Chapter 2

Models, Censoring, and Likelihood for Failure-Time Data

William Q. Meeker and Luis A. Escobar
Iowa State University and Louisiana State University

Based on the authors' text *Statistical Methods for Reliability Data,* John Wiley & Sons Inc. 1998.

January 5, 2006
19h 13min

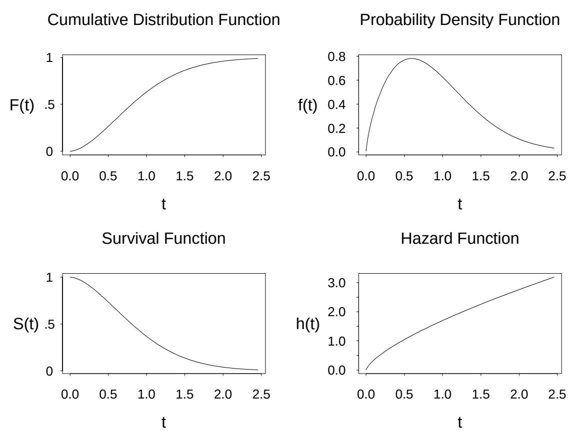
Chapter 2
Models, Censoring, and Likelihood for Failure-Time Data
Objectives

- Describe models for continuous failure-time processes.
- Describe some reliability metrics.
- Describe models that we will use for the discrete data from these continuous failure-time processes.
- Describe common censoring mechanisms that restrict our ability to observe all of the failure times that might occur in a reliability study.
- Explain the principles of likelihood, how it is related to the probability of the observed data, and how likelihood ideas can be used to make inferences from reliability data.

Typical Failure-time cdf, pdf, hf, and sf

$F(t) = 1 - \exp(-t^{1.7})$; $\quad f(t) = 1.7 \times t^{.7} \times \exp(-t^{1.7})$

$S(t) = \exp(-t^{1.7})$; $\quad h(t) = 1.7 \times t^{.7}$



Models for Continuous Failure-Time Processes

T is a nonnegative, continuous random variable describing the failure-time process. The distribution of T can be characterized by any of the following functions:

- The cumulative distribution function (cdf): $F(t) = \Pr(T \leq t)$.
  Example, $F(t) = 1 - \exp(-t^{1.7})$.
- The probability density function (pdf): $f(t) = dF(t)/dt$.
  Example, $f(t) = 1.7 \times t^{.7} \times \exp(-t^{1.7})$.
- Survival function (or reliability function):
  $$S(t) = \Pr(T > t) = 1 - F(t) = \int_t^{\infty} f(x)dx.$$
  Example, $S(t) = \exp(-t^{1.7})$.
- The hazard function: $h(t) = f(t)/[1 - F(t)]$.
  Example, $h(t) = 1.7 \times t^{.7}$

Hazard Function
or Instantaneous Failure Rate Function

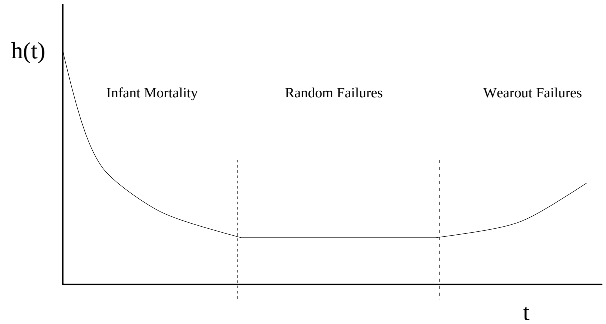
The hazard function $h(t)$ is defined by

$$h(t) = \lim_{\Delta t \to 0} \frac{\Pr(t < T \leq t + \Delta t \mid T > t)}{\Delta t} = \frac{f(t)}{1 - F(t)}.$$

Notes:

- $F(t) = 1 - \exp[-\int_0^t h(x)dx]$, etc.
- $h(t)$ 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 $h(t)$.

