

Parametric Survival Models

Contents

Configuring R	1
1 Parametric Survival Models	2
1.1 Exponential Distribution	2
1.2 Weibull Distribution	3
1.2.1 Properties of Weibull hazard functions	6
1.3 Exponential Regression	10
1.4 Accelerated Failure Time	10
1.4.1 Relationship between PH and AFT models	12
1.5 Dataset: Leukemia treatments	12
1.5.1 Cox semi-parametric model	13
1.5.2 Weibull parametric model	14
1.5.3 Exponential parametric model	14
1.5.4 Diagnostic: complementary log-log survival plot	15
2 Combining left-truncation and interval-censoring	15

Configuring R

Functions from these packages will be used throughout this document:

```
library(conflicted) # check for conflicting function definitions
# library(printr) # inserts help-file output into markdown output
library(rmarkdown) # Convert R Markdown documents into a variety of formats.
library(pander) # format tables for markdown
library(ggplot2) # graphics
library(ggfortify) # help with graphics
library(dplyr) # manipulate data
library(tibble) # `tibble`s extend `data.frame`s
library(magrittr) # `>%` and other additional piping tools
library(haven) # import Stata files
library(knitr) # format R output for markdown
library(tidyr) # Tools to help to create tidy data
library(plotly) # interactive graphics
library(dobson) # datasets from Dobson and Barnett 2018
library(parameters) # format model output tables for markdown
library(haven) # import Stata files
library(latex2exp) # use LaTeX in R code (for figures and tables)
library(fs) # filesystem path manipulations
library(survival) # survival analysis
library(survminer) # survival analysis graphics
library(KMsurv) # datasets from Klein and Moeschberger
library(parameters) # format model output tables for
library(webshot2) # convert interactive content to static for pdf
library(forcats) # functions for categorical variables ("factors")
```

```
library(stringr) # functions for dealing with strings
library(lubridate) # functions for dealing with dates and times
library(broom) # Summarizes key information about statistical objects in tidy tibbles
library(broom.helpers) # Provides suite of functions to work with regression model 'broom::tidy()' t
```

Here are some R settings I use in this document:

```
rm(list = ls()) # delete any data that's already loaded into R

conflicts_prefer(dplyr::filter)
ggplot2::theme_set(
  ggplot2::theme_bw() +
    # ggplot2::labs(col = "") +
    ggplot2::theme(
      legend.position = "bottom",
      text = ggplot2::element_text(size = 12, family = "serif")))

knitr::opts_chunk$set(message = FALSE)
options('digits' = 6)

panderOptions("big.mark", ",")
pander::panderOptions("table.emphasize.rownames", FALSE)
pander::panderOptions("table.split.table", Inf)
conflicts_prefer(dplyr::filter) # use the `filter()` function from dplyr() by default
legend_text_size = 9
run_graphs = TRUE
```

1 Parametric Survival Models

1.1 Exponential Distribution

The exponential distribution is the baseline parametric model in survival analysis. It assumes a constant hazard rate $\lambda > 0$ over time. This constant hazard implies that the event probability in any short interval depends only on the interval duration and not on elapsed time (the memoryless property).

$$\begin{aligned}
f(t) &\stackrel{\text{def}}{=} \lambda e^{-\lambda t} && \text{(probability density function definition for } t \geq 0) \\
\log\{f(t)\} &= \log\{\lambda\} - \lambda t && \text{(taking natural logarithm of density)} \\
F(t) &\stackrel{\text{def}}{=} \int_0^t f(u) du && \text{(cumulative distribution function definition)} \\
&= \int_0^t \lambda e^{-\lambda u} du && \text{(substituting probability density function)} \\
&= [-e^{-\lambda u}]_0^t && \text{(evaluating antiderivative)} \\
&= 1 - e^{-\lambda t} && \text{(evaluating limits of integration)} \\
S(t) &\stackrel{\text{def}}{=} 1 - F(t) && \text{(survival function definition)} \\
&= 1 - (1 - e^{-\lambda t}) && \text{(substituting cumulative distribution function)} \\
&= e^{-\lambda t} && \text{(simplifying terms)} \\
\Lambda(t) &\stackrel{\text{def}}{=} -\log\{S(t)\} && \text{(cumulative hazard function definition)} \\
&= -\log\{e^{-\lambda t}\} && \text{(substituting survival function)} \\
&= \lambda t && \text{(simplifying logarithm of exponential)} \\
\lambda(t) &\stackrel{\text{def}}{=} \frac{\partial}{\partial t} \Lambda(t) && \text{(hazard function definition as derivative of cumulative hazard)} \\
&= \frac{\partial}{\partial t} (\lambda t) && \text{(substituting cumulative hazard)} \\
&= \lambda && \text{(differentiating linear term)} \\
E[T] &\stackrel{\text{def}}{=} \int_0^\infty S(t) dt && \text{(expectation formula for non-negative continuous random variable)} \\
&= \int_0^\infty e^{-\lambda t} dt && \text{(substituting survival function)} \\
&= \left[-\frac{1}{\lambda} e^{-\lambda t} \right]_0^\infty && \text{(evaluating antiderivative)} \\
&= \lambda^{-1} && \text{(evaluating limits of integration)}
\end{aligned}$$

1.2 Weibull Distribution

The Weibull distribution generalizes the exponential distribution by introducing a shape parameter $p > 0$ alongside the scale parameter $\lambda > 0$ (Kalbfleisch and Prentice 2011). This flexibility allows the hazard rate to change monotonically over time, accommodating situations where risk increases or decreases as time elapses.

