Unit 2: Models, Censoring, and Likelihood for Failure-Time Data


Ramón V. León

Unit 2 Objectives

- Describe models for continuous failure-time processes/distributions
  - Time to failure
- Describe models that we will use for the discrete data from these continuous failure-time processes/distributions
  - Data resulting primarily by interval censoring, i.e. failure times that one knows that they fell in an interval
    - Rounding
Unit 2 Objectives

- Describe common censoring mechanisms that restrict our ability to observe all of the failure times that might occur in a reliability study
  - Right censoring: survival times, i.e., one only knows that a failure would have occurred past some survival (censoring) time.
  - Left censoring: one only knows that the failure occurred before some (censoring) time
  - Interval censoring: one only knows that the failure occurred in an interval and not the exact time to failure
Typical Failure-Time Probability Functions

Cumulative Distribution Function
\[ F(t) = P(T \leq t) = 1 - \exp(-t^{1.7}) \]

Probability Density Function
\[ f(t) = \frac{dF(t)}{dt} = 1.7 \times t^{1.7} \times \exp(-t^{1.7}) \]

Survival Function
\[ S(t) = P(T \geq t) = \int_{t}^{\infty} f(t) \, dt = \exp(-t^{1.7}) \]

Hazard Function
\[ h(t) = \frac{f(t)}{S(t)} = \frac{f(t)}{1 - F(t)} = 1.7 \times t^{7} \]
Hazard Function or Instantaneous Failure Rate Function

The hazard function $h(t)$ is defined by

$$h(t) = \lim_{\Delta t \to 0} \frac{P(t < T \leq t + \Delta t | T > t)}{\Delta t}$$

$$= \lim_{\Delta t \to 0} \frac{P(t < T \leq t + \Delta t) / P(T > t)}{\Delta t}$$

$$= \lim_{\Delta t \to 0} \frac{P(t < T \leq t + \Delta t)}{P(T > t) \Delta t}$$

$$= \frac{f(t)}{S(t)} = \frac{f(t)}{1 - F(t)}$$

CDF in Terms of the Hazard Function

$$H(t) = \int_0^t h(x) dx = \int_0^t \frac{f(x)}{1 - F(x)} dx$$

Let $u = F(x) \Rightarrow du = f(x) dx$, so

$$H(t) = \int_0^{F(t)} \frac{du}{1 - u} = -\ln(1 - u)_{0}^{F(t)} = -\ln(1 - F(t)) = -\ln S(t)$$

$$\Rightarrow$$

$$F(t) = 1 - \exp(-H(t)) = 1 - \exp\left(-\int_0^t h(x) dx\right)$$
Hazard Function or Instantaneous Failure Rate Function

- $h(t)$ describes propensity of failure in the next small interval of time given survival to time $t$
  \[ h(t) \times \Delta t \approx \Pr(t < T \leq t + \Delta t \mid T > t). \]
- Some reliability engineers think of modeling in terms of $h(t)$.

Engineers Interpretation of the Hazard Function

The hazard function can be interpreted as a failure rate if there is a large number of items, say $n(t)$, in operation at time $t$. Then

\[ n(t) \times h(t) \Delta t = \text{Expected number of failures in time (} t, t + \Delta t) \]

\[ \Rightarrow \]

\[ h(t) = \text{Expected number of failures per unit of time per unit at risk} \]
FIT Rate

A FIT rate is defined as the hazard function in units of 1/hours multiplied by $10^9$.

Example 2.5

- 165,000 copies of a component
- Hazard rate constant over time at 15 FITs.
  - $h(t) = 15 \times 10^{-9}$ failures per unit per hour for all times $t$ measured in units of hours.
- A prediction for the number of failures from this component in 1 year (8760 hours) of operation is
  - $15 \times 10^{-9} \times 165,000 \times 8760 = 21.7$
Bathtub Curve Hazard Function

Cumulative Hazard Function and Average Hazard Rate

- Cumulative hazard function:
  \[ H(t) = \int_0^t h(x) \, dx. \]
  Notice that, \( F(t) = 1 - \exp \left[ -H(t) \right] = 1 - \exp \left[ - \int_0^t h(x) \, dx \right]. \)

