Statistics for Engineers Lecture 4
Reliability and Lifetime Distributions

Chong Ma

Department of Statistics
University of South Carolina
chongm@email.sc.edu

February 15, 2017
Outline

1. Reliability Analysis
2. Weibull Distribution
3. Reliability Functions
4. QQ plot
Reliability analysis deals with failure time (i.e., lifetime, time-to-event) data. For example,

\[ T = \text{time from start of product service until failure} \]
\[ T = \text{time until a warranty claim} \]
\[ T = \text{number of hours in use/cycles until failure} \]

We call \( T \) a **lifetime random variable** if it measures the time to an “event”; e.g., failure, death, eradication of some infection/condition, etc. Engineers are often involved with reliability studies, because reliability is strongly related to product quality. There are many well-known **lifetime distribution**, including

- exponential
- Weibull
- Gamma, lognormal, inverse Gaussian, Gompertz-Makeham, Birnbaum-Sanders, Extreme value, log-logistic, etc.
Outline

1. Reliability Analysis
2. Weibull Distribution
3. Reliability Functions
4. QQ plot
A random variable $T$ is said to have a **Weibull distribution** with parameters $\beta > 0$ and $\eta > 0$ if its pdf is given by

$$f_T(t) = \frac{\beta}{\eta} \left( \frac{t}{\eta} \right)^{\beta-1} e^{-\left(\frac{t}{\eta}\right)^\beta} I(t > 0)$$

We say $T \sim \text{Weibull}(\beta, \eta)$, where

$\beta = \text{shape parameter}$

$\eta = \text{scale parameter}$

The cdf of $T \sim \text{Weibull}(\beta, \eta)$ exists in closed form such as

$$F_T(t) = (1 - e^{-\left(\frac{t}{\eta}\right)^\beta}) I(t > 0)$$

**Notation:** $I(t > 0) = \begin{cases} 1, & \text{if } t > 0 \\ 0, & \text{if } t \leq 0 \end{cases}$
Weibull Distribution

Given \( T \sim \text{Weibull}(\beta, \eta) \), the mean and variance of \( T \) are obtained by

\[
E(T) = \eta \Gamma \left( 1 + \frac{1}{\beta} \right) \\
\text{Var}(T) = \eta^2 \left\{ \Gamma \left( 1 + \frac{2}{\beta} \right) - \left[ \Gamma \left( 1 + \frac{1}{\beta} \right) \right]^2 \right\}
\]

Remarks

- By changing the values of \( \beta \) and \( \eta \), the Weibull distribution may assume many shapes. Because of this flexibility (and for other reasons), the Weibull distribution is very popular among engineers in reliability applications.

- When \( \beta = 1 \), the Weibull distribution reduces to the exponential \((\lambda = 1/\eta)\) distribution.
Figure 1: PDF and CDF for Weibull distributions with $\beta = 2, \eta = 5$, $\beta = 2, \eta = 10$ and $\beta = 3, \eta = 10$, respectively.
Weibull Distribution

Example The lifetime of a rechargeable battery under constant usage conditions, denoted by $T$ (measured in hours), follows a Weibull distribution with parameters $\beta = 2$ and $\eta = 10$.

(a) What is the mean time to failure?

$$E(T) = 10 \Gamma\left(\frac{3}{2}\right) \approx 8.862 \text{hours}$$

(b) What is the probability that a battery is still functional at time $t = 20$?

$$P(T \geq 20) = 1 - P(T < 20)$$
$$= 1 - F_T(20)$$
$$= 1 - (1 - e^{-(20/10)^2})$$
$$\approx 0.018$$
(c) What is the probability that a battery is still functional at time \( t = 20 \) given that the battery is functional at time \( t = 10 \)?

\[
P(T \geq 20 | T \geq 10) = \frac{P(T \geq 20 \text{ and } T \geq 10)}{P(T \geq 10)} = \frac{P(T \geq 20)}{P(T \geq 10)}
\]

\[
= \frac{1 - F_T(20)}{1 - F_T(10)} \approx e^{-\left(\frac{20}{10}\right)^2} \approx 0.050
\]

Note that

\[0.050 \approx P(T \geq 20 | T \geq 10) \neq P(T \geq 10) = e^{-\left(\frac{20}{10}\right)^2} \approx 0.368\]

(d) What is the 99th percentile of this lifetime distribution? Solve

\[
F_T(\phi_{0.99}) = 1 - e^{-\left(\phi_{0.99}/10\right)^2} = 0.99
\]

and we obtain \( \phi_{0.99} \approx 21.64 \) hours.
We now describe some different, but equivalent, ways of defining the distribution of a (continuous) lifetime random variable $T$.

