This tutorial explains the basic concept of parametric time to event (TTE) models, focusing on commonly used exponential, Weibull, and log-logistic model. TTE data is commonly used as endpoint for treatment effect of a drug or prognosis of diseases. Although non-parametric Kaplan-Meier analysis has been widely used for TTE data analysis, parametric modeling analysis has its own advantages such as ease of simulation, and evaluation of continuous covariate. Accelerated failure time model is introduced as a covariate model for TTE data together with proportional hazard model. Compared to proportional hazard model, accelerated failure time model provides more intuitive results on covariate effect since it states that covariates change TTE whereas in proportional hazard model covariates affect hazard.

**Keywords:** Parametric Time to Event Model; Accelerated Failure Time Model; Proportional Hazard Model

**INTRODUCTION**

Time to event (TTE) data, known generically as survival data in statistics are widely used data type to describe prognosis of diseases or drug effect such as overall survival, progression free survival, time to disease progression, and time to vomit so on. TTE data are unique in that they have time information regarding when an event occurred as well as whether the event occurred or not.

The Kaplan–Meier analysis, is a non-parametric method for TTE data [1]. Survival proportion at each time point is estimated by product-limit method, and confidence interval in the estimate of proportion is calculated by approximating binomial distribution with a normal distribution. Although it is the most commonly used TTE analysis methods with many advantages, it is limited in that it is not adequate for simulation, and statistical significance of continuous covariate such as time varying covariates per se cannot be tested.

This tutorial reviewed essential concept to understand parametric TTE models focusing on exponential, Weibull, and log-logistic models which as far as I know, are the most commonly used TTE models. Furthermore, accelerated failure time model, a covariate model for TTE model is introduced as well as conventional proportional hazard model.
BASIC FUNCTIONS FOR TTE MODELING

There are essential functions to understand TTE data modeling, which are probability function \( f(t) \), survival function \( S(t) \), and hazard function \( h(t) \) as described below.

Probability density is simply a probability that an event occurs at each time, and it can be described as a mathematical formula if the change of probability density over time is assumed to follow certain pattern. Probability density function can be described by multiplying hazard and survival function as following equation.

\[
\text{Probability density: } f(t) = (h(t)) \times (S(t)) \tag{1}
\]

where \( t \) is a particular time, \( h(t) \) and \( S(t) \) are hazard and survival function as described below, respectively.

Survival function \( S(t) \) is defined as \( 1 – \text{cumulative density of an event} \). Cumulative density function is simply an integral of probability density over time and means cumulative probability that an event occurs until time \( t \). Therefore, \( 1 - F(t) \) can be interpreted as the probability that an event has not yet occurred until time \( t \).

\[
\text{Survival: } S(t) = 1 - F(t) = e^{-H(t)} = e^{-h(t)} = P(T \geq t) \tag{2}
\]

where \( F(t) = \int_0^t f(u) du = P(T < t) = 1 - S(t) \); \( H(t) \) is cumulative hazard, expressed as \( \int_0^t h(u) du \); \( T \) is a random variable denoting the time of event, therefore \( P(T \geq t) \) can be interpreted as cumulative probability that an event occurs at time, \( t \) or thereafter.

Hazard is a conditional probability that an event will occur in a small interval \( ([t, t+\Delta t]) \) given that it has not occurred before. Taking \( \Delta t \) to the limit zero, an instantaneous rate of occurrence at each particular time, \( t \) is obtained as described in equation 3. Hazard function can be expressed by \( f(t) \) divided by \( S(t) \).

\[
\text{Hazard: } h(t) = \lim_{\Delta t \to 0} \frac{P(t \leq T < t + \Delta t | T \geq t)}{\Delta t} = \frac{f(t)}{S(t)} \tag{3}
\]

From the relationship among \( f(t), S(t) \) and \( h(t) \), by specifying one of the three functions the other two functions can be also specified.

TTE MODELS

Exponential model
Exponential model describes the survival \( S(t) \) over time as exponential decay model, similar to the concentration over time after intravenous bolus injection of a drug following mono-exponential decay. Only one parameter, \( \lambda \) is required, which determine the degree of decrease in \( S(t) \) over time (Eq. 4).

\[
S(t) = e^{-\lambda t} \tag{4}
\]
From the relationships among the functions, \( f(t) \), and \( h(t) \) can be derived as follows.

\[
\begin{align*}
  h(t) &= \lambda \\
  f(t) &= \lambda e^{-\lambda t}
\end{align*}
\]  

(5)  
(6)

It is evident from the above equation that hazard is constant in exponential survival model. Using the inverse function of equation 4, time to any survival probability value (\( t(S) \)) can be calculated, as well as time to median (50%) survival (\( t(50) \)).

\[
\begin{align*}
  t(S) &= \frac{1}{\lambda} \times \log \left( \frac{1}{S} \right) \\
  t(50) &= -\frac{1}{\lambda} \times \log 2
\end{align*}
\]

(7)  
(8)

**Weibull model**

Weibull model is defined as equation 9, providing more flexibility in \( S(t) \) over time plot than exponential model with an additional parameter (\( \gamma \)). Weibull model is reduced to exponential model when \( \gamma = 1 \).

\[ S(t) = e^{-\lambda t^\gamma} \]

(9)

From the \( S(t), f(t) \), and \( h(t) \) are derived as follows.

\[
\begin{align*}
  h(t) &= \lambda t^\gamma \gamma^{-1} \\
  f(t) &= \lambda t^\gamma e^{-\lambda t^\gamma}
\end{align*}
\]

(10)  
(11)

Constant hazard model assumed in exponential model rarely tenable in reality. As can be guessed in equation 10, hazard changes over time in Weibull model, but either in monotonical increase or decrease.