Bathtub Curve Hazard Function



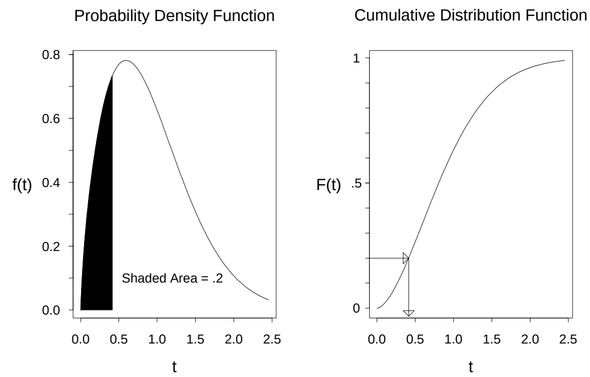
Cumulative Hazard Function and Average Hazard Rate

- 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 $(t_1, t_2]$:
$$\text{AHR}(t_1, t_2) = \frac{\int_{t_1}^{t_2} h(u)du}{t_2 - t_1} = \frac{H(t_2) - H(t_1)}{t_2 - t_1}.$$
If $F(t_2) - F(t_1)$ is small (say less than .1), then
$$\text{AHR}(t_1, t_2) \approx \frac{F(t_2) - F(t_1)}{(t_2 - t_1)\,S(t_1)}.$$

- An important special case arises when $t_1 = 0$,
$$\text{AHR}(t) = \frac{\int_0^t h(u)du}{t} = \frac{H(t)}{t} \approx \frac{F(t)}{t}.$$
Approximation is good for small $F(t)$, say $F(t) < .10$.

Plots Showing that the Quantile Function is the Inverse of the cdf



Distribution Quantiles

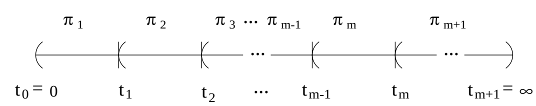
- The p quantile of F is the **smallest** time t_p such that
$$\Pr(T \le t_p) = F(t_p) \ge p, \quad \text{where } 0 < p < 1.$$
$t_{.20}$ is the time by which 20% of the population will fail. For, $F(t) = 1 - \exp(-t^{1.7})$, $p = F(t_p)$ gives $t_p = [-\log(1-p)]^{1/1.7}$ and $t_{.2} = [-\log(1 - .2)]^{1/1.7} = .414$.

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

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

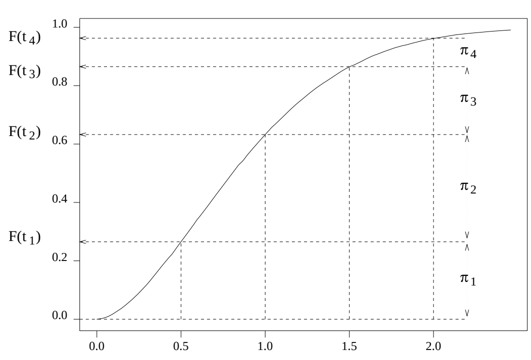
- t_p is also know ad B100p (e.g., $t_{.10}$ is also known as B10).

Partitioning of Time into Non-Overlapping Intervals



Times need **not** be equally spaced.

Graphical Interpretation of the π's



Models for Discrete Data from a Continuous Time Processes

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.

Define,
$$\begin{aligned} \pi_i &= \Pr(t_{i-1} < T \le t_i) = F(t_i) - F(t_{i-1}) \\ p_i &= \Pr(t_{i-1} < T \le t_i \mid T > t_{i-1}) = \frac{F(t_i) - F(t_{i-1})}{1 - F(t_{i-1})} \end{aligned}$$
Because the π_i values are multinomial probabilities, $\pi_i \ge 0$ 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 $0 \le p_i \le 1$

Models for Discrete Data from a Continuous Time Processes–Continued

It follows that,

$$S(t_{i-1}) = \Pr(T > t_{i-1}) = \sum_{j=i}^{m+1} \pi_j$$

$$\pi_i = p_i S(t_{i-1})$$

$$S(t_i) = \prod_{j=1}^{i} \left(1 - p_j\right), \quad i = 1, \ldots, m+1$$

We view $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_{m+1})$ or $\boldsymbol{p} = (p_1, \ldots, p_m)$ as the non-parametric parameters.

Probabilities for the Multinomial Failure Time Model Computed from $F(t) = 1 - \exp(-t^{1.7})$

t_i	$F(t_i)$	$S(t_i)$	π_i	p_i	$1 - p_i$
0.0	.000	1.000			
0.5	.265	.735	.265	.265	.735
1.0	.632	.368	.367	.500	.500
1.5	.864	.136	.231	.629	.371
2.0	.961	.0388	.0976	.715	.285
∞	1.000	.000	.0388	1.000	.000
			1.000		

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).
- Inspections 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 (simple).
  - non independent (difficult).
- Simple analysis requires **non-informative** censoring assumption.

Likelihood (Probability of the Data) as a Unifying Concept

- Likelihood provides a general and versatile method of estimation.
- Model/Parameters combinations with relatively 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 no general software).

