generating function)}
\end{aligned}
$$

Further, if we let $Z = \frac{Y-\mu}{\sigma} \sim N(0,1)$, then $Y = \mu + \sigma Z$.

Thus: The model $Y \sim N(\mu, \sigma)$ is a **location–scale family**, defined by the standardized random variable Z, with *location parameter* μ and *scale parameter* σ.

1. Consider $Y = \mu + \sigma Z$ where $Z \sim N(0,1)$. What is the distribution of Y? Why are the names *location* parameter and *scale* parameter appropriate for, respectively, μ and σ?
2. The normal distribution is sometimes used as a lifetime distribution (in fact it is a possible choice in MINITAB). What is a possible problem with this distribution?

The lifetime T has a lognormal distribution with parameters μ and σ if $Y \equiv \ln T$ is normally distributed, $Y \sim N(\mu, \sigma)$.

We can hence write

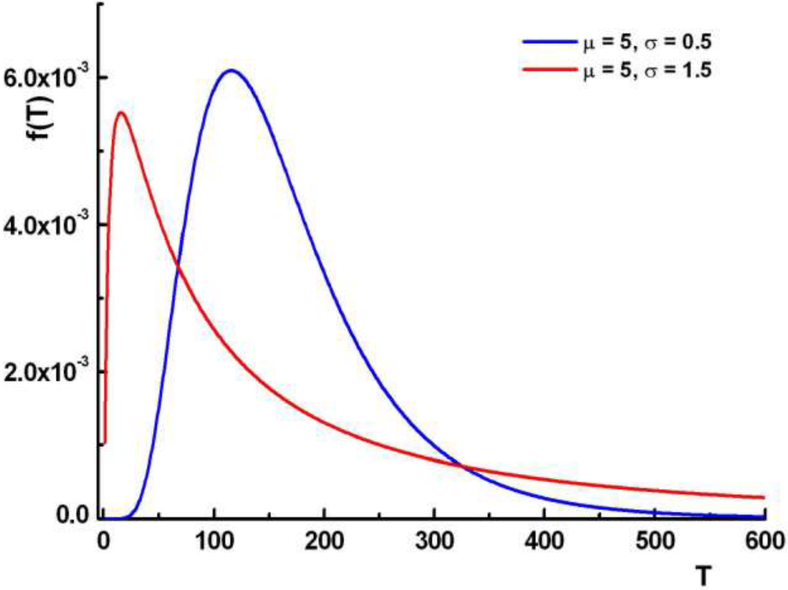
$$Y = \ln \mathrm{T} = \mu + \sigma Z \qquad (*)$$

where $Z \sim N(0, 1)$.

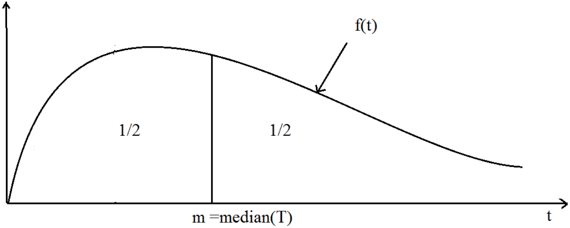
Here μ is called the *location parameter* and σ is called the *scale parameter* of the lognormal distribution.

Because of (*) we say that the lognormal distribution is a **log-location-scale family** of distributions, meaning that the log of T defines a *location-scale* family.

LOGNORMAL DISTRIBUTION (CONT.)



EXERCISE



$m = median(T)$ is defined by $F(m) = R(m) = 1/2$.

Compute the median m when T is

1. Exponentially distributed with parameter θ, i.e. $T \sim \text{Expon}(\theta)$
2. $T \sim \text{Weib}(\alpha, \theta)$
3. $T \sim \text{lognormal}(\mu, \sigma)$

Recall:

$$T \sim \text{lognormal}(\mu, \sigma) \iff \ln T \sim N(\mu, \sigma)$$

Thus

$$\begin{aligned} R(t) &= P(T > t) = P(\ln T > \ln t) \\ &= 1 - \Phi\left(\frac{\ln t - \mu}{\sigma}\right) \end{aligned}$$

and

$$\begin{aligned} f(t) &= -R'(t) = \phi\left(\frac{\ln t - \mu}{\sigma}\right) \cdot \frac{1}{t\sigma} \\ &= \frac{1}{\sqrt{2\pi}} \cdot \frac{1}{\sigma} \cdot \frac{1}{t} e^{-\frac{(\ln t - \mu)^2}{2\sigma^2}} \quad \text{for } t > 0 \end{aligned}$$

