Comparing the Accelerated Failure Time and Porportional Hazards Models for the Weibull Distribution

Gavin Pu

2024-03-27

Introduction

The Weibull distribution is the only distribution that is a member of both the accelerated failure time (AFT) model family and the proportional hazards (PH) model family. A question that arises from this statement is what differences these two parametrizations of the Weibull distribution may show. AFT models are fully parametric, while PH models are usually semi-parametric under the Cox PH model approach in which the baseline hazard function is left unspecified.

In survival analysis, important functions of a time-to-event response include the survival function (the probability that a subject will survive past a given time) and cumulative hazard function (the accumulated instantaneous risk of experiencing failure for a subject at risk of failing at a given time). Other informative statistics related to time-to-event responses are quantiles, such as the median survival time.

To compare the AFT and Cox PH model formulations for the Weibull distribution, we will first simulate follow-up times by using a Weibull distribution to generate failure times and an exponential distribution to generate censoring times. We will then fit AFT and Cox PH models to the simulated data and compare their estimated survival functions, cumulative hazard functions, and quartiles with those from the original Weibull distribution and from the commonly used Kaplan-Meier (KM) estimate.

Data Generation

```r
library(survival)
```

To ensure reproducibility of the randomly generated data, we will specify a seed for the random number generator before conducting the simulation.

```r
set.seed(475)
```

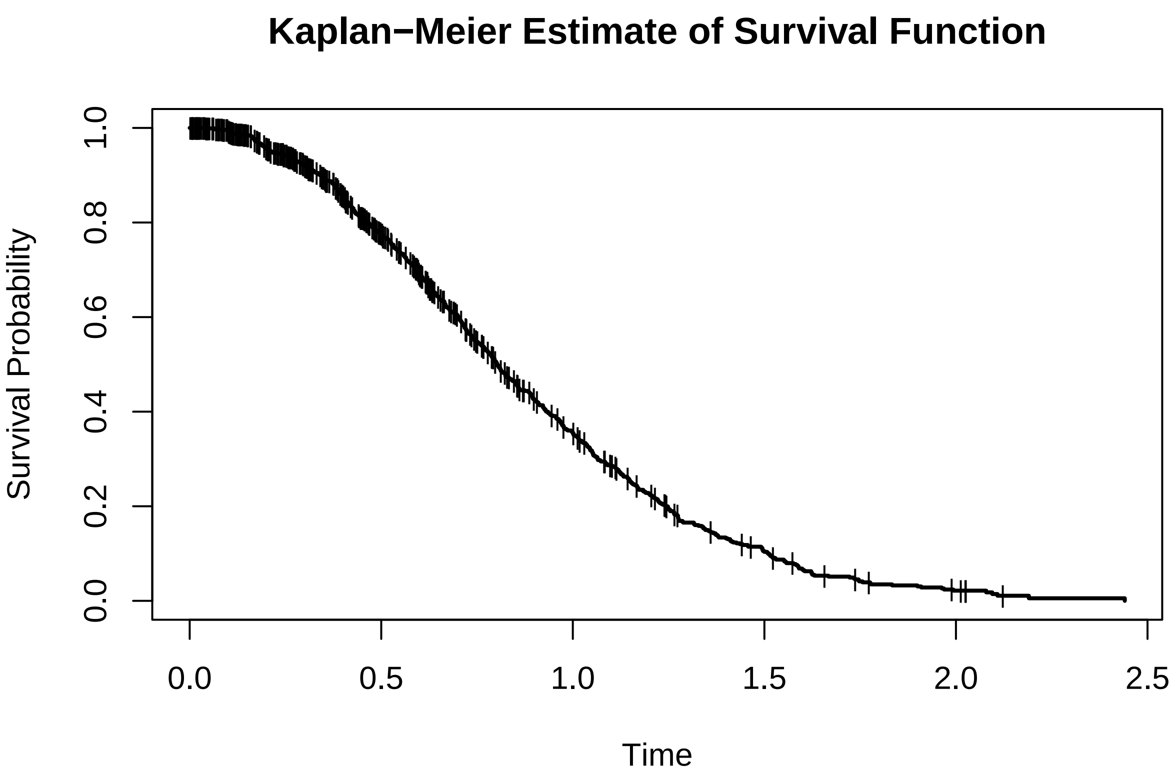
We will generate $n = 1000$ observations. For each subject $i \in \{1, \dots, 1000\}$, we will sample a failure time $T_i \sim \text{Weibull}(2, 1)$ and censoring time $C_i \sim \text{Exp}(0.5)$. The subject's follow-up time can be calculated as $X_i = \min\{T_i, C_i\}$. Each subject also has an indicator δ_i for their censoring status. If $\delta_i = 1$, the subject experienced the event of interest. Conversely, if $\delta_i = 0$, the subject was censored. Our final dataset will consist of only the follow-up times and status indicators.

```r
# Number of observations
n <- 1000

# Failure times
T <- rweibull(n, shape = 2, scale = 1)

# Censoring times
C <- rexp(n, rate = 0.5)
```

```r
# Follow-up times
X <- pmin(T, C)

# Indicators for whether the event was observed (status)
delta <- ifelse(T <= C, 1, 0)

# Data frame with follow-up times and statuses
data <- data.frame(time = X, status = delta)
```

