Previously, we have focused on multiplicative intensity models, where
\[ h(t \mid z) = h_0(t) \cdot g(z). \]

These can also be expressed as
\[
\begin{align*}
H(t \mid z) &= H_0(t) \cdot g(z) \\
\text{or } S(t \mid z) &= \exp(-H(t)g(z)) = \left(\exp(-H_0(t))\right)^{g_0(t)} \\
\text{or } S(t \mid z) &= (S_0(t))^{g(z)}.
\end{align*}
\]

Consider instead the model where, for the 2-sample problem,
\[
\begin{align*}
\text{patients in group 0} & : \sim T_0 \\
\text{patients in group 1} & : \sim T_1 \\
\text{and } [\mathcal{L}(T_1) = \mathcal{L}(\phi T_0) \quad (\phi > 0)].
\end{align*}
\]

That is, survival time in group 1 is distributed as $\phi T_0$, where survival time in group 0 is distributed as $T_0$.

Let $z = 0, 1$ denote group.

Then
\[
S_1(t) = P(T_1 > t) = P(\phi T_0 > t) = P(T_0 > t/\phi) = S_0(t/\phi).
\]

For simplicity, let $\psi = 1/\phi$. Then the accelerated failure time model for the 2-sample problem can be defined by any of the following 3 equations:
\[
\begin{align*}
S_1(t) &= S_0(\psi t) \\
\text{or } f_1(t) &= \psi f_0(\psi t) \\
\text{or } h_1(t) &= \psi h_0(\psi t)
\end{align*}
\]
Let’s generalize to a vector of covariates $z$. Let $g(z)$ be some non-negative function such that $g(0) = 1$.

Then we can generalize (27.1) to

$$S(t \mid z) = S_0(t \cdot g(z))$$  \hspace{1cm} (27.2)

e.g. $z = \text{scalar and binary}$

$$S(t \mid z = 0) = S_0(t \cdot 1) = S_0(t)$$

and

$$S(t \mid z = 1) = S_0(t \cdot g(1)) = S_0(\psi t)$$

where $\psi = g(1)$.

$\rightarrow$ This is just (27.1).

Note that (27.2) corresponds to

$$f(t \mid z) = g(z) \cdot f_0(t \cdot g(z))$$

or

$$h(t \mid z) = g(z) \cdot h_0(t \cdot g(z)).$$

**Special Case:** Suppose $g(z) = e^{\beta z}$.

Then, for example, $h(t \mid z) = e^{\beta z}h_0(te^{\beta z})$.

Let’s express this in terms of random variables:

Suppose $T_0 \sim S_0(t)$

Then if $T \overset{\text{def}}{=} \frac{T_0}{g(z)}$,  \hspace{1cm} (27.3)

$$S(t \mid z) \overset{\text{def}}{=} P(T > t \mid z) = P[T_0 > t \cdot g(z)] = S_0(t \cdot g(z)),$$

which is what (27.2) specifies. This also follows if for given $Z = z$,

$$T \sim \frac{T_0}{g(z)}.$$  \hspace{1cm} (27.4)

Thus, in terms of r.v.’s we can specify an AFT model as
\[ \ln T = \ln T_0 - \ln (g(z)) \]  
(27.5)

or, less stringently, given \( Z = z \),

\[ \ln T \sim \ln T_0 - \ln (g(z)) ; \]  
(27.6)

equivalently, given \( Z = z \),

\[ \ln T + \ln (g(z)) \sim \ln T_0. \]  
(27.7)

Let \( \mu_0 \defeq E(\ln T_0) \) and \( \epsilon \defeq \ln T_0 - \mu_0 \).

Then (27.5) is equivalent to

or
\[
\begin{align*}
\ln T &= \ln T_0 - \ln (g(z)) = -\ln (g(z)) + \mu_0 + \ln T_0 - \mu_0 \\
\ln T &= \mu_0 - \ln g(z) + \epsilon,
\end{align*}
\]  
(27.8)

where the random variable \( \epsilon \) has mean 0 and a distribution that’s independent of \( z \).

(27.6) is equivalent to, for given \( Z = z \),

\[ \ln(T) \sim \mu_0 - \ln(g(z)) + \epsilon \]

with \( \epsilon = \ln(T_0) - \mu_0 \). Now define

\[ \tilde{\epsilon} = \ln(T) - \mu_0 + \ln(g(z)). \]  
(27.9)

Then

\[ \ln(T) = \mu_0 - \ln(g(z)) + \tilde{\epsilon}. \]

Also in this more general case, for given \( Z = z \),

\[ \tilde{\epsilon} = \ln(T) - \mu_0 + \ln(g(z)) \sim \ln(T_0) - \ln(g(z)) - \mu_0 + \ln(g(z)) = \ln(T_0) - \mu_0, \]

a mean zero random variable whose distribution does not depend on \( Z = z \).

**Special Case:** \( g(z) = e^{\beta z} \). Then

\[ \ln T = \mu_0 - \beta z + \tilde{\epsilon}, \quad \tilde{\epsilon} \sim (0, \text{some variance}). \]
Connection with PH Models

PH Model: \( S(t \mid z) = (S_0(t))^{g_1(z)} \) for some \( g_1(\cdot) \)

AFT Model: \( S(t \mid z) = S_0(g_2(z) \cdot t) \cdot g_2(z) \).

These in general are different models. They coincide iff \( T_0 \sim \text{Weibull} \) (cf: Cox & Oakes, 1984).

