Chapter 2

Models, Censoring, and Likelihood for Failure-Time Data

W. Q. Meeker, L. A. Escobar, and F. G. Pascual
Iowa State University, Louisiana State University, and Washington State University.

Based on Meeker, Escobar, and Pascual (2021): *Statistical Methods for Reliability Data, Second Edition*, John Wiley & Sons Inc.

May 24, 2021
10h 49min

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

Topics discussed in this chapter:

- 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.

Chapter 2

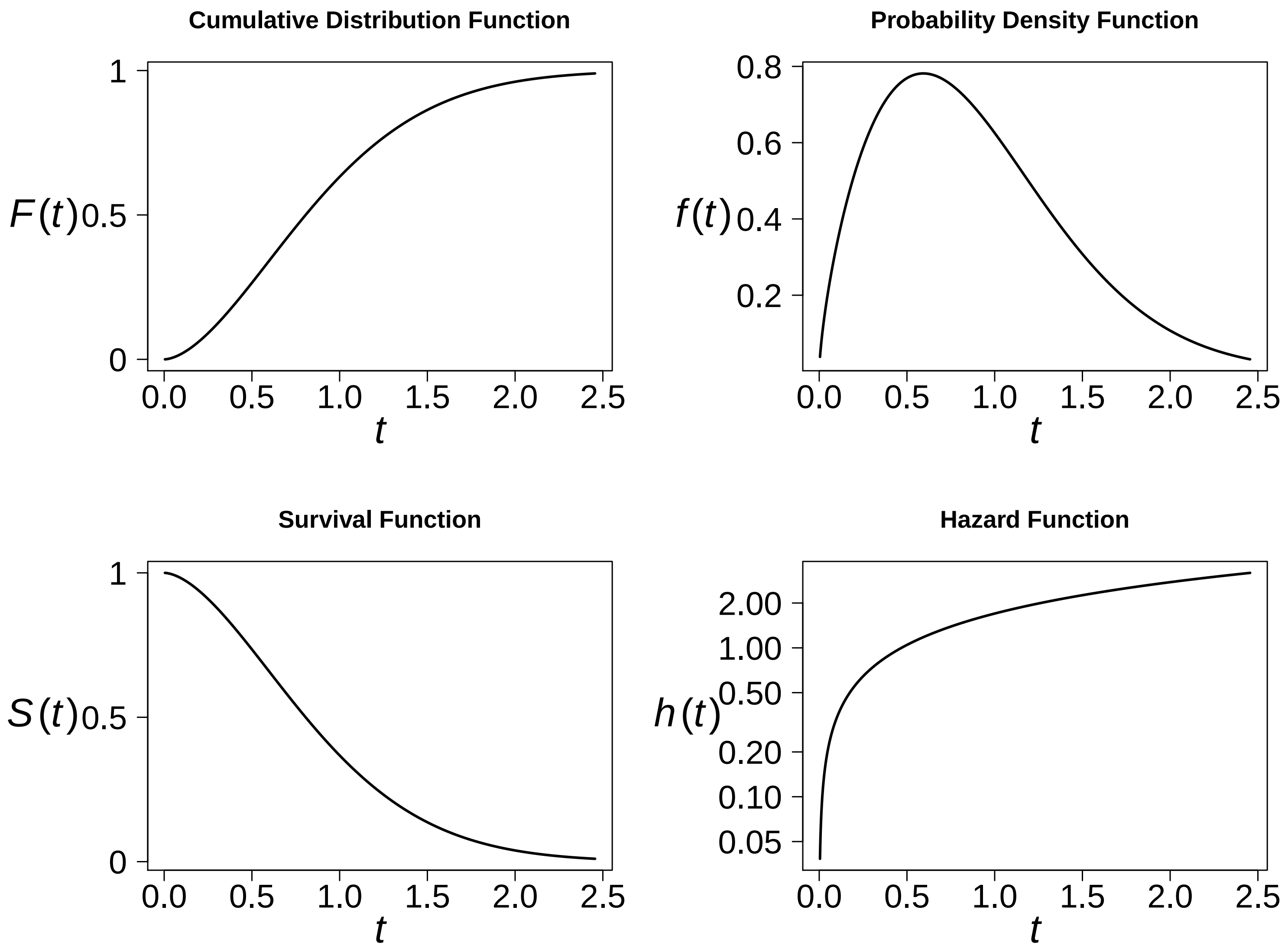
Segment 1

Failure-time Models and Metrics

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

$$F(t) = 1 - \exp(-t^{1.7}); \quad f(t) = 1.7 \times t^{0.7} \times \exp(-t^{1.7})$$