- **cumulative distribution function (cdf)**
  
  $$F_T(t) = P(T \leq t)$$
  
  which can be interpreted as the proportion of units that have failed by time $t$.

- **survivor function**

  $$S_T(t) = P(T > t) = \bar{F}_T(t) = 1 - F_T(t)$$
  
  Which can be interpreted as the proportion of units that have **not** failed by time $t$.

- **probability density function**

  $$f_T(t) = \frac{d}{dt} F_T(t) = -\frac{d}{dt} S_T(t)$$
  
  Where $F_T(t) = \int_0^t f_T(t)dt$ and $S_T(t) = \int_t^\infty f_T(t)dt$. 
The **hazard function** of a lifetime random variable $T$ is

$$h_T(t) = \lim_{\epsilon \to 0} \frac{P(t \leq T \leq t + \epsilon | T \geq t)}{\epsilon} = \frac{f_T(t)}{S_T(t)}$$

The hazard function is not a probability; rather, it is a **probability rate**, indicating **how the rate of failure varies with time.** Interprets:

- **Distribution with increasing** hazard functions are seen in units where some kind of aging or “wear out” take place, that is, getting **weaker** over time.
- **Distribution with decreasing** hazard functions correspond to the population getting **stronger** over time.
- The hazard function may decrease initially, stay constant over a period of time, and then increase. This corresponds to a population whose units get stronger initially (defective individuals “die out” early), exhibit random failures for a period of time (constant hazard), and then eventually the population weakens (e.g., due to old age, etc). These hazard functions are **bathtub-shaped**.
Hazard Functions

Given that \( T \sim \text{Weibull}(\beta, \eta) \), then

- The pdf of \( T \) is

\[
f_T(t) = \frac{\beta}{\eta} \left( \frac{t}{\eta} \right)^{\beta-1} e^{-\left(\frac{t}{\eta}\right)^\beta} I(t > 0)
\]

- The cdf of \( T \) is

\[
F_T(t) = (1 - e^{-\left(\frac{t}{\eta}\right)^\beta}) I(t > 0)
\]

- The survivor function of \( T \) is

\[
S_T(t) = e^{-\left(\frac{t}{\eta}\right)^\beta} I(t > 0)
\]

- The hazard function of \( T \) is

\[
h_T(t) = \frac{f_T(t)}{S_T(t)} = \frac{\frac{\beta}{\eta} \left( \frac{t}{\eta} \right)^{\beta-1} e^{-\left(\frac{t}{\eta}\right)^\beta}}{e^{-\left(\frac{t}{\eta}\right)^\beta}} = \frac{\beta}{\eta} \left( \frac{t}{\eta} \right)^{\beta-1}
\]
Hazard Functions

Figure 2: Weibull hazard functions with $\eta = 1$ and $\beta = 3, 1.5, 1, \text{ and } 0.5$. 
Hazard Functions

Remarks It is easy to see that for a Weibull distribution

- $h_T(t)$ is **increasing** if $\beta > 1$ (wear out; population gets weaker)
- $h_T(t)$ is **constant** if $\beta = 1$ (random failures; exponential distribution)
- $h_T(t)$ is **decreasing** if $\beta < 1$ (infant mortality; population gets stronger)

Example The data below are times, denoted by $T$ (measured in months), to the first failure for 20 electric carts used for internal delivery and transportation in a large manufacturing facility.