As a preliminary assessment of these data, we will fit and visualize the KM estimate, a non-parametric estimator of the survival function. Censored subjects are marked with a vertical bar.

```r
# KM estimate
km_fit <- survfit(Surv(time, status) ~ 1, data = data)

# Plot of KM estimate
plot(km_fit, conf.int = FALSE, mark.time = TRUE, pch = "|", lwd = 2, cex = 0.75,
     main = "Kaplan-Meier Estimate of Survival Function", xlab = "Time",
     ylab = "Survival Probability")
```



Censoring appears to occur most frequently from $t = 0$ to $t = 0.5$. As t increases beyond this range, the frequency of censoring seems to gradually decrease.

Simulation

AFT Model

The survival function for an AFT model fitted on our simulated dataset will have the form

$$S(t) = \exp\left(-\exp\left(\frac{\log(t) - \beta_0}{\tau}\right)\right),$$

where β_0 and τ are the intercept and scale parameters respectively. Additionally, the cumulative hazard function will be

$$H(t) = \exp\left(\frac{\log(t) - \beta_0}{\tau}\right).$$

```r
# AFT model
model_aft <- survreg(Surv(time, status) ~ 1, data = data, dist = "weibull")
summary(model_aft)
```

```
##
## Call:
## survreg(formula = Surv(time, status) ~ 1, data = data, dist = "weibull")
##               Value Std. Error      z      p
## (Intercept) -0.0148     0.0190  -0.78   0.44
## Log(scale)  -0.7039     0.0283 -24.87 <2e-16
##
## Scale= 0.495
##
## Weibull distribution
## Loglik(model)= -473.9   Loglik(intercept only)= -473.9
## Number of Newton-Raphson Iterations: 8
## n= 1000
```

From the fitted model, the estimated parameters are $\hat{\beta}_0 = -0.0148$ and $\hat{\tau} = 0.495$. Using these estimates, we can derive the survival and cumulative hazard functions under the AFT model.

```r
# AFT parameters
beta0_aft <- model_aft$coefficients["(Intercept)"]
tau_aft <- model_aft$scale

# AFT survival function
S_aft <- function(t) {
  W_aft <- (log(t) - beta0_aft) / tau_aft
  return(exp(-exp(W_aft)))
}

# AFT cumulative hazard function
H_aft <- function(t) {
  W_aft <- (log(t) - beta0_aft) / tau_aft
  return(exp(W_aft))
}
```

Cox PH Model

The hazard function for a Cox PH model fitted on our simulated dataset will have the form

$$h(t) = h_0(t),$$

where $h_0(t)$ is the baseline hazard function that is typically left unspecified. For our simulation, however, we will obtain an estimate of $h_0(t)$ in order to derive the survival and cumulative hazard functions under the Cox PH model.

```r
# Cox PH model
model_coxph <- coxph(Surv(time, status) ~ 1, data = data)
summary(model_coxph)
```

```
## Call:  coxph(formula = Surv(time, status) ~ 1, data = data)
```

```
##
## Null model
##   log likelihood= -3981.555
##   n= 1000
```