$$S(t) = \exp(-t^{1.7}); \quad h(t) = 1.7 \times t^{0.7}$$



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, $f(t) = 1.7 \times t^{0.7} \times \exp(-t^{1.7})$.

- 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, $h(t) = 1.7 \times t^{0.7}$.

Hazard Function

The hazard function is defined by

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

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

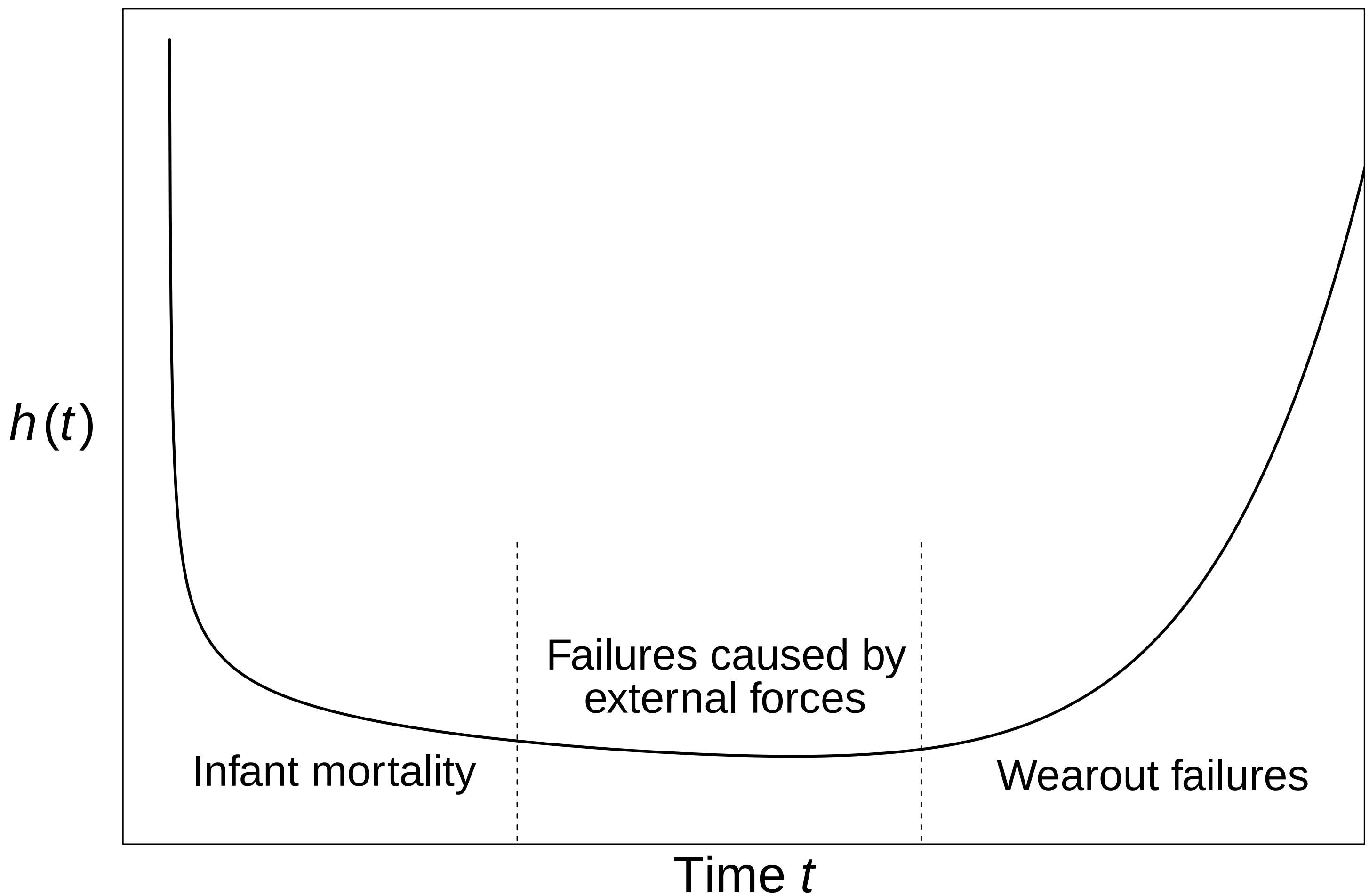
Notes:

- $F(t) = 1 - \exp[-\int_0^t h(x)dx]$.
- $h(t)$ 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 and Average Hazard

- Cumulative hazard function:

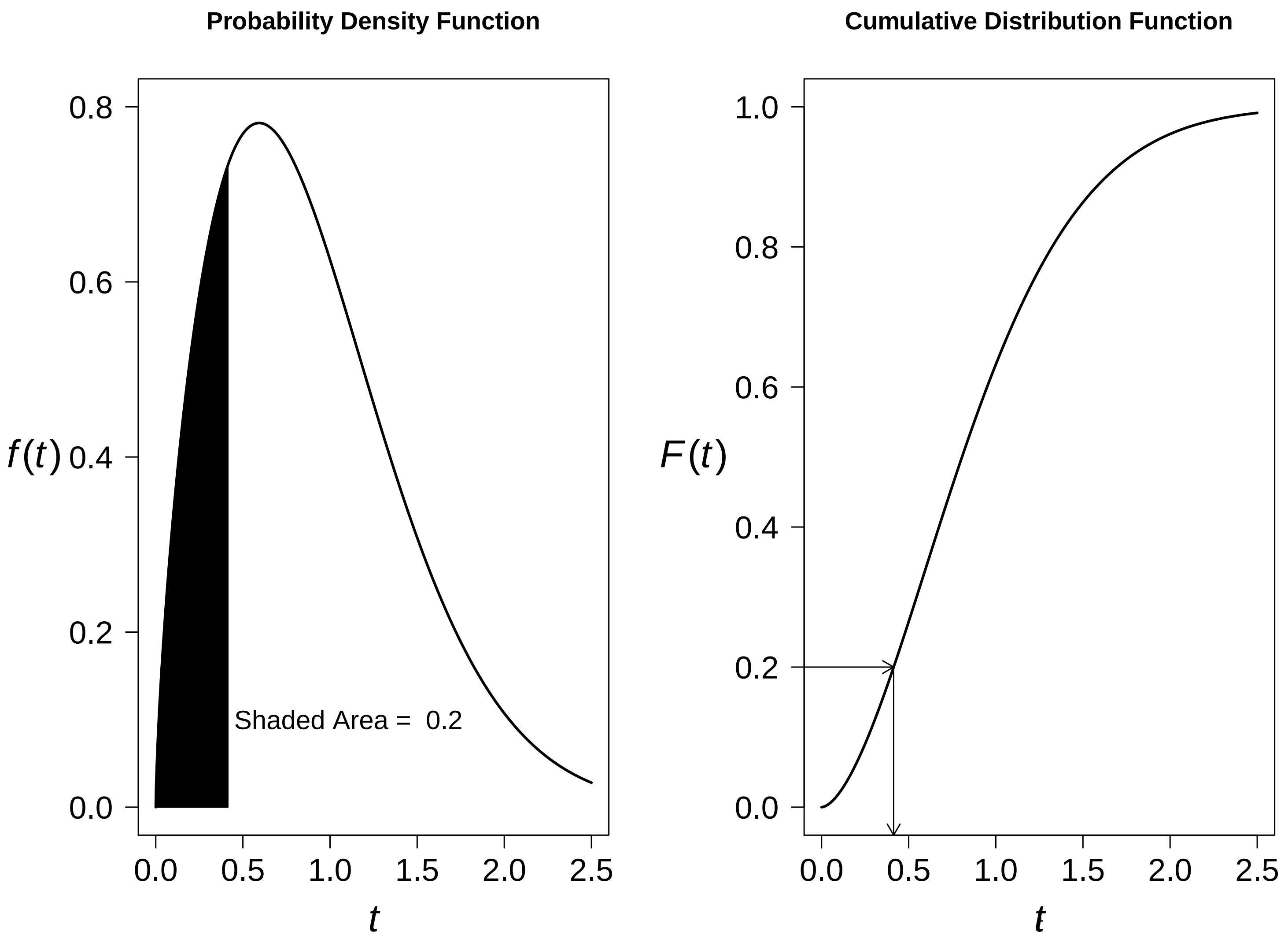
$$H(t) = \int_0^t h(x)\,dx.$$

Note that, $F(t) = 1 - \exp[-H(t)] = 1 - \exp\Big[-\int_0^t h(x)\,dx\Big].$

- The average hazard rate in the interval $(t_1, t_2]$:

$$\text{AHR}(t_1, t_2) = \frac{\int_{t_1}^{t_2} h(u)du}{t_2 - t_1} = \frac{H(t_2) - H(t_1)}{t_2 - t_1}.$$

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

$$\Pr(T \le t_p) = F(t_p) \ge p, \quad \text{where } 0 < p < 1.$$

  $t_{0.20}$ 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 $t_{0.2} = [-\log(1-0.2)]^{1/1.7} = 0.414$.

- 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 the standard convention.

- t_p is also known as B100p (e.g., $t_{0.10}$ is also known as B10).

Chapter 2

Segment 2

Distribution of Remaining Life

Distribution of Lifetime Conditional on Survival to t_0

- Consider the conditional (left-truncated) distribution

$$\Pr(T \le t | T > t_0) = \frac{F(t) - F(t_0)}{1 - F(t_0)}, \quad t \ge t_0$$

with corresponding pdf

$$\frac{f(t)}{1 - F(t_0)}, \quad t \ge t_0.$$

- This distribution is useful to describe the **age** at which a unit will fail, conditional on survival to age t_0.

Distribution of Remaining Life

- Certain applications require consideration of the distribution of **remaining life**:
  - Prediction of future field failures for a population of units that have been in service.
  - Assessment of expected remaining life of particular units.
  - Assessment of used-asset value.

