STAT 331

Accelerated Failure Time Models

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

$$\begin{array}{rrcl} & H(t \mid z) & = & H_0(t) \cdot g(z) \\ \text{or} & S(t \mid z) & = & e^{-H(t|z)} = \left(e^{-H_0(t)}\right)^{g_0(t)} \\ \text{or} & S(t \mid z) & = & (S_0(t))^{g(z)}. \end{array}$$

Consider instead the model where, for the 2-sample problem,

$$\begin{array}{rcl} \text{patients in group 0} & : & \sim T_0 \\ \text{patients in group 1} & : & \sim T_1 \end{array}$$

$$\text{and} \quad \boxed{\mathcal{L}(T_1) = \mathcal{L}(\varphi T_0)} \quad (\varphi > 0).$$

That is, survival time in group 1 is distributed as φT_0, where survival time in group 0 is distributed as T_0.

Let $z = 0, 1$ denote group.

Then

$$\begin{aligned} S_1(t) = P(T_1 > t) &= P(\varphi T_0 > t) \\ &= P(T_0 > t/\varphi). \\ &= S_0(t/\phi). \end{aligned}$$

For simplicity, let $\psi = 1/\varphi$. Then the accelerated failure time model for the 2-sample problem can be defined by any of the following 3 equations:

$$\left.\begin{array}{rrcl} & S_1(t) & = & S_0(\psi t) \\ \text{or} & f_1(t) & = & \psi f_0(\psi t) \\ \text{or} & h_1(t) & = & \psi h_0(\psi t) \end{array}\right\}. \tag{27.1}$$

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

$$S(t \mid z) = S_0(t \cdot g(z)) \tag{27.2}$$

e.g. z = scalar and binary

$$\begin{array}{rrcll} & S(t \mid z=0) & = & S_0(t \cdot 1) & = S_0(t) \\ \text{and} & S(t \mid z=1) & = & S_0(t \cdot g(1)) & = S_0(\psi t) \\ & & & \text{where } \psi & = g(1). \end{array}$$

$\longrightarrow$ This is just (27.1).

Note that (27.2) corresponds to

$$\text{or} \begin{array}{rcl} f(t \mid z) & = & g(z) \cdot f_0(t \cdot g(z)) \\ h(t \mid z) & = & g(z) \cdot h_0(t \cdot g(z)). \end{array}$$

<u>Special Case:</u> Suppose $g(z) = e^{\beta z}$.

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

Let's express this in terms of random variables:

Suppose $T_0 \sim S_0(t)$

Then if $T \stackrel{\text{def}}{=} \frac{T_0}{g(z)}$, (27.3)

$$\begin{aligned} S(t \mid z) \stackrel{\text{def}}{=} P(T > t \mid z) &= P[T_0 > t \cdot g(z)] \\ &= S_0(t \cdot g(z)), \end{aligned}$$

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

$$T \sim \frac{T_0}{g(z)}. \tag{27.4}$$

Thus, in terms of r.v.'s we can specify an AFT model as

$$\ln T = \ln T_0 - \ln\left(g\left(z\right)\right) \tag{27.5}$$

or, less stringently, given $Z = z$,

$$\ln T \sim \ln T_0 - \ln\left(g\left(z\right)\right); \tag{27.6}$$

equivalently, given $Z = z$,

$$\ln T + \ln\left(g\left(z\right)\right) \sim \ln T_0. \tag{27.7}$$

Let $\mu_0 \stackrel{\text{def}}{=} E\left(ln\ T_0\right)$ and $\epsilon \stackrel{\text{def}}{=} \ln T_0 - \mu_0$.

Then (27.5) is equivalent to

$$\begin{aligned} \ln T &= \ln T_0 - \ln\left(g\left(z\right)\right) = -\ln\left(g\left(z\right)\right) + \mu_0 + \ln T_0 - \mu_0 \\ \text{or} \quad \ln T &= \mu_0 - \ln g\left(z\right) + \epsilon, \end{aligned} \tag{27.8}$$

where the random variable ϵ has mean 0 and a distibution 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

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

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 $Z = z$.