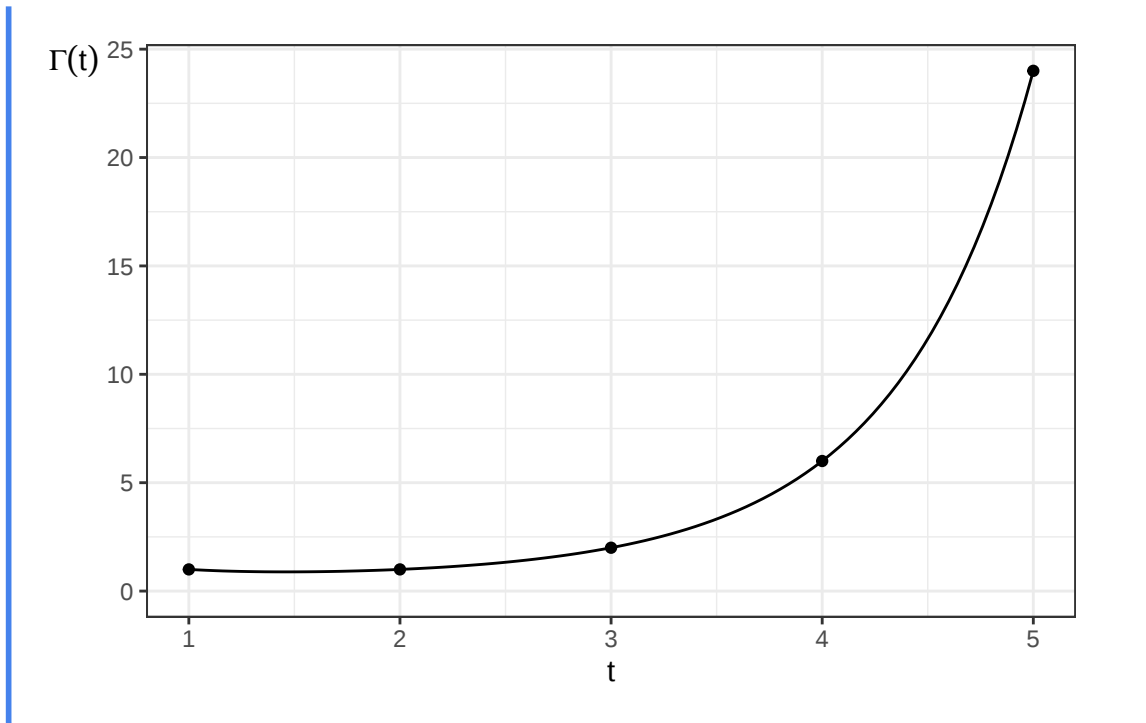
Using the parameterization of Kalbfleisch and Prentice (2011):

$$\begin{aligned}
f(t) &\stackrel{\text{def}}{=} \lambda p (\lambda t)^{p-1} e^{-(\lambda t)^p} && \text{(probability density function for } t \geq 0) \\
F(t) &\stackrel{\text{def}}{=} \int_0^t f(u) du && \text{(cumulative distribution function definition)} \\
&= \int_0^t \lambda p (\lambda u)^{p-1} e^{-(\lambda u)^p} du && \text{(substituting density function)} \\
&= [-e^{-(\lambda u)^p}]_0^t && \text{(evaluating antiderivative via substitution } w = (\lambda u)^p) \\
&= 1 - e^{-(\lambda t)^p} && \text{(evaluating limits of integration)} \\
S(t) &\stackrel{\text{def}}{=} 1 - F(t) && \text{(survival function definition)} \\
&= 1 - (1 - e^{-(\lambda t)^p}) && \text{(substituting cumulative distribution function)} \\
&= e^{-(\lambda t)^p} && \text{(simplifying terms)} \\
\Lambda(t) &\stackrel{\text{def}}{=} -\log\{S(t)\} && \text{(cumulative hazard function definition)} \\
&= -\log\{e^{-(\lambda t)^p}\} && \text{(substituting survival function)} \\
&= (\lambda t)^p && \text{(simplifying logarithm of exponential)} \\
\log\{\Lambda(t)\} &= \log\{(\lambda t)^p\} && \text{(taking natural logarithm of cumulative hazard)} \\
&= p \log\{\lambda t\} && \text{(applying exponent rule for logarithms)} \\
&= p \log\{\lambda\} + p \log\{t\} && \text{(applying product rule for logarithms)} \\
\lambda(t) &\stackrel{\text{def}}{=} \frac{\partial}{\partial t} \Lambda(t) && \text{(hazard function definition)} \\
&= \frac{\partial}{\partial t} ((\lambda t)^p) && \text{(substituting cumulative hazard)} \\
&= p(\lambda t)^{p-1} \cdot \lambda && \text{(applying power and chain rules of differentiation)} \\
&= \lambda p (\lambda t)^{p-1} && \text{(rearranging factors)} \\
E[T] &\stackrel{\text{def}}{=} \int_0^\infty S(t) dt && \text{(expectation formula for non-negative random variable)} \\
&= \int_0^\infty e^{-(\lambda t)^p} dt && \text{(substituting survival function)} \\
&= \frac{1}{\lambda p} \int_0^\infty u^{\frac{1}{p}-1} e^{-u} du && \text{(substituting } u = (\lambda t)^p \implies t = \lambda^{-1} u^{1/p}) \\
&= \lambda^{-1} \cdot \Gamma\left(1 + \frac{1}{p}\right) && \text{(applying definition of Gamma function } \Gamma(z))
\end{aligned}$$

i Note

Recall from calculus:

- $\Gamma(t) \stackrel{\text{def}}{=} \int_{u=0}^\infty u^{t-1} e^{-u} du$
- $\Gamma(t) = (t-1)!$ for integers $t \in \mathbb{Z}$
- It is implemented by the `gamma()` function in R.