For a Cox PH model, the `survfit` function can be used to find the survival and cumulative hazard functions directly.

```r
# Cox PH survival and cumulative hazard functions
coxph_fit <- survfit(model_coxph)
```

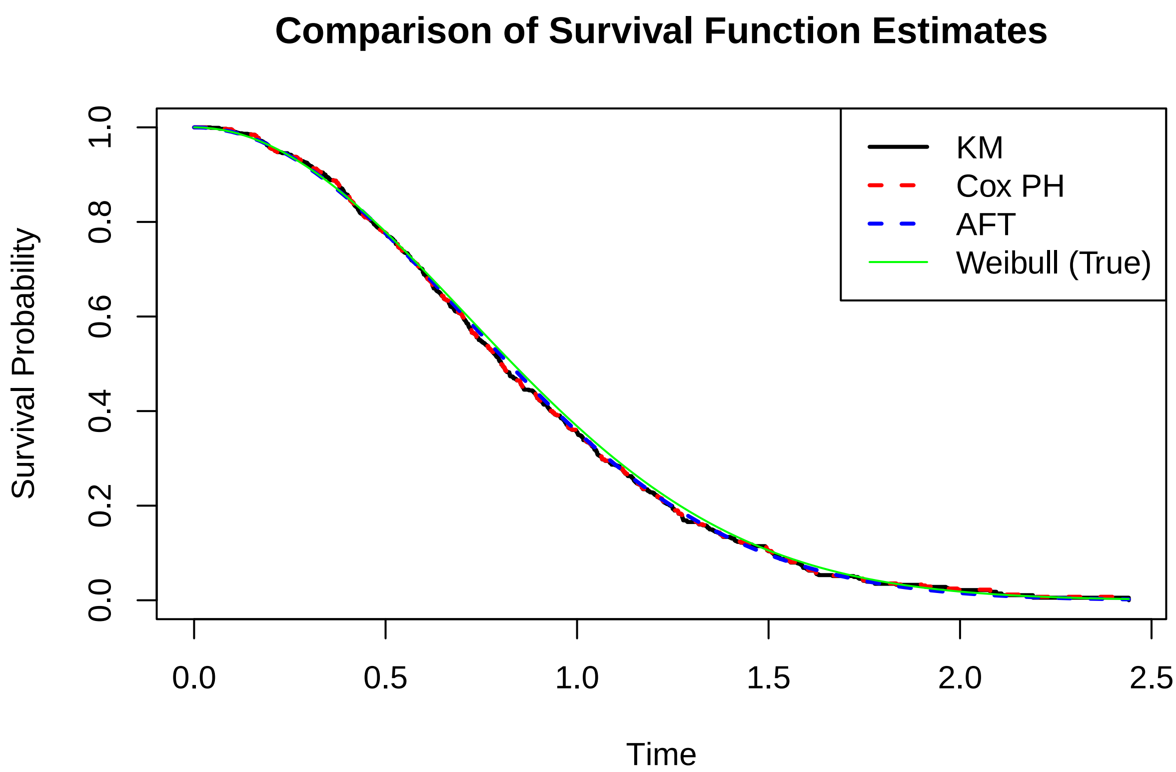
Model Comparison

As an additional benchmark on how well our models fit these data, we will calculate the true survival and cumulative hazard functions of the Weibull$(2, 1)$ distribution used to generate the failure times.

```r
# True survival function of Weibull distribution
S_weibull <- function(t) {
  return(1 - pweibull(t, shape = 2, scale = 1))
}

# True cumulative hazard function of Weibull distribution
H_weibull <- function(t) {
  return(-log(1 - pweibull(t, shape = 2, scale = 1)))
}
```

First, we will compare the survival functions of the AFT and Cox PH models with the KM estimate and the true survival function of the Weibull distribution.