- Average hazard rate in interval \((t_1, t_2)\):
  \[ \text{AHR}(t_1, t_2) = \frac{\int_{t_1}^{t_2} h(u) \, du}{t_2 - t_1} = \frac{H(t_2) - H(t_1)}{t_2 - t_1}. \]
Practical Interpretation of Average Hazard Rate

\[ AHR(t_1, t_2) = \frac{F(t_2) - F(t_1)}{t_2 - t_1} = \frac{P(t_1 \leq T \leq t_2)}{t_2 - t_1} \]

if \( F(t_2) = P(T \leq t_2) \) is small, say less than 0.1

In particular,

\[ AHR(t) = \int_0^t \frac{h(u)du}{t} = \frac{H(t)}{t} \approx \frac{F(t)}{t} = \frac{P(T \leq t)}{t} \]

if \( F(t) = P(T \leq t) \) is small, say less than 0.1

Derivation

\[ \int_0^t f(u)du \leq \int_0^t f(u)du \leq \int_0^t f(u)du \]

\[ \Rightarrow \]

\[ \frac{F(t_2) - F(t_1)}{S(t_2)} \leq \int_0^t h(u)du = H(t_2) - H(t_1) \leq \frac{F(t_2) - F(t_1)}{S(t_2)} \]

\[ \Rightarrow \]

\[ \frac{1}{S(t_2)} \left[ \frac{F(t_2) - F(t_1)}{t_2 - t_1} \right] \leq \frac{H(t_2) - H(t_1)}{t_2 - t_1} = AHR(t_1, t_2) \leq \frac{1}{S(t_2)} \left[ \frac{F(t_2) - F(t_1)}{t_2 - t_1} \right] \]

So if \( (F(t_1) \leq F(t_2)) \) is small \( S(t_2) \leq S(t_1) \) is close to 1

\[ \Rightarrow \]

\[ AHR(t_1, t_2) = \left[ \frac{F(t_2) - F(t_1)}{t_2 - t_1} \right] = \frac{P(t_1 \leq T \leq t_2)}{t_2 - t_1} \]
Quantile Function

The $p$ quantile of $F$ is the smallest time $t_p$ such that

$$P(T \leq t_p) = F(t_p) \geq p,$$

where $0 < p < 1$

When $F(t)$ is constant over some intervals, there can be more than one solution $t$ to the equation $F(t) \geq p$. Taking $t_p$ equal to the smallest $t$ value satisfying $F(t) \geq p$ is a standard convention.

Distribution Quantiles
Simple Quantile Calculation

When \( F(t) \) is strictly increasing there is a unique value \( t_p \) that satisfies \( F(t_p) = p \), and we write
\[
t_p = F^{-1}(p).
\]

Example:
\( t_{20} \) is the time by which 20\% of the population will fail. For,
\[
F(t) = 1 - \exp(-t^{1.7}), \quad p = F(t_p) \text{ gives } t_p = \left[-\log(1-0.2)\right]^{1/1.7}
\]
and \( t_{2.2} = \left[-\log(1-0.2)\right]^{1/1.7} = 0.414. \)

Terminology:

\( t_{10} \) is also known as B10)

Models for Discrete Data from a Continuous Time Process

All data are discrete! Partition \((0, \infty)\) into \( m + 1 \) intervals depending on inspection times and roundoff as follows:
\[
(t_0, t_1), (t_1, t_2), \ldots, (t_{m-1}, t_m), (t_m, t_{m+1})
\]
where \( t_0 = 0 \) and \( t_{m+1} = \infty \). Observe that the last interval is of infinite length.
Partitioning of Time into Non-Overlapping Intervals

\[ \pi_1 \quad \pi_2 \quad \pi_3 \ldots \pi_{m-1} \quad \pi_m \quad \pi_{m+1} \]

\[ t_0 = 0 \quad t_1 \quad t_2 \ldots \quad t_{m-1} \quad t_m \quad t_{m+1} = \infty \]

Times need **not** be equally spaced.

The \( \pi \)'s are the probabilities of failure in the intervals.

Graphical Interpretations of the \( \pi \)'s
Nonparametric Parameters

Define,

\[ \pi_i = \Pr(t_{i-1} < T \leq t_i) = F(t_i) - F(t_{i-1}) \]
\[ p_i = \Pr(t_{i-1} < T \leq t_i | T > t_{i-1}) = \frac{F(t_i) - F(t_{i-1})}{1 - F(t_{i-1})} = \frac{\pi_i}{S(t_{i-1})} \]

Because the \( \pi_i \) values are multinomial probabilities, \( \pi_i \geq 0 \) and \( \sum_{j=1}^{m+1} \pi_j = 1 \). Also, \( p_{m+1} = 1 \) but the only restriction on \( p_1, \ldots, p_m \) is \( 0 \leq p_i \leq 1 \)