Here are some Weibull density functions, with $\lambda = 1$ and p varying:

```
library(ggplot2)
lambda <- 1
ggplot() +
  geom_function(
    aes(col = "0.25"),
    fun = \(x) dweibull(x, shape = 0.25, scale = 1 / lambda)
  ) +
  geom_function(
    aes(col = "0.5"),
    fun = \(x) dweibull(x, shape = 0.5, scale = 1 / lambda)
  ) +
  geom_function(
    aes(col = "1"),
    fun = \(x) dweibull(x, shape = 1, scale = 1 / lambda)
  ) +
  geom_function(
    aes(col = "1.5"),
    fun = \(x) dweibull(x, shape = 1.5, scale = 1 / lambda)
  ) +
  geom_function(
    aes(col = "2"),
    fun = \(x) dweibull(x, shape = 2, scale = 1 / lambda)
  ) +
  geom_function(
    aes(col = "5"),
    fun = \(x) dweibull(x, shape = 5, scale = 1 / lambda)
  ) +
  theme_bw() +
  xlim(0, 2.5) +
  ylab("f(t)") +
  theme(axis.title.y = element_text(angle = 0)) +
  theme(legend.position = "bottom") +
  guides(
```

```
col =
  guide_legend(
    title = "p",
    label.theme = element_text(size = 12)
  )
)
```

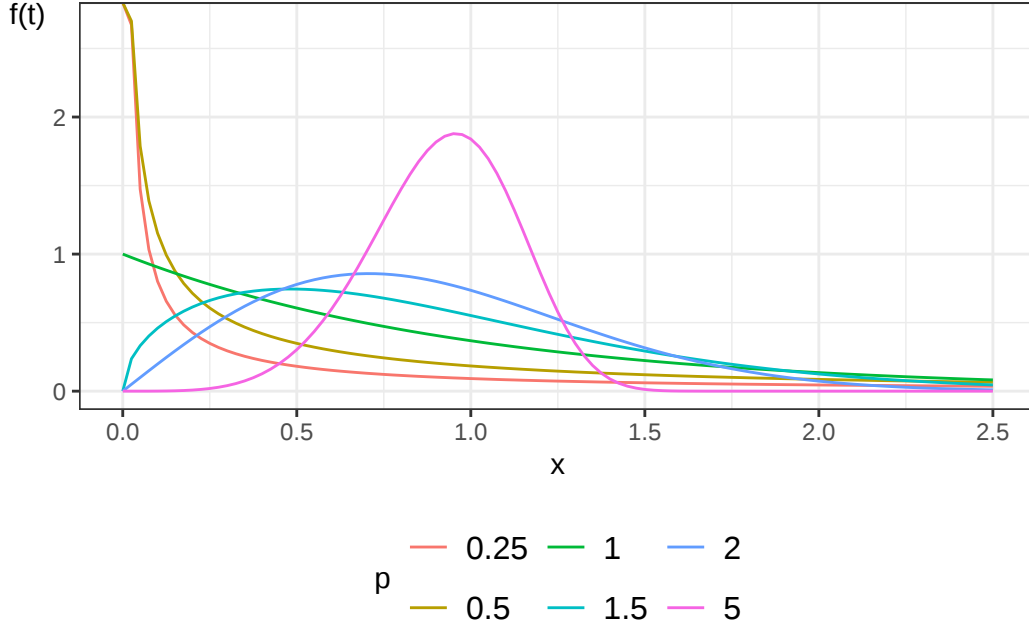


Figure 1: Density functions for Weibull distribution

1.2.1 Properties of Weibull hazard functions

Theorem 1.1 (Properties of Weibull hazard functions). *If T has a Weibull distribution, then:*

- When $p = 1$, the Weibull distribution simplifies to the exponential distribution.
- When $p > 1$, the hazard is strictly increasing: $\lambda'(t) > 0$.
- When $p < 1$, the hazard is strictly decreasing: $\lambda'(t) < 0$.
- $\log \Lambda(t)$ is a straight line relative to $\log t$: $\log \Lambda(t) = p \log \lambda + p \log t$.

i Proof

Proof. We prove each property of the Weibull hazard function $\lambda(t) = \lambda p (\lambda t)^{p-1}$:

1. **Simplification to Exponential distribution when $p = 1$:** Setting $p = 1$ in the Weibull hazard function yields:

$$\begin{aligned} \lambda(t) &= \lambda(1)(\lambda t)^{1-1} && \text{(substituting } p = 1 \text{ into Weibull hazard formula)} \\ &= \lambda \cdot 1 \cdot (\lambda t)^0 && \text{(simplifying exponent } 1 - 1 = 0) \\ &= \lambda && \text{(since } (\lambda t)^0 = 1 \text{ for } t > 0) \end{aligned}$$

This constant value is the hazard function of the exponential distribution with parameter λ .

2. **Increasing hazard when $p > 1$:** Differentiating $\lambda(t) = \lambda p t^{p-1}$ with respect to t gives:

$$\begin{aligned} \lambda'(t) &\stackrel{\text{def}}{=} \frac{\partial}{\partial t} (\lambda p t^{p-1}) && \text{(definition of hazard derivative)} \\ &= \lambda p (p-1) t^{p-2} && \text{(applying power rule of differentiation)} \end{aligned}$$

Since $\lambda > 0$, $p > 0$, and $t > 0$, the sign of $\lambda'(t)$ depends entirely on the factor $(p - 1)$. When $p > 1$, $(p - 1) > 0$, so $\lambda'(t) > 0$ for all $t > 0$, meaning the hazard function is strictly increasing over time.

3. **Decreasing hazard when $p < 1$:** Using the derivative $\lambda'(t) = \lambda^p p(p - 1)t^{p-2}$: When $p < 1$, $(p - 1) < 0$, so $\lambda'(t) < 0$ for all $t > 0$, meaning the hazard function is strictly decreasing over time.
4. **Linear relationship between $\log \Lambda(t)$ and $\log t$:** Taking the natural logarithm of the cumulative hazard function $\Lambda(t) = (\lambda t)^p$:

$$\begin{aligned}\log\{\Lambda(t)\} &\stackrel{\text{def}}{=} \log\{(\lambda t)^p\} && \text{(substituting Weibull cumulative hazard)} \\ &= p \log\{\lambda t\} && \text{(applying logarithm exponent rule } \log(a^b) = b \log a) \\ &= p \log\{\lambda\} + p \log\{t\} && \text{(applying logarithm product rule } \log(ab) = \log a + \log b)\end{aligned}$$

Defining $y \stackrel{\text{def}}{=} \log\{\Lambda(t)\}$ and $x \stackrel{\text{def}}{=} \log\{t\}$, this equation takes the linear form $y = a + bx$ with intercept $a = p \log\{\lambda\}$ and slope $b = p$.