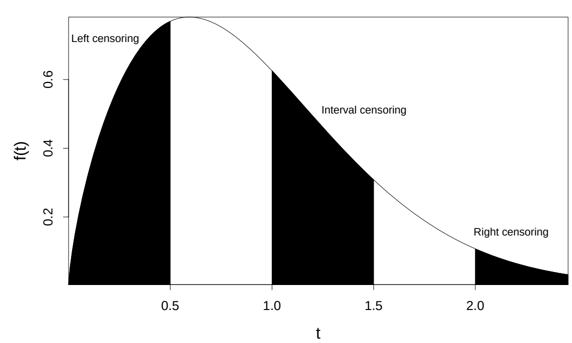
Determining the Likelihood (Probability of the Data)

The form of the likelihood will depend on:

- Question/focus of study.
- Assumed model.
- Measurement system (form of available data).
- Identifiability/parameterization.

Likelihood (Probability of the Data) Contributions for Different Kinds of Censoring

$\Pr(\textbf{Data}) = \prod_{i=1}^{n} \Pr(\textbf{data}_i) = \Pr(\textbf{data}_1) \times \cdots \times \Pr(\textbf{data}_n)$



Likelihood Contributions for Different Kinds of Censoring with $F(t) = 1 - \exp(-t^{1.7})$

- Interval-censored observations:

$$L_i(\boldsymbol{p}) = \int_{t_{i-1}}^{t_i} f(t)\,dt = F(t_i) - F(t_{i-1}).$$

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:

$$L_i(\boldsymbol{p}) = \int_0^{t_i} f(t)\,dt = F(t_i) - F(0) = F(t_i).$$

If a failure is found at the first inspection time $t = .5$, $L_i = F(.5) = .265$.

- Right-censored observations:

$$L_i(\boldsymbol{p}) = \int_{t_i}^{\infty} f(t)\,dt = F(\infty) - F(t_i) = 1 - F(t_i).$$

If a unit has not failed by the last inspection at $t = 2$, $L_i = 1 - F(2) = .0388$.

Likelihood for Life 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, data$_i$, in $(t_{i-1}, t_i]$ is

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

- Assuming that the censoring is at t_i

Type of Censoring	Characteristic	Number of Cases	Likelihood of Responses $L_i(\boldsymbol{\pi}; \text{data}_i)$
Left at t_i	$T \le t_i$	ℓ_i	$[F(t_i)]^{\ell_i}$
Interval	$t_{i-1} < T \le t_i$	d_i	$[F(t_i) - F(t_{i-1})]^{d_i}$
Right at t_i	$T > t_i$	r_i	$[1 - F(t_i)]^{r_i}$

Likelihood: Probability of the Data

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

$$\begin{aligned} L(\boldsymbol{\pi}; \text{DATA}) &= \mathcal{C} \prod_{i=1}^{n} L_i(\boldsymbol{\pi}; \text{data}_i) \\ &= \mathcal{C} \prod_{i=1}^{m+1} [F(t_i)]^{\ell_i} [F(t_i) - F(t_{i-1})]^{d_i} [1 - F(t_i)]^{r_i} \end{aligned}$$

where $n = \sum_{j=1}^{m+1} \left(d_j + r_j + \ell_j\right)$ and $\mathcal{C}$ is a constant depending on the sampling inspection scheme but not on $\boldsymbol{\pi}$. So we can take $\mathcal{C} = 1$.

- Want to find $\boldsymbol{\pi}$ so that $L(\boldsymbol{\pi})$ is large.

Likelihood for Arbitrary Censored Data

- In general, the ith observation consists of an interval $(t_i^L, t_i]$, $i = 1, \ldots, n$ $(t_i^L < t_i)$ that contains the time event T for the ith individual.

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

- Assuming that the censoring is at t_i

Type of Censoring	Characteristic	Likelihood of a single Response $L_i(\boldsymbol{\pi}; \text{data}_i)$
Left at t_i	$T \le t_i$	$F(t_i)$
Interval	$t_i^L < T \le t_i$	$F(t_i) - F(t_i^L)$
Right at t_i	$T > t_i$	$1 - F(t_i)$

Likelihood for General Reliability Data

- The total likelihood for the DATA with n independent observations is

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

- Some of the observations may have multiple occurrences. Let $(t_j^L, t_j]$, $j = 1, \ldots, k$ be the distinct intervals in the DATA and let w_j be the frequency of observation of $(t_j^L, t_j]$. Then

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

- In this case the nonparametric parameters $\boldsymbol{\pi}$ 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 $\mathcal{C}$.