HAZARD FUNCTION OF THE LOGNORMAL DISTRIBUTION

$$z(t) = \frac{f(t)}{R(t)} = \frac{\phi\left(\frac{\ln t - \mu}{\sigma}\right)/(t\sigma)}{1 - \Phi\left(\frac{\ln t - \mu}{\sigma}\right)}$$

Let $T \sim \text{lognormal}(\mu, \sigma)$. Then $Y = \ln T \sim N(\mu, \sigma)$, and hence

$$M_Y(t) = E(e^{tY}) = e^{\mu t + \frac{1}{2}\sigma^2 t^2}$$

Thus:

$$\begin{aligned} E(T) &= E(e^Y) = M_Y(1) = e^{\mu + \frac{1}{2}\sigma^2} \\ E(T^2) &= E(e^{2Y}) = M_Y(2) = e^{2\mu + 2\sigma^2} \\ \text{Var}(T) &= e^{2\mu+2\sigma^2} - e^{2\mu+\sigma^2} = e^{2\mu+\sigma^2}(e^{\sigma^2} - 1) \end{aligned}$$

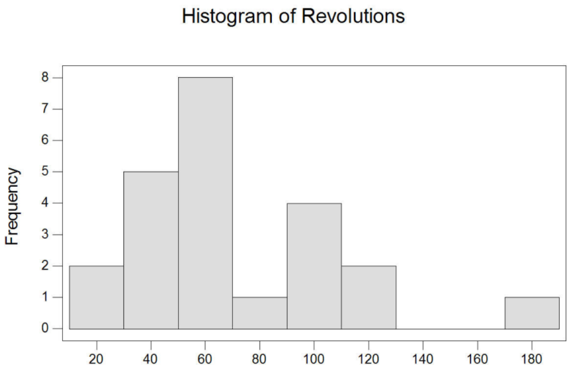
On the other hand,

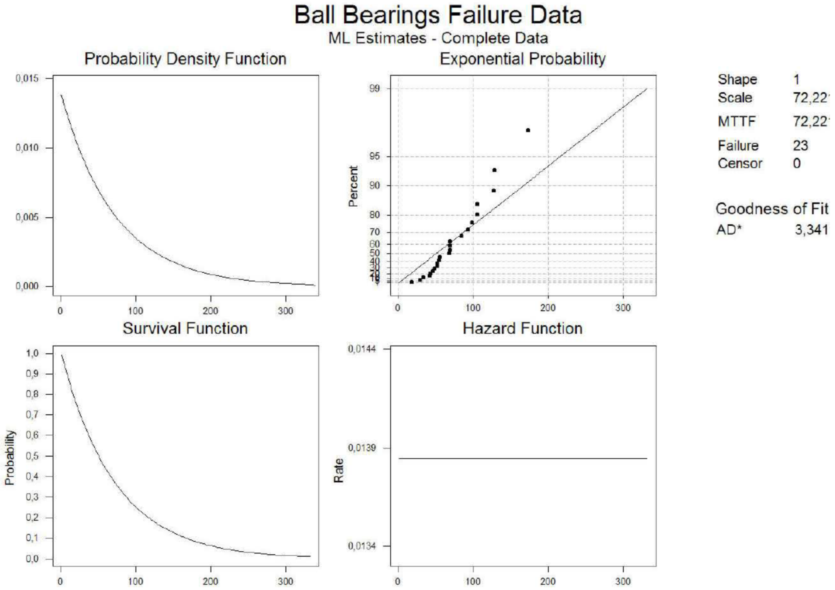