□

Exm

Example 1.1 (Evaluating Weibull hazard shapes under different shape parameters). Consider a clinical trial evaluating time to relapse in months following cancer treatment, modeled using a Weibull distribution with scale parameter $\lambda = 0.1$. We evaluate the hazard rate $\lambda(t)$ at months $t = 1, 6, 12$ under three distinct shape parameters:

1. **Decreasing hazard ($p = 0.5$):** The hazard rate formula gives:

$$\lambda(t) = (0.1)(0.5)(0.1t)^{-0.5} = \frac{0.05}{\sqrt{0.1t}}$$

- At $t = 1$ month: $\lambda(1) = \frac{0.05}{\sqrt{0.1}} \approx 0.1581$ events per month.
 - At $t = 6$ months: $\lambda(6) = \frac{0.05}{\sqrt{0.6}} \approx 0.0645$ events per month.
 - At $t = 12$ months: $\lambda(12) = \frac{0.05}{\sqrt{1.2}} \approx 0.0456$ events per month. The risk of relapse is highest immediately following treatment and declines over time.
2. **Constant hazard ($p = 1.0$):** The hazard rate formula reduces to $\lambda(t) = \lambda = 0.1000$ events per month for all t . The risk of relapse remains constant regardless of time elapsed.
 3. **Increasing hazard ($p = 2.0$):** The hazard rate formula gives:

$$\lambda(t) = (0.1)(2.0)(0.1t)^{1.0} = 0.02t$$

- At $t = 1$ month: $\lambda(1) = 0.02(1) = 0.0200$ events per month.
- At $t = 6$ months: $\lambda(6) = 0.02(6) = 0.1200$ events per month.
- At $t = 12$ months: $\lambda(12) = 0.02(12) = 0.2400$ events per month. The risk of relapse increases linearly over time.

Exercise 1.1. Verify the calculations in Example 1.1 and explain how the shape parameter p affects cumulative hazard curves.

The Weibull distribution provides more flexibility than the exponential. Figure 2 shows some Weibull hazard functions, with $\lambda = 1$ and p varying:

```
library(ggplot2)
library(eha)
lambda <- 1

ggplot() +
  geom_function(
```

```

    aes(col = "0.25"),
    fun = \(x) hweibull(x, shape = 0.25, scale = 1 / lambda)
  ) +
  geom_function(
    aes(col = "0.5"),
    fun = \(x) hweibull(x, shape = 0.5, scale = 1 / lambda)
  ) +
  geom_function(
    aes(col = "1"),
    fun = \(x) hweibull(x, shape = 1, scale = 1 / lambda)
  ) +
  geom_function(
    aes(col = "1.5"),
    fun = \(x) hweibull(x, shape = 1.5, scale = 1 / lambda)
  ) +
  geom_function(
    aes(col = "2"),
    fun = \(x) hweibull(x, shape = 2, scale = 1 / lambda)
  ) +
  theme_bw() +
  xlim(0, 2.5) +
  ylab(expr(lambda)) +
  theme(axis.title.y = element_text(angle = 0)) +
  theme(legend.position = "bottom") +
  guides(
    col =
      guide_legend(
        title = "p",
        label.theme = element_text(size = 12)
      )
  )
)

```

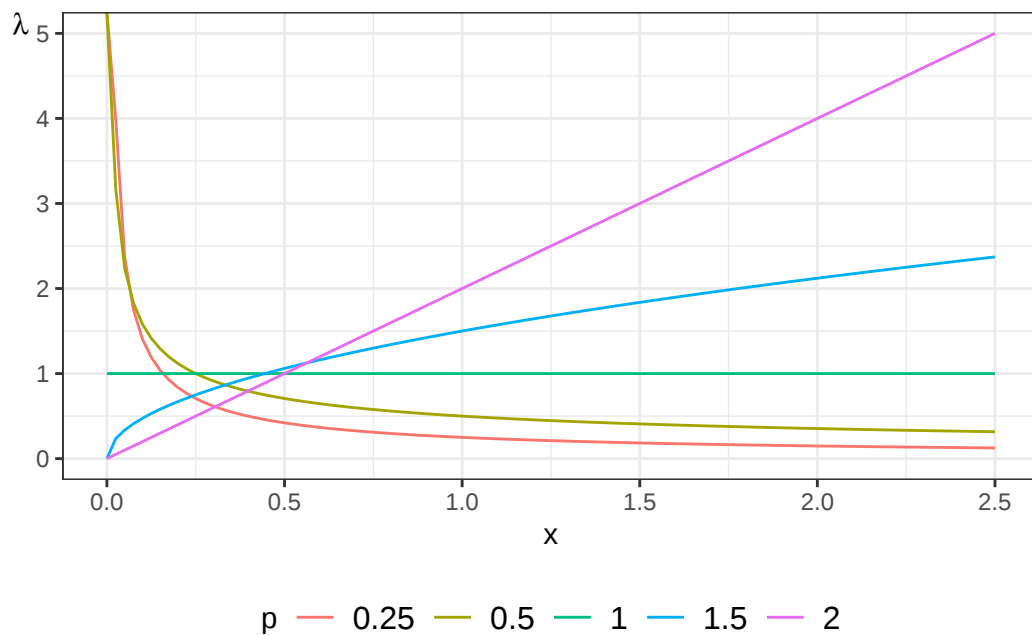


Figure 2: Hazard functions for Weibull distribution

```

library(ggplot2)
lambda <- 1

ggplot() +
  geom_function(
    aes(col = "0.25"),
    fun = pweibull,
    args = list(shape = 0.25, scale = 1 / lambda, lower.tail = FALSE)
  ) +
  geom_function(
    aes(col = "0.5"),
    fun = pweibull,
    args = list(shape = 0.5, scale = 1 / lambda, lower.tail = FALSE)
  ) +
  geom_function(
    aes(col = "1"),
    fun = pweibull,
    args = list(shape = 1, scale = 1 / lambda, lower.tail = FALSE)
  ) +
  geom_function(
    aes(col = "1.5"),
    fun = pweibull,
    args = list(shape = 1.5, scale = 1 / lambda, lower.tail = FALSE)
  ) +
  geom_function(
    aes(col = "2"),
    fun = pweibull,
    args = list(shape = 2, scale = 1 / lambda, lower.tail = FALSE)
  ) +
  theme_bw() +
  xlim(0, 2.5) +
  ylab("S(t)") +
  theme(axis.title.y = element_text(angle = 0)) +
  theme(legend.position = "bottom") +
  guides(
    col =
      guide_legend(
        title = "p",
        label.theme = element_text(size = 12)
      )
  )
)