Notice:

\[ S(t_{i-1}) = P(T > t_{i-1}) = \sum_{j=1}^{m+1} \pi_j \]
\[ \pi_i = p_i S(t_{i-1}) \]

A Important Derivation

\[ S(t_{i-1}) - S(t_i) = F(t_i) - F(t_{i-1}) = \pi_i = p_i S(t_{i-1}) \]
\[ \Rightarrow \]
\[ (1 - p_i)S(t_{i-1}) = S(t_i) \]
\[ \Rightarrow \text{by induction} \]
\[ S(t_i) = \prod_{j=1}^{i} (1 - p_j), \quad i = 1, \ldots, m + 1 \]
Nonparametric Parameters

Since

\[ F(t_i) = 1 - \prod_{j=1}^{i} (1 - p_j), \quad i = 1, \ldots, m + 1 \]

and

\[ F(t_i) = \sum_{j=1}^{i} \pi_j, \quad i = 1, \ldots, m + 1 \]

we view \( \pi = (\pi_1, \ldots, \pi_{m+1}) \) or \( p = (p_1, \ldots, p_{m+1}) \) as the nonparametric parameters.

Example Calculation of the Nonparametric Parameters

Probabilities for the Multinomial Failure Time Model

Computed from \( F(t) = 1 - \exp(-t^{1.7}) \)

<table>
<thead>
<tr>
<th>( t_i )</th>
<th>( F(t_i) )</th>
<th>( S(t_i) )</th>
<th>( \pi_i )</th>
<th>( p_i )</th>
<th>( 1 - p_i )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>.000</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.5</td>
<td>.265</td>
<td>.735</td>
<td>.265</td>
<td>.265</td>
<td>.735</td>
</tr>
<tr>
<td>1.0</td>
<td>.632</td>
<td>.368</td>
<td>.367</td>
<td>.500</td>
<td>.500</td>
</tr>
<tr>
<td>1.5</td>
<td>.864</td>
<td>.136</td>
<td>.231</td>
<td>.629</td>
<td>.371</td>
</tr>
<tr>
<td>2.0</td>
<td>.961</td>
<td>.0388</td>
<td>.0976</td>
<td>.715</td>
<td>.285</td>
</tr>
<tr>
<td>( \infty )</td>
<td>1.000</td>
<td>.000</td>
<td>.0388</td>
<td>1.000</td>
<td>.000</td>
</tr>
</tbody>
</table>

\( p_i = \frac{\pi_i}{S(t_{i-1})} \)
Examples of Censoring Mechanisms

Censoring restricts our ability to observe $T$. Some sources of censoring are:

- Fixed time to end test (lower bound on $T$ for unfailed units)
- Inspection times (upper and lower bounds on $T$)
- Staggered entry of units into service leads to multiple censoring
- Multiple failure modes (also known as competing risks, and resulting in multiple right censoring):
  - Independent failure modes (simple)
  - Non-independent failure modes (difficult)

Simple analysis requires non-informative censoring assumptions

Likelihood (Probability of the Data) as a Unifying Concept

- Likelihood provides a general and versatile method of estimation
- Model/Parameters combinations with large likelihood are plausible
- Allows for censored, interval, and truncated data
- Theory is simple in regular models
- Theory more complicated in non-regular models (but concepts are similar)
- Limitation: can be computationally intensive (still not general software)
Determining the Likelihood (Probability of the Data)

The form of the likelihood will depend on

- Question and focus of the study
- Assumed model
- Measurement system (form of available data)
- Identifiability/parametrization

Likelihood Contributions for Different Kinds of Censoring
Likelihood Contributions for Different Kinds of Censoring

Example: $F(t) = 1 - \exp(-t^{1.7})$

- Interval-censored observations:
  $$L_i(p) = \int_{t_{i-1}}^{t_i} f(t) \, dt = F(t_i) - F(t_{i-1}).$$
  If a unit is still operating at $t = 1.0$ but has failed at $t = 1.5$ inspection, $L_i = F(1.5) - F(1.0) = .231$.

- Left-censored observations:
  $$L_i(p) = \int_0^{t_i} f(t) \, dt = F(t_i) - F(0) = F(t_i).$$
  If a failure is found at the first inspection time $t = .5$, $L_i = F(.5) = .265$.