$$median(T) = e^{\mu}$$

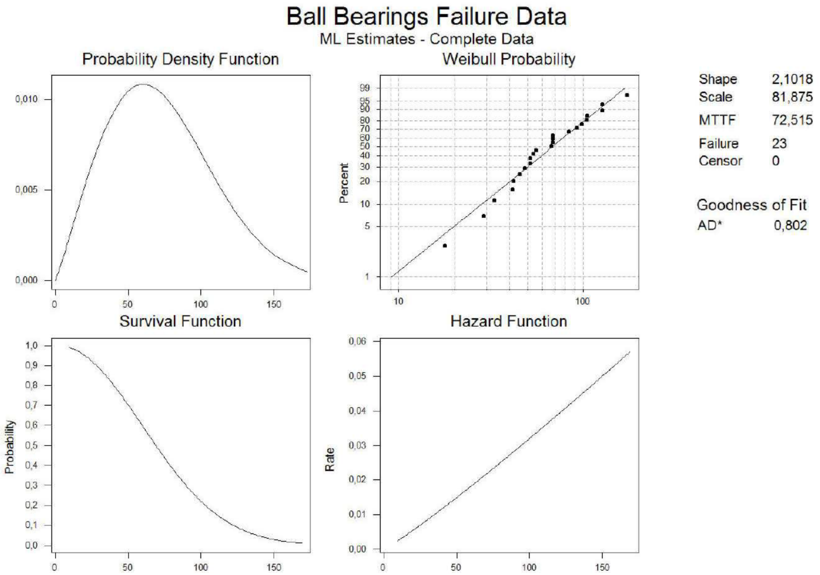
since $P(T \le e^{\mu}) = P(\ln T \le \mu) = P(Y \le \mu) = 1/2$.

17,88	28,92	33,00	41,52	42,12	45,60	48,40	51,84
51,96	54,12	55,56	67,80	68,64	68,64	68,88	84,12
93,12	98,64	105,12	105,84	127,92	128,04	173,40	

Question: *How can we fit a parametric lifetime model to these data?*

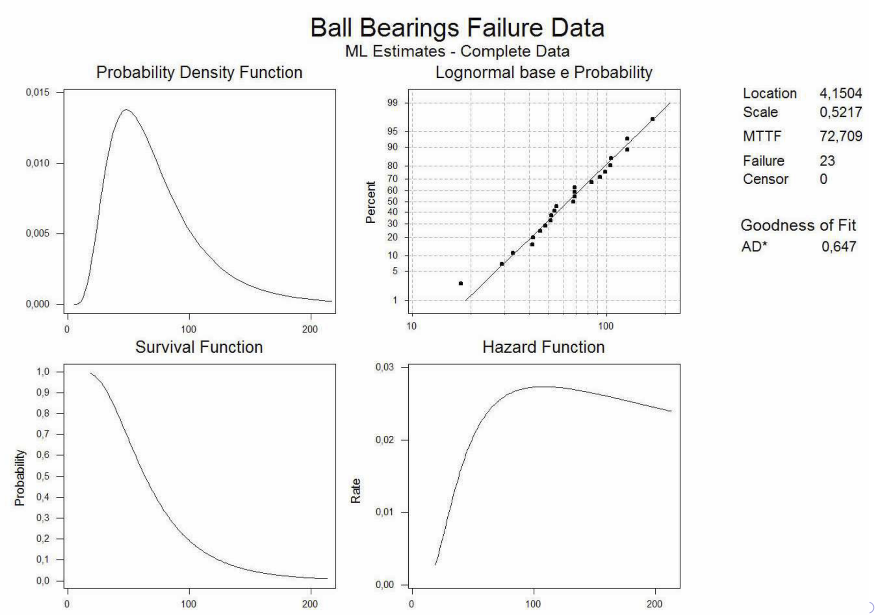


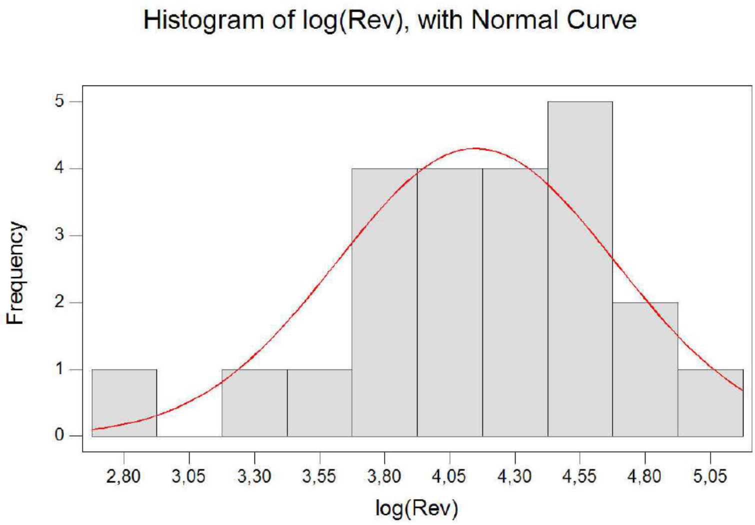

Ball Bearings Failure Data
ML Estimates - Complete Data
Probability Density Function
Exponential Probability
Survival Function
Hazard Function
Percent
Probability
Rate
Shape 1
Scale 72,22
MTTF 72,22
Failure 23
Censor 0
Goodness of Fit
AD* 3,341