<table>
<thead>
<tr>
<th>3.9</th>
<th>4.2</th>
<th>5.4</th>
<th>6.5</th>
<th>7.0</th>
<th>8.8</th>
<th>9.2</th>
<th>11.4</th>
<th>14.3</th>
<th>15.1</th>
</tr>
</thead>
<tbody>
<tr>
<td>15.3</td>
<td>15.5</td>
<td>17.9</td>
<td>18.0</td>
<td>19.0</td>
<td>19.0</td>
<td>23.9</td>
<td>24.8</td>
<td>26.0</td>
<td>34.2</td>
</tr>
</tbody>
</table>

In this example, we will assume that a $\textit{Weibull}(\beta, \eta)$ model for $T = \text{time to cart failure}$ (in months)

Because the model parameters $\beta$ and $\eta$ are not given to us, our first task is to estimate them using the data at hand.
To estimate the parameters $\beta$ and $\eta$, we form the **likelihood function**

$$L(\beta, \eta) = \prod_{i=1}^{20} f_T(t_i) = \prod_{i=1}^{20} \frac{\beta}{\eta} \left(\frac{t_i}{\eta}\right)^{\beta-1} e^{-\left(\frac{t_i}{\eta}\right)^\beta}$$

$$= \left(\frac{\beta}{\eta^\beta}\right)^{20} \left(\prod_{i=1}^{20} t_i\right)^{\beta-1} e^{-\sum_{i=1}^{20} \left(\frac{t_i}{\eta}\right)^\beta}$$

To obtain the values of $\beta$ and $\eta$ that “most closely agree” with the data, we need maximize $L(\beta, \eta)$ with respect to $\beta$ and $\eta$, that is,

$$\max_{\beta, \eta} L(\beta, \eta)$$

The maximizer $\hat{\beta}$ and $\hat{\eta}$, that is, maximizing the likelihood function $L(\beta, \eta)$, are the **maximum likelihood estimates** of $\beta$ and $\eta$. 
For the cart data, the maximum likelihood estimates of $\beta$ and $\eta$ are

$$\hat{\beta} \approx 1.99, \hat{\eta} \approx 16.94$$

The $\hat{\beta} \approx 1.99$ estimate suggests that there is “wear out” taking place among the carts, that is, the population of carts gets weaker as time passes.

(a) Using the estimated Weibull ($\beta \approx 1.99, \eta \approx 16.94$) distribution as a model for future cart lifetimes, find the probability that a cart will “survive” past 20 months.

$$P(T > 20) = 1 - P(T \leq 20) = 1 - F_T(20)$$

$$= 1 - (1 - e^{-(20/16.94)^{1.99}}) \approx 0.249$$

(b) Using the estimated distribution, find the 90th percentile of the cart lifetimes.

$$F_T(\phi_{0.90}) = 1 - e^{-(\phi_{0.90}/16.94)^{1.99}} = 0.90$$

Solving for $\phi_{0.90}$ gives $\phi_{0.90} \approx 25.75$. Only ten percent of the cart lifetime would exceed this value.
Outline

1. Reliability Analysis
2. Weibull Distribution
3. Reliability Functions
4. QQ plot
A quantile-quantile plot (QQ plot) is a graphical display that can help assess the appropriateness of a model (distribution). Here is how the plot is constructed:

- On the vertical axis, we plot the observed data, ordered from low to high.
- On the horizontal axis, we plot the (ordered) theoretical quantiles from the distribution (model) assumed for the observed data.

Our intuition should suggest the following:

- If the observed data “agree” with the distribution’s theoretical quantiles, then the QQ plot should look like a straight line (the distribution is a good choice).
- If the observed data do not “agree” with the theoretical quantiles, then the QQ plot should have curvature in it (the distribution is not a good choice).
Figure 3: Cart data - Weibull qq plot. The observed data are plotted versus the theoretical quantiles from a Weibull distribution with $\hat{\beta} \approx 1.99$ and $\hat{\eta} \approx 16.94$. 
Quantile-quantile plots

Remarks: When it comes to interpreting QQ plot, we look for **general agreement**. The observed data usually never line up perfectly with the model’s quantiles.

**Interpretation for the cart data - Weibull distribution**

- There is a **general agreement** with the cart data and the quantiles from the Weibull distribution.

- The straight line is formed from the 25th percentile and 75th percentiles of the observed data and the assumed model.

- The bands composed by red dashed lines can be used to determine to what extent the skewness of the scatter plot is allowed to not suspect the model.
  - If all of the data fall within the bands, then there is no reason to suspect the model.
  - detect outliers, that is, observations not grossly consistent with the assumed model.