```

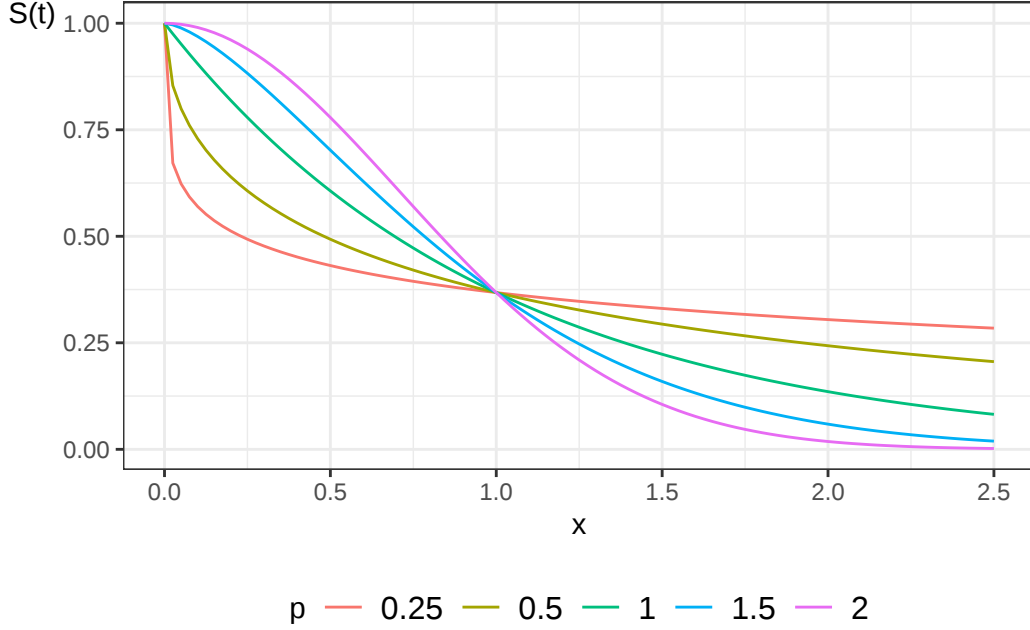


Figure 3: Survival functions for Weibull distribution

1.3 Exponential Regression

For each subject i , define a linear predictor $\eta(\tilde{x}) \stackrel{\text{def}}{=} \beta_0 + (\beta_1 x_1 + \dots + \beta_p x_p) = \tilde{x} \cdot \tilde{\beta}$. Exponential regression models the hazard rate using a log link function:

$$\begin{aligned}
 \lambda(t \mid \tilde{x}) &\stackrel{\text{def}}{=} \exp\{\eta(\tilde{x})\} && \text{(exponential regression hazard specification)} \\
 \lambda_0(t) &\stackrel{\text{def}}{=} \lambda(t \mid \tilde{0}) && \text{(baseline hazard rate definition at } \tilde{x} = \tilde{0} \text{)} \\
 &= \exp\{\eta(\tilde{0})\} && \text{(substituting } \tilde{x} = \tilde{0} \text{ into hazard function)} \\
 &= \exp\{\beta_0 + (\beta_1 \cdot 0 + \dots + \beta_p \cdot 0)\} && \text{(evaluating linear predictor at zero vector)} \\
 &= \exp\{\beta_0\} && \text{(simplifying exponent)}
 \end{aligned}$$

When there are no additional predictors ($\tilde{x} = \tilde{0}$), the baseline hazard is constant: $\lambda = \exp\{\beta_0\}$. This model uses a log link as in a generalized linear model for Poisson or rate outcomes. Because the hazard rate does not depend on time t , the hazard ratio comparing two covariate vectors $\tilde{x}$ and $\tilde{x}^*$ is constant over time:

$$\begin{aligned}
 \theta_\lambda(t \mid \tilde{x} : \tilde{x}^*) &\stackrel{\text{def}}{=} \frac{\lambda(t \mid \tilde{x})}{\lambda(t \mid \tilde{x}^*)} && \text{(hazard ratio definition)} \\
 &= \frac{\exp\{\eta(\tilde{x})\}}{\exp\{\eta(\tilde{x}^*)\}} && \text{(substituting exponential regression hazard rates)} \\
 &= \exp\{\eta(\tilde{x}) - \eta(\tilde{x}^*)\} && \text{(applying quotient rule for exponentials)} \\
 &= \exp\{(\tilde{x} - \tilde{x}^*) \cdot \tilde{\beta}\} && \text{(expressing difference in linear predictors)}
 \end{aligned}$$

Thus, exponential regression models possess proportional hazards trivially.