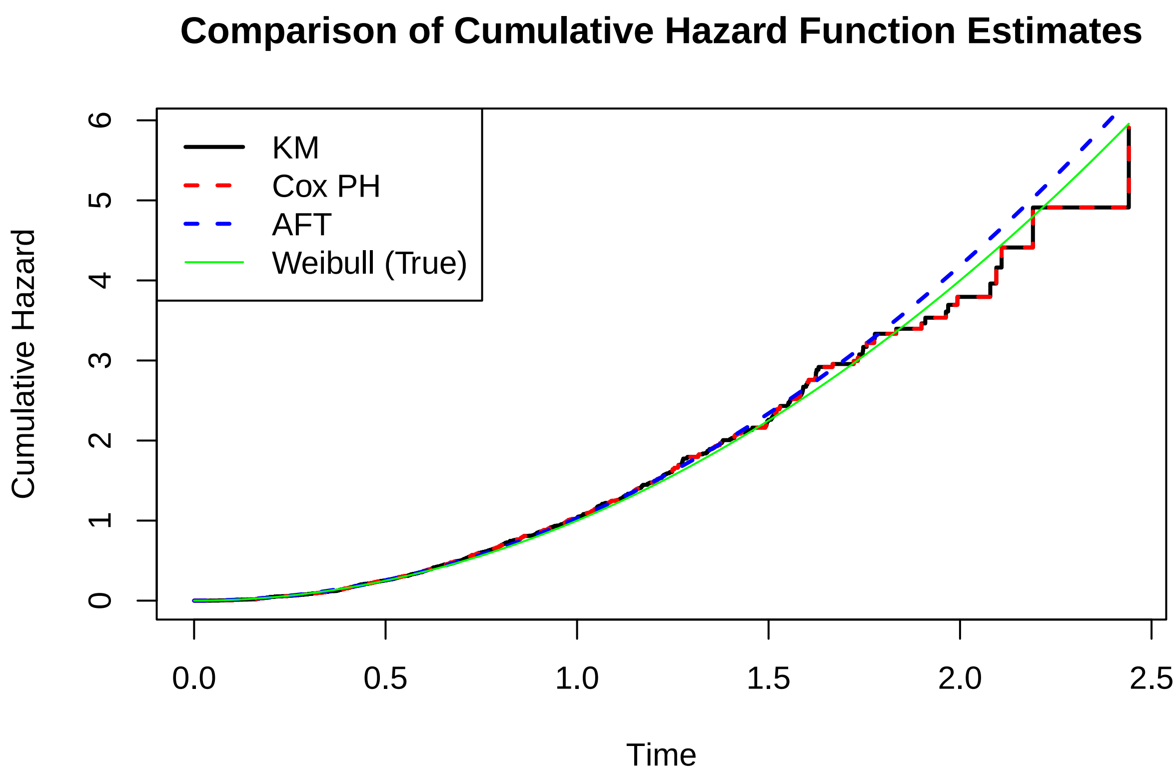
```r
# Plot of survival function estimates
plot(km_fit, conf.int = FALSE, lwd = 2,
     main = "Comparison of Survival Function Estimates",
     xlab = "Time", ylab = "Survival Probability")
lines(coxph_fit, col = "red", lty = 2, lwd = 2, conf.int = FALSE)
curve(S_aft, add = TRUE, col = "blue", lty = 2, lwd = 2)
curve(S_weibull, add = TRUE, col = "green")
legend("topright", legend = c("KM", "Cox PH", "AFT", "Weibull (True)"),
       col = c("black", "red", "blue", "green"), lty = c(1, 2, 2, 1),
       lwd = c(2, 2, 2, 1))
```



All four survival functions appear to be extremely similar. In particular, the KM estimate and the survival function derived from the Cox PH model are almost exactly the same. Near $t = 1$ is when the true survival function of the Weibull distribution differs the most from the three other survival function estimates.