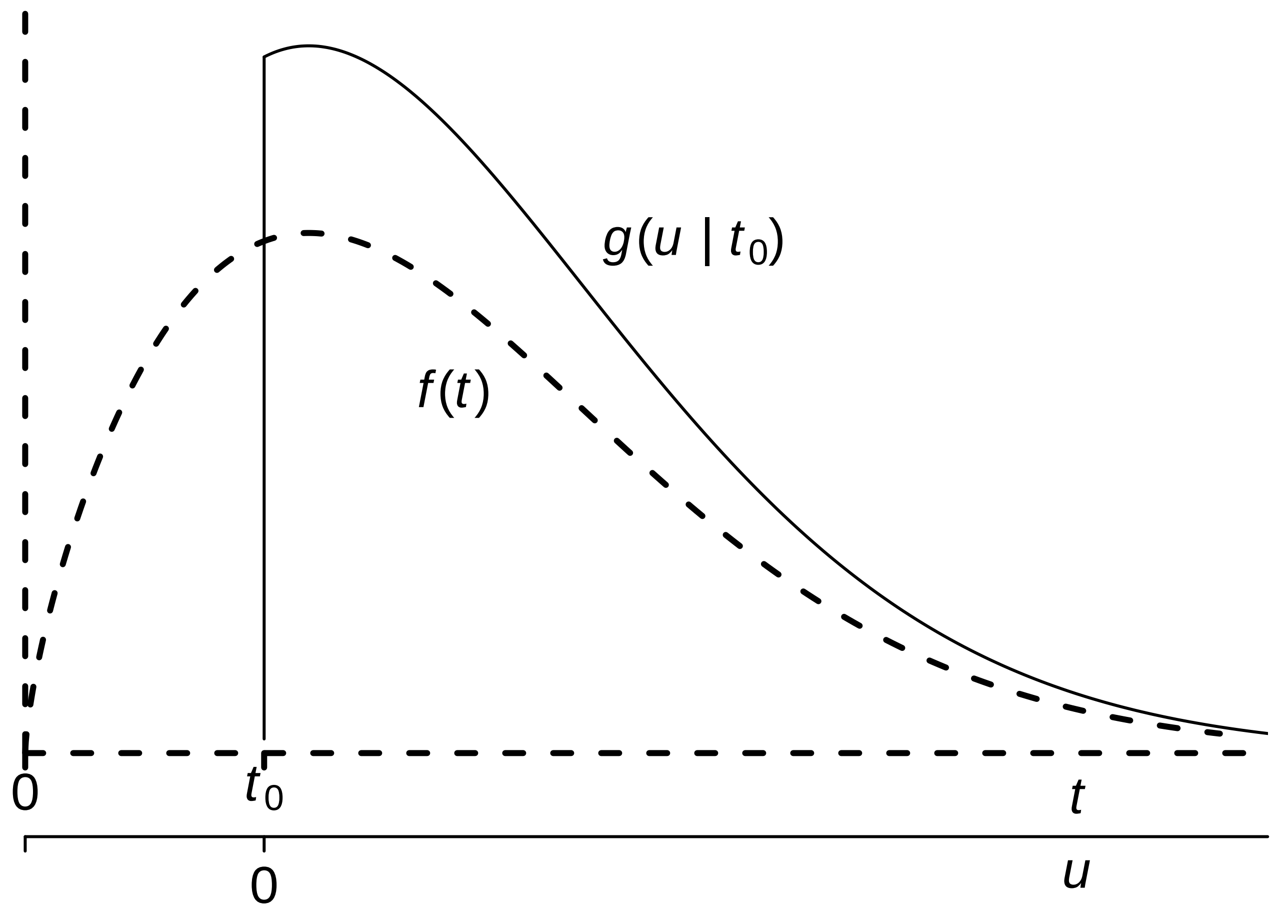
- Consider a unit with failure time T that has survived until t_0. The distribution $G(u|t_0)$ of remaining life $U = T - t_0$ is the probability that the unit will fail within the next u time units, given that $T > t_0$. That is:

$$\begin{aligned} G(u|t_0) &= \Pr(U \leq u|T > t_0) = \Pr(T - t_0 \leq u|T > t_0) \\ &= \Pr(T \leq u + t_0|T > t_0) = \frac{F(u + t_0) - F(t_0)}{1 - F(t_0)}, \quad u \geq 0. \end{aligned}$$

and the corresponding pdf is

$$g(u|t_0) = \frac{f(u + t_0)}{1 - F(t_0)}, \quad u \geq 0.$$

Distribution of Remaining Life



Example: Distribution of Remaining Life of an Automobile Transmission

- Suppose the lifetime (in thousands of miles) of an automobile transmission has a cdf

$$F(t) = 1 - \exp\left[-(t/140)^2\right], \quad t \geq 0.$$

- The cdf of the remaining life of an automobile transmission that has been in service for $t_0 = 95$ thousand miles is

$$G(u|95) = \Pr(U \leq u|95) = \frac{F(u+95) - F(95)}{1 - F(95)}$$
$$= 1 - \exp\left[\left(\frac{95}{140}\right)^2 - \left(\frac{u+95}{140}\right)^2\right], \quad u \geq 0.$$

- The probability that the automobile transmission will survive the next $u = 45$ thousand miles is