1.4 Accelerated Failure Time

In a proportional hazards (PH) model, covariates act multiplicatively on the baseline hazard function:

$$\begin{aligned}
\lambda(t \mid \tilde{x}) &\stackrel{\text{def}}{=} \lambda_0(t) \cdot \theta_\lambda(\tilde{x}) && (\text{proportional hazards hazard rate definition}) \\
\Lambda(t \mid \tilde{x}) &\stackrel{\text{def}}{=} \int_0^t \lambda(u \mid \tilde{x}) du && (\text{cumulative hazard function definition}) \\
&= \int_0^t \lambda_0(u) \cdot \theta_\lambda(\tilde{x}) du && (\text{substituting proportional hazard rate}) \\
&= \theta_\lambda(\tilde{x}) \cdot \Lambda_0(t) && (\text{factoring constant hazard multiplier out of integral}) \\
S(t \mid \tilde{x}) &\stackrel{\text{def}}{=} \exp\{-\Lambda(t \mid \tilde{x})\} && (\text{survival function in terms of cumulative hazard}) \\
&= \exp\{-\theta_\lambda(\tilde{x}) \cdot \Lambda_0(t)\} && (\text{substituting cumulative hazard}) \\
&= [\exp\{-\Lambda_0(t)\}]^{\theta_\lambda(\tilde{x})} && (\text{applying exponent laws}) \\
&= [S_0(t)]^{\theta_\lambda(\tilde{x})} && (\text{substituting baseline survival function } S_0(t))
\end{aligned}$$

An alternative modeling framework is the **Accelerated Failure Time (AFT)** model. Rather than multiplying the hazard rate, covariates accelerate or decelerate the progression of time itself:

$$S(t \mid \tilde{x}) \stackrel{\text{def}}{=} S_0(t \cdot \theta_\lambda(\tilde{x}))$$

where $\theta_\lambda(\tilde{x}) \stackrel{\text{def}}{=} \exp\{\eta(\tilde{x})\}$, $\eta(\tilde{x}) \stackrel{\text{def}}{=} \beta_1 x_1 + \dots + \beta_p x_p$, and $S_0(t) \stackrel{\text{def}}{=} P(T \geq t \mid \tilde{x} = \tilde{0})$ is the baseline survival function.

We derive the expected survival time under the AFT model using change-of-variables integration:

$$\begin{aligned}
E[T \mid \tilde{x}] &\stackrel{\text{def}}{=} \int_0^\infty S(t \mid \tilde{x}) dt && (\text{expectation formula for non-negative random variable}) \\
&= \int_0^\infty S_0(t \cdot \theta_\lambda(\tilde{x})) dt && (\text{substituting AFT survival function}) \\
&= \int_0^\infty S_0(u) \cdot \frac{du}{\theta_\lambda(\tilde{x})} && (\text{substituting } u = t \cdot \theta_\lambda(\tilde{x}) \implies dt = \frac{du}{\theta_\lambda(\tilde{x})}) \\
&= \theta_\lambda(\tilde{x})^{-1} \cdot \int_0^\infty S_0(u) du && (\text{factoring constant } \theta_\lambda(\tilde{x})^{-1} \text{ out of integral}) \\
&= \theta_\lambda(\tilde{x})^{-1} \cdot E[T \mid \tilde{x} = \tilde{0}] && (\text{substituting baseline expectation formula})
\end{aligned}$$

Thus, the mean survival time for a subject with covariates $\tilde{x}$ equals the baseline mean survival time divided by the acceleration factor $\theta_\lambda(\tilde{x}) = \exp\{\eta(\tilde{x})\}$.

We also derive the cumulative hazard and hazard rate functions under the AFT formulation:

$$\begin{aligned}
\Lambda(t \mid \tilde{x}) &\stackrel{\text{def}}{=} -\log\{S(t \mid \tilde{x})\} && (\text{cumulative hazard function definition}) \\
&= -\log\{S_0(t \cdot \theta_\lambda(\tilde{x}))\} && (\text{substituting AFT survival function}) \\
&= \Lambda_0(t \cdot \theta_\lambda(\tilde{x})) && (\text{substituting baseline cumulative hazard}) \\
\lambda(t \mid \tilde{x}) &\stackrel{\text{def}}{=} \frac{\partial}{\partial t} \Lambda(t \mid \tilde{x}) && (\text{hazard function definition}) \\
&= \frac{\partial}{\partial t} (\Lambda_0(t \cdot \theta_\lambda(\tilde{x}))) && (\text{substituting AFT cumulative hazard}) \\
&= \Lambda'_0(t \cdot \theta_\lambda(\tilde{x})) \cdot \frac{\partial}{\partial t} (t \cdot \theta_\lambda(\tilde{x})) && (\text{applying chain rule of differentiation}) \\
&= \lambda_0(t \cdot \theta_\lambda(\tilde{x})) \cdot \theta_\lambda(\tilde{x}) && (\text{since } \Lambda'_0(u) = \lambda_0(u)) \\
&= \theta_\lambda(\tilde{x}) \cdot \lambda_0(t \cdot \theta_\lambda(\tilde{x})) && (\text{rearranging factors})
\end{aligned}$$

1.4.1 Relationship between PH and AFT models

If the baseline distribution is exponential with parameter λ , substituting $S_0(t) = \exp\{-\lambda t\}$ into the AFT model yields:

$$\begin{aligned} S(t \mid \tilde{x}) &= \exp\{-\lambda \cdot t \cdot \theta_\lambda(\tilde{x})\} && \text{(substituting exponential baseline into AFT formula)} \\ &= [\exp\{-\lambda t\}]^{\theta_\lambda(\tilde{x})} && \text{(applying exponent laws)} \\ &= [S_0(t)]^{\theta_\lambda(\tilde{x})} && \text{(substituting exponential baseline survival function)} \end{aligned}$$

This derivation demonstrates that the exponential AFT model is mathematically identical to the exponential PH model.

For a Weibull baseline distribution with scale λ and shape p , the hazard and survival functions are $\lambda(t) = \lambda p(\lambda t)^{p-1}$ and $S(t) = \exp\{-(\lambda t)^p\}$. Specifying a proportional hazards model with subject-specific multiplier $\theta_i \stackrel{\text{def}}{=} \exp\{\eta_i\}$ gives hazard rate $\lambda(t) = \lambda p(\lambda t)^{p-1} \theta_i$. Defining an adjusted scale parameter $\lambda^* \stackrel{\text{def}}{=} \lambda \theta_i^{1/p}$, the resulting survival function is:

$$\begin{aligned} S^*(t) &\stackrel{\text{def}}{=} \exp\{-(\lambda^* t)^p\} && \text{(Weibull survival formula with scale parameter } \lambda^*) \\ &= \exp\left\{-\left(\lambda \theta_i^{1/p} t\right)^p\right\} && \text{(substituting } \lambda^* = \lambda \theta_i^{1/p}) \\ &= \exp\{-(\lambda t)^p \cdot \theta_i\} && \text{(simplifying exponent } (\theta_i^{1/p})^p = \theta_i) \\ &= [\exp\{-(\lambda t)^p\}]^{\theta_i} && \text{(applying exponent laws)} \\ &= [S_0(t)]^{\theta_i} && \text{(substituting baseline Weibull survival function)} \\ &= S_0(t \cdot \theta_i^{1/p}) && \text{(expressing as AFT model with acceleration factor } \theta_i^{1/p}) \end{aligned}$$

This derivation proves that the Weibull model is simultaneously a proportional hazards model and an accelerated failure time model.