Parametric Analysis - if no censoring

\[
\ln T_i = \mu - \beta z_i + \tilde{\epsilon}_i \quad i = 1, 2, \ldots, n
\]

\( \tilde{\epsilon}_i \text{ iid } \sim (0, \sigma^2) \)

some distribution

Least Squares: find \((\hat{\mu}, \hat{\beta})\) such that

\[
\sum_{i=1}^n \left( \ln T_i - \hat{\mu} - \hat{\beta} z_i \right)^2 \text{ is minimized.}
\]

\( \longrightarrow \) no specific parametric assumptions needed

\( \longrightarrow \) efficient if \( \ln T_i \sim \text{normal} \)

(or \( T_i \sim \text{lognormal} \))

\( \longrightarrow \) not efficient if \( \ln T_i \) is not normal

e.g. ARE = .61 if \( T_i \sim \text{Exp} \) (see Cox & Oakes, 1984).

Parametric Analysis - with censoring

\( \leftrightarrow \) see Buckley, J & James, I, 1979, Biometrika, Vol. 66, p 429–36.

Semiparametric Inference

Suppose \( \ln T = \mu_0 - \beta^* z + \epsilon^* \)

\( \epsilon^* \sim (0, \text{ some variance}) \)

Let
\[ \epsilon = \epsilon^* + \mu_0 \]
\[ \beta = -\beta^* \]
so that we can also write
\[ \ln T = \beta z + \epsilon, \quad \epsilon \sim S_0(\cdot) \quad \{ \text{mean not zero} \} \]

Let \( C_1, C_2, \ldots, C_n \) be potential censoring times (independent and independent of the \( T_i \) given \( Z_i \)).

Observe \((U_i, \delta_i, Z_i)\) \( i = 1, 2, \ldots, n \)
\[
\begin{align*}
U_i &= \min (T_i, C_i) \\
\delta_i &= 1 \left( T_i \leq C_i \right)
\end{align*}
\]
usual setting.

Define \( e_i(\beta) = \ln U_i - \beta Z_i \) and note that \( \min (\ln T_i - \beta Z_i, \ln C_i - \beta Z_i) = \min (e_i, d_i) \) where \( d_i = \ln C_i - \beta Z_i \).

Also, \( \epsilon_i \leq d_i \Leftrightarrow \ln T_i \leq \ln C_i \Leftrightarrow T_i \leq C_i \). Thus, \( 1_{\{e_i \leq d_i\}} \equiv \delta_i \).

Thus, consider the “data”
\[
(U_i^*, \delta_i) \quad i = 1, 2, \ldots, n
\]
where
\[ U_i^* = \min (\epsilon_i, d_i) \quad \text{Note } \epsilon_i \perp d_i | Z_i. \]

Thus
\[
\begin{align*}
U_i^* &= \min (\ln T_i, \ln C_i) - \beta Z_i \\
&= \ln (\min (T_i, C_i)) - \beta Z_i \\
&= \ln \left( U_i \right) - \beta Z_i = e_i(\beta).
\end{align*}
\]

Thus, we’ve created “data” where the underlying survival time \( \epsilon_i \) doesn’t depend on \( Z_i \)! (more later)

Recall the PH model and Cox’s PL score function
\[
n^{-1/2} U(\beta^*) = \cdots = n^{-1/2} \sum_{i=1}^{n} \int_0^{\infty} \left( Z_i - \frac{\sum_{j=1}^{n} Y_j(u) Z_j e^{\beta^* Z_j}}{\sum Y_j(u) e^{\beta^* Z_j}} \right) dN_i(u). \]
If there is no covariate effect (i.e., $\beta^* = 0$), $n^{-1/2}U(\beta^*)$ reduces to

$$n^{-1/2} \sum_i \int_0^\infty \left( Z_i - \frac{\sum_j Y_j(u) Z_j}{\sum_j Y_j(u)} \right) dN_i(u).$$

Also, we know $n^{-1/2}U(0) \xrightarrow{\mathcal{L}} N(0, \text{some variance})$ when $\beta^* = 0$.

Let’s consider this statistic and apply it to the “data” $(U_i^*, \delta_i)$. Note: for the true $\beta$, the underlying “survival times” $\epsilon_i$ have a distribution that does not depend on $Z_i$, and therefore follow a Cox proportional hazards model with parameter $\beta^* = 0$.

$$S(\beta) \overset{\text{def}}{=} n^{-1/2} \sum_i \int_0^\infty \left( Z_i - \frac{\sum_j Z_j \mathbb{1}[e_j(\beta) \geq u]}{\sum_j \mathbb{1}[e_j(\beta) \geq u]} \right) d1 \left[ e_i(\beta) \leq u \text{ and } \delta_i = 1 \right]$$

or

$$S(\beta) = n^{-1/2} \sum_{i=1}^n \delta_i \left( Z_i - \frac{\sum_j Z_j \mathbb{1}[e_j(\beta) \geq e_i(\beta)]}{\sum_j \mathbb{1}[e_j(\beta) \geq e_i(\beta)]} \right).$$

Note

$$\epsilon_j(\beta) \geq \epsilon_i(\beta) \iff \ln U_j - \beta Z_j \geq \ln U_i - \beta Z_i \iff \ln \left( \frac{U_j}{U_i} \right) \geq \beta (Z_j - Z_i).$$

Thus, in summary

$$S(\beta) \xrightarrow{\mathcal{L}} N(0, \text{some var}).$$

Then use this as an estimating equation for $\beta$, i.e. find $\hat{\beta}$ s.t. $S(\hat{\beta}) = 0$.

Then, it can be shown that

$$\sqrt{n} \left( \hat{\beta} - \beta \right) \xrightarrow{\mathcal{L}} N(0, \text{some var}).$$

General References to AFT Models


**Exercises**

1. Give an example of an accelerated failure time model involving 2 covariates: $Z_1=$ treatment group, and $Z_2=$ age. Use $T$ to denote survival time.