Next, we will compare the cumulative hazard functions derived using the same four methods.

```
# Plot of cumulative hazard function estimates
plot(km_fit, conf.int = FALSE, lwd = 2,
     main = "Comparison of Cumulative Hazard Function Estimates",
     xlab = "Time", ylab = "Cumulative Hazard", cumhaz = TRUE)
lines(coxph_fit, col = "red", lty = 2, lwd = 2, conf.int = FALSE, cumhaz = TRUE)
curve(H_aft, add = TRUE, col = "blue", lty = 2, lwd = 2)
curve(H_weibull, add = TRUE, col = "green")
legend("topleft", legend = c("KM", "Cox PH", "AFT", "Weibull (True)"),
       col = c("black", "red", "blue", "green"), lty = c(1, 2, 2, 1),
       lwd = c(2, 2, 2, 1))
```



The four cumulative hazard functions match each other closely from $t = 0$ to $t = 0.5$. Around $t = 1$, the true cumulative hazard function appears to diverge slightly from the three estimated cumulative hazard functions. Beyound $t = 1.5$, the cumulative hazard functions exhibit noticeable differences from each other, except for the cumulative hazard functions derived from the Cox PH model and the KM estimate which are almost the same at all times.

Finally, we will calculate the first quartile (Q1), median, and third quartile (Q3) using each of the four methods.

```
# Data frame of quartiles for each model
quartiles <- t(data.frame(
  km = quantile(km_fit, probs = c(0.25, 0.5, 0.75), conf.int = FALSE),
  coxph = quantile(coxph_fit, probs = c(0.25, 0.5, 0.75), conf.int = FALSE),
  aft = qsurvreg(c(0.25, 0.5, 0.75), mean = beta0_aft, scale = tau_aft,
                 distribution = "weibull"),
  weibull = qweibull(c(0.25, 0.5, 0.75), shape = 2, scale = 1)
))
rownames(quartiles) <- c("KM", "Cox PH", "AFT", "Weibull (True)")

# Table of quartiles for each model
knitr::kable(quartiles, col.names = c("Method", "Q1", "Median", "Q3"),
             caption = "Comparison of Estimated Quartiles")
```

Table 1: Comparison of Estimated Quartiles

Method	Q1	Median	Q3
KM	0.5325383	0.8020367	1.152643
Cox PH	0.5325383	0.8021095	1.152643
AFT	0.5320066	0.8219400	1.158117
Weibull (True)	0.5363600	0.8325546	1.177410

The quartiles of the true Weibull distribution are slightly greater than the quartiles estimated by the other three methods. Out of the four quartile estimates found, those derived from the KM estimate and the Cox PH model are the most similar to each other.

Discussion

Even with many censored observations, the AFT and Cox PH survival functions resembled the true Weibull survival function closely. The estimates of the cumulative hazard function, however, showed more differences from the true Weibull cumulative hazard function, especially during later times. As the plot of the KM estimate shows, this may be because there were very few observations sampled with a follow-up time past $t = 2$. The quartiles of the AFT and Cox PH models were also similar to each other and to the quartiles from the underlying Weibull distribution of the failure times.

The Cox PH model and the KM estimate were almost identical, which can be explained by how our simulated data contained no covariates. The Cox PH model is normally a semi-parametric model where the covariates and their corresponding coefficients form the parametric component of the model. Due to the absence of covariates, the Cox PH model used to fit these data only contained a non-parametric component, namely the baseline hazard function which is normally left unspecified. The Cox PH model effectively became a non-parametric model for our simulated data. This could explain why it is extremely similar to the KM estimate, which is also non-parametric.

In conclusion, there do not appear to be very significant differences between the AFT and Cox PH models fitted to these data; it is difficult to say which of the two models is necessarily a better fit. In practice, choosing a model for time-to-event data under a Weibull distribution also depends on other factors, such as whether a fully parametric or non-parametric model is preferred. Additionally, the two models have different assumptions that may be satisfied for different datasets. The AFT model assumes that covariates have a multiplicative effect on survival time, while the Cox PH model assumes that covariates have a constant effect over time. Nonetheless, our generated data appear to be modeled well by both the AFT model and the Cox PH model.