In terms of log survival time $Y \stackrel{\text{def}}{=} \log T$, the Weibull AFT model takes a log-linear regression form:

$$\begin{aligned} Y &= \alpha - \sigma \eta + \sigma W && \text{(log-linear AFT regression formulation)} \\ \alpha &\stackrel{\text{def}}{=} -\log\{\lambda\} && \text{(intercept parameter definition)} \\ \sigma &\stackrel{\text{def}}{=} \frac{1}{p} && \text{(scale parameter definition as inverse shape } p) \end{aligned}$$

where W follows the standard Gumbel (extreme value) distribution. The estimated AFT regression coefficients $\hat{\beta}_{\text{AFT}}$ produced by R's `survreg()` function relate directly to the PH regression coefficients $\hat{\beta}_{\text{PH}}$ produced by `coxph()` via the identity $\hat{\beta}_{\text{AFT}} = -\sigma \hat{\beta}_{\text{PH}}$.

The exponential and Weibull distributions are the only continuous distributions that are simultaneously proportional hazards models and accelerated failure time models. Other parametric families commonly used for survival data (log-logistic, log-normal, and generalized gamma) are AFT families rather than PH families; the log-logistic distribution is additionally a proportional-odds model.

1.5 Dataset: Leukemia treatments

To illustrate semi-parametric and parametric survival models, we analyze remission survival times from a clinical trial of 42 pediatric leukemia patients (Kalbfleisch and Prentice 2011). Half of the patients received a new therapy (6-mercaptopurine, 6-MP) and half received a standard control treatment.

The variables in this dataset include:

- **survt**: Remission duration in weeks (T).
- **status**: Relapse status (**relapse** = 1 or **censored** = 0).
- **rx**: Treatment assignment (**new** vs **standard**).

- **sex**: Patient sex (female vs male).
- **surv**: Survival outcome object created with `Surv(time = survt, event = (status == "relapse"))`.

```
library(haven)
library(survival)
anderson <-
  paste0(
    "http://web1.sph.emory.edu/dkleinb/allDatasets",
    "/surv2datasets/anderson.dta"
  ) |>
  read_dta() |>
  mutate(
    status = status |>
      case_match(
        1 ~ "relapse",
        0 ~ "censored"
      ),
    sex = sex |>
      case_match(
        0 ~ "female",
        1 ~ "male"
      ),
    rx = rx |>
      case_match(
        0 ~ "new",
        1 ~ "standard"
      ),
    surv = Surv(time = survt, event = (status == "relapse"))
  )

print(anderson)
```

1.5.1 Cox semi-parametric model

We fit a Cox proportional hazards model comparing standard therapy to new therapy:

```
anderson_cox0 <- coxph(
  formula = surv ~ rx,
  data = anderson
)
summary(anderson_cox0)
#> Call:
#> coxph(formula = surv ~ rx, data = anderson)
#>
#> n= 42, number of events= 30
#>
#>      coef exp(coef) se(coef)      z Pr(>|z|)
#> rxstandard 1.572    4.817    0.412 3.81 0.00014 ***
#> ---
#> Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
#>
#>      exp(coef) exp(-coef) lower .95 upper .95
#> rxstandard    4.82      0.208     2.15    10.8
#>
#> Concordance= 0.69 (se = 0.041 )
#> Likelihood ratio test= 16.4 on 1 df,  p=5e-05
```

```
#> Wald test          = 14.5   on 1 df,    p=1e-04
#> Score (logrank) test = 17.2   on 1 df,    p=3e-05
```

The estimated log hazard ratio $\hat{\beta}_{\text{PH}}$ for standard treatment relative to new treatment is positive, indicating that standard therapy is associated with a higher hazard of relapse (shorter remission duration).

1.5.2 Weibull parametric model

We next fit a parametric Weibull accelerated failure time model using `survreg()`:

```
anderson_weib <- survreg(
  formula = surv ~ rx,
  data = anderson,
  dist = "weibull"
)
summary(anderson_weib)
#>
#> Call:
#> survreg(formula = surv ~ rx, data = anderson, dist = "weibull")
#>              Value Std. Error      z      p
#> (Intercept)  3.516      0.252 13.96 < 2e-16
#> rxstandard  -1.267      0.311 -4.08 4.5e-05
#> Log(scale)  -0.312      0.147 -2.12  0.034
#>
#> Scale= 0.732
#>
#> Weibull distribution
#> Loglik(model)= -106.6   Loglik(intercept only)= -116.4
#>   Chisq= 19.65 on 1 degrees of freedom, p= 9.3e-06
#> Number of Newton-Raphson Iterations: 5
#> n= 42
```

In `survreg()`, regression parameters are reported on the log-time scale ($Y = \log T = \alpha + \tilde{\beta}_{\text{AFT}}\tilde{x} + \sigma W$). The estimated coefficient $\hat{\beta}_{\text{AFT}}$ for standard treatment is negative, indicating a reduction in log survival time (faster progression to relapse). The scale estimate $\hat{\sigma}$ corresponds to $1/p$, and satisfies $\hat{\beta}_{\text{AFT}} = -\hat{\sigma}\hat{\beta}_{\text{PH}}$.