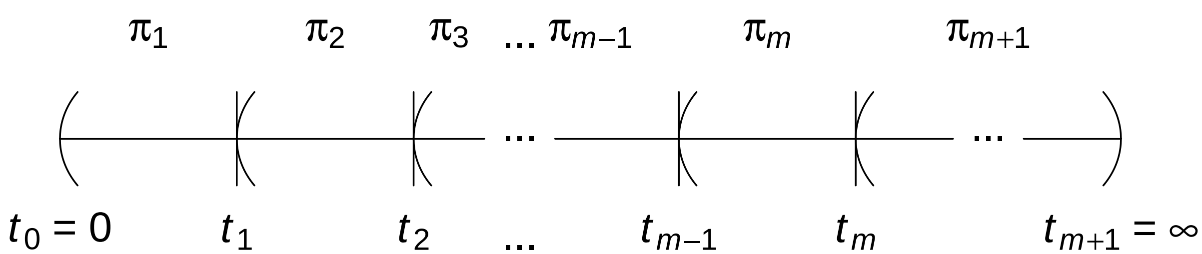
$$\Pr(U \geq 45|95) = 1 - \Pr(U \leq u|95) = \exp\left[\left(\frac{95}{140}\right)^2 - \left(\frac{45+95}{140}\right)^2\right]$$
$$= \exp\left[\left(\frac{95}{140}\right)^2 - 1\right] = 0.583$$

Chapter 2

Segment 3

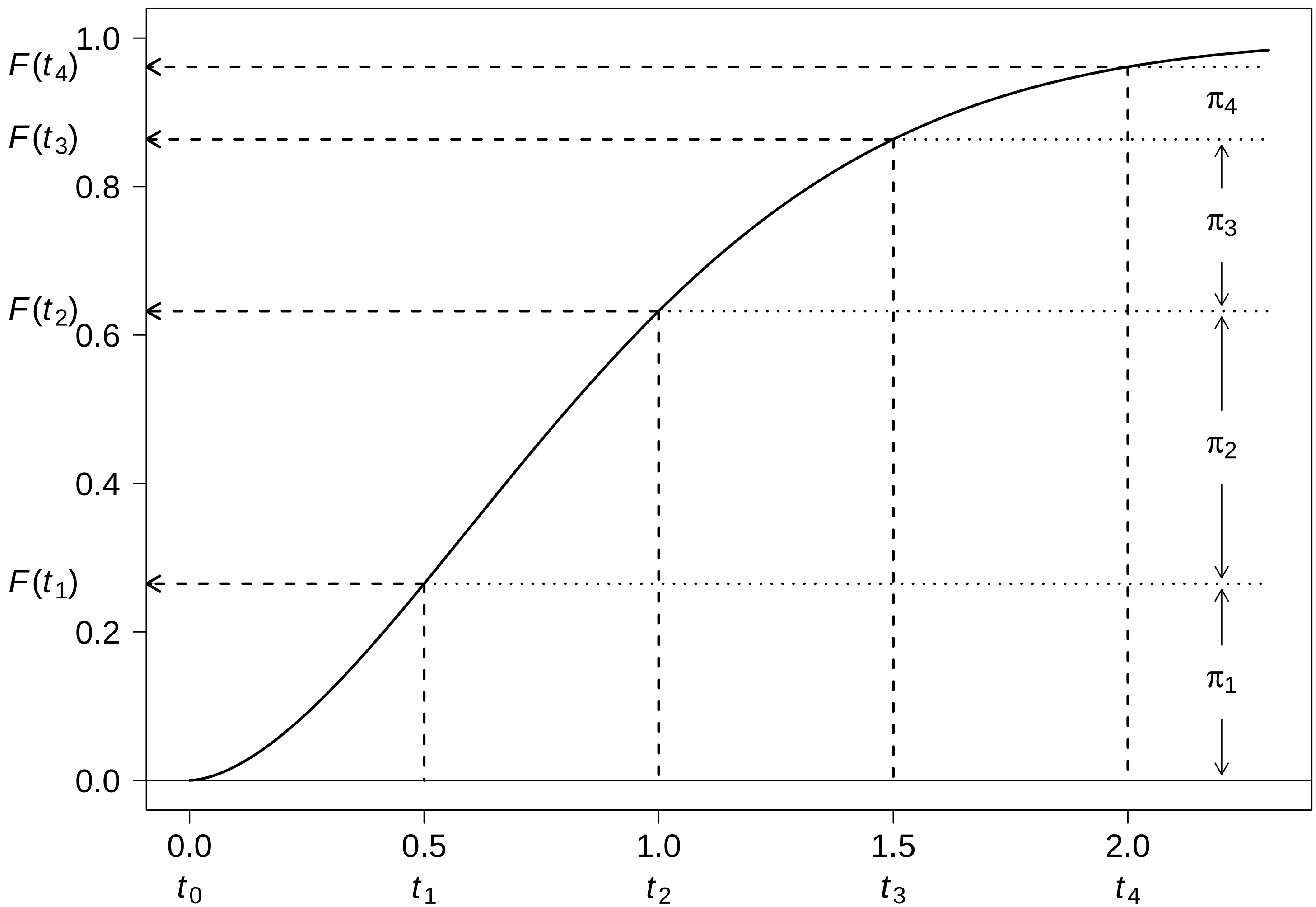
A Nonparametric Model for Failure-Time Data