Time to any survival probability value (\( t(S) \)), and time to median survival (\( t(50) \)) are derived as follows.

\[
\begin{align*}
  t(S) &= \left( \frac{1}{\lambda} \times \log \left( \frac{1}{S} \right) \right)^{\gamma^{-1}} \\
  t(50) &= \left( \frac{1}{\lambda} \times \log 2 \right)^{\gamma^{-1}}
\end{align*}
\]

(12)  
(13)

**Log-logistic model**

Log-logistic model is defined as equation 14 with two parameters. Log-logistic model is structurally distinctive one from Weibull model, and there is no full-nested model relationship between log-logistic and exponential model.

\[ S(t) = \left[ 1 + \left( \frac{t}{\lambda} \right)^\gamma \right]^{-1} \]

(14)

From the \( S(t), f(t), \) and \( h(t) \) for log-logistic model are derived as follows.

\[ h(t) = \frac{\left( \frac{t}{\lambda} \right) \gamma \gamma^{-1}}{1 + \left( \frac{t}{\lambda} \right)^\gamma} \]

(15)
From equation 15, hazard in log-logistic model also change over time as in the case of Weibull model. However, hazard in log-logistic model can also describe non-monotonical changes such as initial increase in hazard then decrease, and vice versa, whereas hazard change over time in Weibull model is limited to monotonical ones.

Time to a particular survival probability value \( t(S) \), and time to median survival \( t(50) \) in log-logistic model are derived as follows.

\[
t(S) = \lambda \left( \frac{1 - S}{S} \right)^{1/\gamma} \\
t(50) = \lambda
\]

Important functions for exponential, Weibull, and log-logistic models are summarized in Table 1.

### COVARIATE MODELS

#### Proportional hazard model

Proportional hazard model assumes that covariates affect the hazard in the form of multiplication. Equation 19 states that the hazard with covariate \( X_i (h_i(t|X_i)) \) is equal to baseline hazard \( h_0(t) \) multiplied by covariate effect, \( e^{(X_i^* \times \beta_i)} \), where \( X_i^* \) are covariates and \( \beta_i \) are parameters relating the covariate and hazard. Exponentiation of covariates multiplied by their relating parameters keeps the values positive. From equation 19, the relationship between survival function \( S_i(t|X_i) \) with covariate and baseline survival \( S_0(t) \) is derived as equation 20.

\[
h_i(t|X_i) = h_0(t) \times e^{(X_i^* \times \beta_i)} \\
S_i(t|X_i) = S_0(t) \times e^{(X_i^* \times \beta_i)}
\]
where \( X'_i \times \beta_i = \beta_1 \times X_{1i} + \beta_2 \times X_{2i} + \beta_3 \times X_{3i} + \ldots \times X \ldots \) are covariates, and \( \beta_1, \beta_2 \) are parameters relating the respective covariate to hazard.

**Accelerated failure time model**

Unlike proportional hazard model, accelerated failure time model states that covariates affect the time (Eq. 21). For example, the lifetime of an individual on the new treatment is described as \( e^{(X'_i \times \beta_i)} \) times the lifetime that the individual would have experienced under the standard treatment. \( \frac{1}{e^{(X'_i \times \beta_i)}} \) is acceleration factor (AF), which is a ratio of time-quantiles corresponding to any fixed value of \( S(t) \) AF allows to evaluate the effect of predictor variables on the survival time, which is more intuitive. When death is the endpoint of interest, AF > 1 corresponds to acceleration in time to death, indicating earlier death (worse prognosis) in patients with new treatment compared to standard treatment, whereas AF < 1 indicates more prolonged survival in those with new treatment. From equation 21, relationship between survival with covariate and baseline survival (Eq. 22), and one between hazard with covariate and baseline hazard (Eq. 23) are derived.

Unlike proportional hazard model, the regression parameter estimates from accelerated failure time models are known to be robust to omitted covariates. They are also less affected by the choice of probability distribution [2-4]

\[
T_i = T_{0i} \times e^{(X'_i \times \beta_i)} \tag{21}
\]

\[
S_i(t|X_i) = S_0(\frac{t}{e^{(X'_i \times \beta_i)}}) \tag{22}
\]

\[
h_i(t|X_i) = e^{(X'_i \times \beta_i)} \times h_0(\frac{t}{e^{(X'_i \times \beta_i)}}) \tag{23}
\]

**REFERENCES**

1. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. J Am Stat Assoc 1958;53:457-481.

2. Patel K, Kay R, Rowell L. Comparing proportional hazards and accelerated failure time models: an application in influenza. Pharm Stat 2006;5:213-224.

3. Keiding N, Andersen PK, Klein JP. The role of frailty models and accelerated failure time models in describing heterogeneity due to omitted covariates. Stat Med 1997;16:215-224.

4. Bradburn MJ, Clark TG, Love SB, Altman DG. Survival analysis part II: multivariate data analysis--an introduction to concepts and methods. Br J Cancer 2003;89:431-436.

https://tcpharm.org

https://doi.org/10.12793/tcp.2021.29.e7