1.5.3 Exponential parametric model

Setting the scale parameter to $\sigma = 1$ ($p = 1$) yields the exponential parametric model:

```
anderson_exp <- survreg(
  formula = surv ~ rx,
  data = anderson,
  dist = "exp"
)
summary(anderson_exp)
#>
#> Call:
#> survreg(formula = surv ~ rx, data = anderson, dist = "exp")
#>              Value Std. Error      z      p
#> (Intercept)  3.686      0.333 11.06 < 2e-16
#> rxstandard  -1.527      0.398 -3.83 0.00013
#>
#> Scale fixed at 1
#>
#> Exponential distribution
#> Loglik(model)= -108.5   Loglik(intercept only)= -116.8
```

```
#> Chisq= 16.49 on 1 degrees of freedom, p= 4.9e-05
#> Number of Newton-Raphson Iterations: 4
#> n= 42
```

In the exponential model, $\hat{\beta}_{\text{AFT}} = -\hat{\beta}_{\text{PH}}$, so the AFT coefficient is equal in magnitude and opposite in sign to the Cox proportional hazards coefficient.

1.5.4 Diagnostic: complementary log-log survival plot

To check whether a Weibull proportional hazards assumption is reasonable, we inspect the complementary log-log transformation of the Kaplan-Meier survival curves, $\log(-\log \hat{S}(t)) = p \log \lambda + p \log t$:

```
library(survminer)
survfit(
  formula = surv ~ rx,
  data = anderson
) |>
ggsurvplot(fun = "cloglog")
```

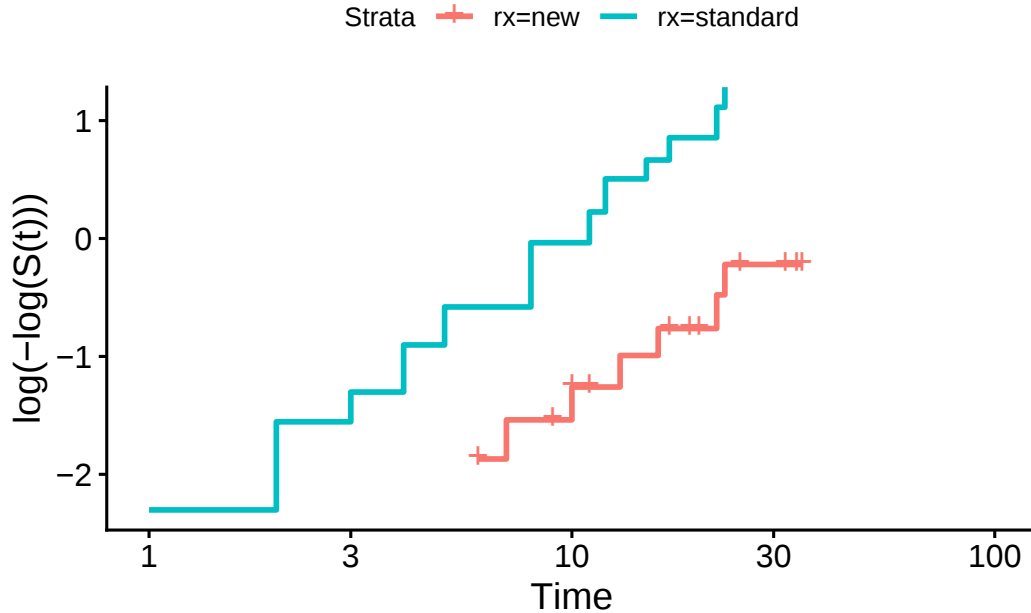


Figure 4: Complementary log-log survival curves by treatment group

If the complementary log-log curves are approximately linear and parallel across treatment groups, then a Weibull proportional hazards model is appropriate.

2 Combining left-truncation and interval-censoring

In observational studies and clinical trials, complex observation schemes can introduce both **left-truncation** (delayed entry) and **interval-censoring**. Left-truncation occurs when subjects enter the risk set only after surviving past an entry time $L_{\text{entry}} > 0$. Interval-censoring occurs when the exact event time T_i is unknown, but is known to fall within an interval $(L_i, R_i]$.

Standard software tools in R make distinct trade-offs between these two observation mechanisms:

`coxph` does left truncation but not left (or interval) censoring; `survreg` does interval censoring but not left truncation (or time dependent covariates).

In particular, `coxph()` constructs risk sets dynamically at each observed failure time t_j using counting process format (`time1 = entry`, `time2 = exit`), enabling seamless handling of left-truncation. However, because partial likelihood relies on ordered exact failure times, standard Cox models cannot easily accommodate interval-censored data.

Conversely, `survreg()` accommodates interval-censored outcomes via `Surv(time1 = L, time2 = R, type = "interval")` by maximizing the interval parametric likelihood $P(L_i < T_i \leq R_i) = F_0(R_i) - F_0(L_i)$. However, standard `survreg()` assumes all subjects enter observation at time 0, and does not condition likelihood terms on surviving past a delayed entry time L_{entry} .

When a study contains both delayed entry and interval-censored event times, specialized tooling is needed. Among the following R packages, only `flexsurv` handles both features together; the other two are listed because they cover interval censoring well and are the usual starting points, not because they solve the truncation half:

- **flexsurv**: Fits flexible parametric survival models and allows custom likelihood specifications that incorporate both left-truncation conditioning and interval-censored bounds.
- **icenReg**: Fits regression models for interval-censored data: Cox proportional-hazards, proportional-odds, and accelerated failure time. Semi-parametric and fully parametric forms are both available, but the fully parametric option covers the AFT models only. It does not itself provide left-truncation conditioning.
- **interval**: Fits nonparametric survival curves (NPMLE) for interval-censored data and provides weighted logrank and Wilcoxon-type tests (Fay and Shaw 2010). It is nonparametric throughout, so it supports neither regression nor truncation.

Fay, Michael P., and Pamela A. Shaw. 2010. “Exact and Asymptotic Weighted Logrank Tests for Interval Censored Data: The interval R Package.” *Journal of Statistical Software* 36 (2): 1–34. <https://doi.org/10.18637/jss.v036.i02>.

Kalbfleisch, John D, and Ross L Prentice. 2011. *The Statistical Analysis of Failure Time Data*. John Wiley & Sons.

¹<https://stat.ethz.ch/pipermail/r-help/2015-August/431733.html>