Partitioning of Time into Non-Overlapping Intervals



Times need **not** be equally spaced.

Graphical Interpretation of the π's



Models for Discrete Data from 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 $t_{m+1} = \infty$. The last interval is of infinite length.

Define,

$$\pi_i = \text{Pr}(t_{i-1} < T \leq t_i) = F(t_i) - F(t_{i-1})$$

$$p_i = \text{Pr}(t_{i-1} < T \leq t_i \mid T > t_{i-1}) = \frac{F(t_i) - F(t_{i-1})}{1 - F(t_{i-1})}$$

Because the π_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 $0 \leq p_i \leq 1$

Models for Discrete Data from Continuous Time Processes–Continued

Following from the previous result,

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

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

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

Either $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_{m+1})$ or $\boldsymbol{p} = (p_1, \ldots, p_m)$ can be used as **"nonparametric 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	0.000	1.000			
0.5	0.265	0.735	0.265	0.265	0.735
1.0	0.632	0.368	0.367	0.500	0.500
1.5	0.864	0.136	0.231	0.629	0.371
2.0	0.961	0.0388	0.0976	0.715	0.285
∞	1.000	0.000	0.0388	1.000	0.000
			1.000		

Chapter 2

Segment 4

Censoring and Likelihood for Failure-Time Data

Examples of Censoring Mechanisms

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

- Fixed time to end a life 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 other random censoring mechanisms resulting in multiple right censoring:
  - Independent (simple).
  - Non independent (complicated).
- Simple modeling and analysis require **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).