- Right-censored observations:
  $$L_i(p) = \int_{t_i}^{\infty} f(t) \, dt = F(\infty) - F(t_i) = 1 - F(t_i).$$
  If a unit has not failed by the last inspection at $t = 2$, $L_i = 1 - F(2) = .0388$. 
Likelihood for Life Data

\[ d_i = 2 = \text{number of observations interval censored in } t_{i-1} \text{ and } t_i \]

\[ l_i = 3 = \text{number of observations left-censored at } t_i \]

\[ r_i = 2 = \text{number of observations right-censored at } t_i \]

Likelihood for Life Table Data

- For a life table the data are: the number of failures \((d_i)\), right censored \((r_i)\), and left censored \((l_i)\) units on each of the nonoverlapping interval \((t_{i-1}, t_i]\), \(i = 1, \ldots, m+1, t_0 = 0\).

- The likelihood (probability of the data) for a single observation, \(d_{i}, \text{ in } (t_{i-1}, t_i]\) is

\[
L_i(\pi; \text{data}_i) = F(t_i; \pi) - F(t_{i-1}; \pi).
\]

- Assuming that the censoring is at \(t_i\)

<table>
<thead>
<tr>
<th>Type of Censoring</th>
<th>Characteristic</th>
<th>Number of Cases</th>
<th>Likelihood of Responses (L_i(\pi; \text{data}_i))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Left at (t_i)</td>
<td>(T \leq t_i)</td>
<td>(l_i)</td>
<td>([F(t_i)]^{l_i})</td>
</tr>
<tr>
<td>Interval (t_{i-1} &lt; T \leq t_i)</td>
<td>(d_i)</td>
<td>([F(t_i) - F(t_{i-1})]^{d_i})</td>
<td></td>
</tr>
<tr>
<td>Right at (t_i)</td>
<td>(T &gt; t_i)</td>
<td>(r_i)</td>
<td>([1 - F(t_i)]^{r_i})</td>
</tr>
</tbody>
</table>
Likelihood: Probability of the Data

- The total likelihood, or joint probability of the DATA, for \( n \) independent observations is

\[
L(\pi; \text{DATA}) = c \prod_{i=1}^{n} L_i(\pi; \text{data}_i)
\]

\[
= c \prod_{i=1}^{m+1} [F(t_i)]^{d_i} [F(t_i) - F(t_{i-1})]^{r_i} [1 - F(t_i)]^{e_i}
\]

where \( n = \sum_{j=1}^{m+1} (d_j + r_j + e_j) \) and \( c \) is a constant depending on the sampling inspection scheme but not on \( \pi \). So we can take \( c = 1 \).

- Want to find \( \pi \) so that \( L(\pi) \) is large.

Likelihood for Arbitrary Censored Data

- In general, the \( i \)th observation consists of an interval \((t_{l,i}, t_i]\), \( i = 1, \ldots, n \) \((t_{l,i} < t_i)\) that contains the time event \( T \) for the \( i \)th individual.

The intervals \((t_{l,i}, t_i]\) may overlap and their union may not cover the entire time line \((0, \infty)\). In general \( t_{l,i} \neq t_{i-1} \).

- Assuming that the censoring is at \( t_i \)

<table>
<thead>
<tr>
<th>Type of Censoring</th>
<th>Characteristic</th>
<th>Likelihood of a single Response ( L_i(\pi; \text{data}_i) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Left at ( t_i )</td>
<td>( T \leq t_i )</td>
<td>( F(t_i) )</td>
</tr>
<tr>
<td>Interval</td>
<td>( t_{l,i} &lt; T \leq t_i )</td>
<td>( F(t_i) - F(t_{l,i}) )</td>
</tr>
<tr>
<td>Right at ( t_i )</td>
<td>( T &gt; t_i )</td>
<td>( 1 - F(t_i) )</td>
</tr>
</tbody>
</table>
Likelihood for General Reliability Data

• The total likelihood for the DATA with \( n \) independent observations is

\[ L(\pi; \text{DATA}) = \prod_{i=1}^{n} L_i(\pi; \text{data}_i). \]

• Some of the observations may have multiple occurrences. Let \( (t_{ij}^j, t_{ij}), j = 1, \ldots, k \) be the distinct intervals in the DATA and let \( w_j \) be the frequency of observation of \( (t_{ij}^j, t_{ij}) \). Then

\[ L(\pi; \text{DATA}) = \prod_{j=1}^{k} \left[ L_j(\pi; \text{data}_j) \right]^{w_j}. \]

• In this case the nonparametric parameters \( \pi \) correspond to probabilities of a partition of \((0, \infty)\) determined by the data (Examples later).

Other Topics in Chapter 2

• Random censoring
• Overlapping censoring intervals
• Likelihood with censoring in the intervals
• How to determine C