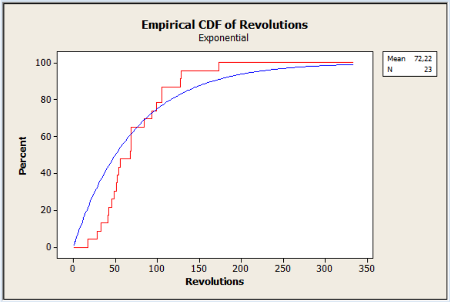
Ball Bearings Failure Data
ML Estimates - Complete Data
Probability Density Function
Weibull Probability
Survival Function
Hazard Function
Percent
Probability
Rate
Shape 2,1018
Scale 81,875
MTTF 72,515
Failure 23
Censor 0
Goodness of Fit
AD* 0,802

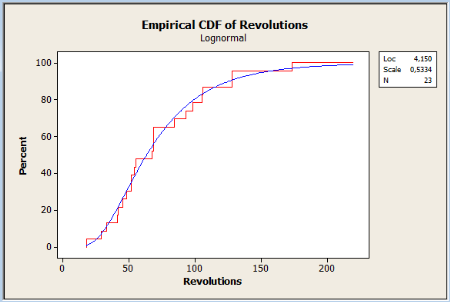
BB-DATA: LOGNORMAL DISTRIBUTION (MINITAB)




Histogram of log(Rev), with Normal Curve
Frequency
0
1
2
3
4
5
2,80
3,05
3,30
3,55
3,80
4,05
4,30
4,55
4,80
5,05
log(Rev)

BB-DATA: EMPIRICAL DISTRIBUTION COMPARED TO PARAMETRIC FITS





Model	$\widehat{\text{MTTF}}$	$\widehat{\text{STD}(T)}$	$\widehat{\text{med}(T)}$	$\hat{\alpha}$	$\hat{\theta}$	$\hat{\mu}$	$\hat{\sigma}$
Exp	72.221	72.221	50.060		72.221		
Weib	72.515	36.250	68.773	2.102	81.875		
Logn	72.710	40.664	63.458			4.150	0.522
Norm	72.221	36.667	72.221			72.221	36.667

Method: Maximum likelihood.

MOTIVATION FOR EXPONENTIAL DISTRIBUTION

- Simplest distribution used in the analysis of reliability data.
- Has the important characteristic that its hazard function is constant (does not depend on time t).
- Popular distribution for some kinds of electronic components (e.g., capacitors or robust, high-quality integrated circuits).
- Might be useful to describe failure times for components that exhibit physical wearout only after expected technological life of the system, in which the component would be replaced.

MOTIVATION FOR WEIBULL DISTRIBUTION

- The theory of extreme values shows that the Weibull distribution can be used to model the minimum of a large number of independent positive random variables from a certain class of distributions.
  - Failure of the weakest link in a chain with many links with failure mechanisms (e.g. fatigue) in each link acting approximately independently.
  - Failure of a system with a large number of components in series and with approximately independent failure mechanisms in each component.
- The more common justification for its use is empirical: the Weibull distribution can be used to model failure-time data with a decreasing or an increasing hazard function.

- The lognormal distribution is a common model for failure times.
- It can be justified for a random variable that arises from the product of a number of identically distributed independent positive random quantities (remember central limit theorem for sum of normals).
- It has been suggested as an appropriate model for failure time caused by a degradation process with combinations of random rates that combine multiplicatively.
- Widely used to describe time to fracture from fatigue crack growth in metals.

- Definition of
  - MTTF$= E(T)$
  - Var(T)
  - SD$(T) = \sqrt{\text{Var}\, T}$
- Presentation of distributions:
  - *Weibull-distribution*, Weib(α, θ)
  - Normal distribution, $N(\mu, \sigma)$
  - Lognormal distribution, lognormal(μ, σ)
- Definition of
  - *location-scale family*, e.g. $N(\mu, \sigma)$
  - *log-location-scale family*, e.g. lognormal(μ, σ)