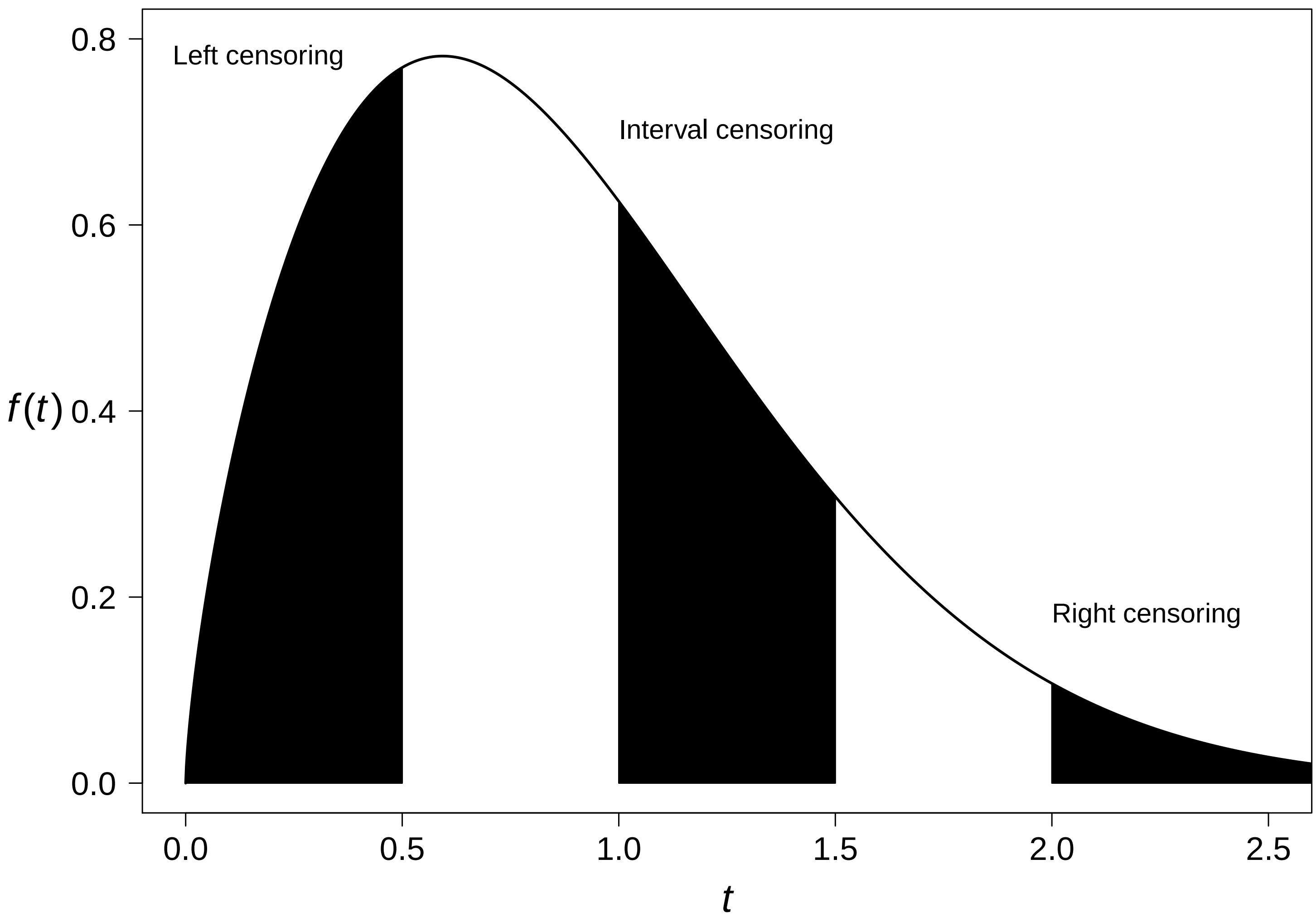
Determining the Likelihood (Probability of the Data)

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

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



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

- Interval-censored observations:

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

If a unit is still operating at $t = 1.0$ but has failed at $t = 1.5$ inspection, $L_i = F(1.5) - F(1.0) = 0.231$.

- Left-censored observations:

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

If a failure is found at the first inspection time $t = 0.5$, $L_i = F(0.5) = 0.265$.

- Right-censored observations:

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

If a unit has not failed by the last inspection at $t = 2$, $L_i = 1 - F(2) = 0.0388$.

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 (note that $F(t)$ depends on either $\boldsymbol{\pi}$ or $\boldsymbol{p}$):

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	$\left[F(t_i) - F(t_{i-1})\right]^{d_i}$
Right at t_i	$T > t_i$	r_i	$[1 - F(t_i)]^{r_i}$

Likelihood: Probability of the Failure-time Data

- The total likelihood, or joint probability of the DATA, for n **independent** observations is (note that $F(t)$ depends on either $\boldsymbol{\pi}$ or $\boldsymbol{p}$):

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

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 $L(\boldsymbol{\pi}; \text{DATA})$ is large.

Likelihood for Arbitrary Censored Failure-Time Data

- In general, observation i 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 individual i.

  The intervals $(t_i^L, t_i]$ may overlap and their union may not cover the entire timeline $(0, \infty)$. 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	$T \leq t_i$	$F(t_i)$
Interval	$t_i^L < T \leq t_i$	$F(t_i) - F(t_i^L)$
Right at t_i	$T > t_i$	$1 - F(t_i)$

Likelihood for General Failure-Time 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 (e.g., identical observations). Let $(t_j^L, t_j]$, $j = 1, \ldots, k$ be the distinct intervals in the DATA and let w_j be the number of (frequency or weight for) observations in $(t_j^L, t_j]$. 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 in $\boldsymbol{\pi}$ correspond to probabilities of a partition of $(0, \infty)$ determined by the data.

References

Meeker, W. Q., L. A. Escobar, and F. G. Pascual (2021). *Statistical Methods for Reliability Data* (Second Edition). Wiley. [1]