<u>Special Case:</u> - $g\left(z\right) = e^{\beta z}$. Then

$$\ln T = \mu_0 - \beta z + \tilde{\epsilon}, \qquad \tilde{\epsilon} \sim \left(0, \text{some variance}\right).$$

Connection with PH Models

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

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

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

Parametric Analysis - if no censoring

$$\ln T_i = \mu - \beta z_i + \tilde{\epsilon}_i \qquad i = 1, 2, \ldots, n$$
$$\tilde{\epsilon}_i \text{ iid } \sim \underbrace{\left(0, \sigma^2\right)}_{\text{some distribution}}$$

Least Squares: find $\left(\widehat{\mu}, \widehat{\beta}\right)$ such that

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

$\longrightarrow$ no specific parametric assumptions needed

$\longrightarrow$ efficient if $\ln T_i \sim$ normal
(or $T_i \sim$ lognormal)

$\longrightarrow$ not efficient if $\ln T_i$ is not normal
e.g. ARE = .61 if $T_i \sim$ Exp (see Cox & Oakes, 1984).

Parametric Analysis - with censoring
$\hookrightarrow$ see Buckley, J & James, I, 1979, Biometrika, Vol. 66, p 429–36.

Semiparametric Inference

Suppose $\ln T = \mu_0 - \beta^* z + \epsilon^*$
$\epsilon^* \sim (0,\ \textit{some variance})$

Let

$$
\begin{aligned}
\epsilon &= \epsilon^* + \mu_0 \\
\beta &= -\beta^* \qquad\qquad \text{so that we can also write} \\
\ln T &= \beta z + \epsilon, \quad \epsilon \sim S_0(\cdot) \quad \{ \text{ mean not zero}
\end{aligned}
$$

n patients

$$
i^{th} \;:\; \ln T_i = \beta z_i + \epsilon_i, \quad \epsilon_i \text{ iid } \sim S_0(\cdot).
$$

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

<u>Observe</u> $(U_i, \delta_i, Z_i) \qquad i = 1, 2, \ldots, n$

$$
\left.\begin{aligned}
U_i &= \min(T_i, C_i) \\
\delta_i &= 1(T_i \le C_i)
\end{aligned}\right\} \quad \text{usual setting.}
$$

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

Also, $\epsilon_i \le d_i \Leftrightarrow \ln T_i \le \ln C_i \Leftrightarrow T_i \le C_i$. Thus, $1_{\{\epsilon_i \le d_i\}} \equiv \delta_i$.

Thus, consider the "data"

$$
(U_i^*, \delta_i) \qquad i = 1, 2, \ldots, n
$$

where

$$
U_i^* = \min(\epsilon_i, d_i) \qquad\qquad \underline{\text{Note}}\ \epsilon_i \perp d_i | Z_i.
$$

Thus

$$
\begin{aligned}
U_i^* &= \min(\ln T_i, \ln C_i) - \beta Z_i \\
&= \ln(\min(T_i, C_i)) - \beta Z_i \\
&= \ln(U_i) - \beta Z_i = e_i(\beta).
\end{aligned}
$$

Thus, we've created "data" where the underlying survival time ϵ_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^*, δ_i). Note: for the true β, the underlying "survival times" ϵ_i have a distribution that does not depend on Z_i, and therefore follow a Cox propotional hazards model with parameter $\beta^* = 0$.

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

or

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

Note

$$\begin{aligned} e_j(\beta) &\geq e_i(\beta) \\ \Leftrightarrow \quad \ln U_j - \beta Z_j &\geq \ln U_i - \beta Z_i \\ \Leftrightarrow \quad \ln\left(\frac{U_j}{U_i}\right) &\geq \beta(Z_j - Z_i). \end{aligned}$$

Thus, in summary

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

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

Then, it can be shown that

$$\widehat{\beta} \xrightarrow{p} \beta$$

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

<u>Note</u>: Close connection to rank tests - Tsiatis, 1990, <u>Ann. Stat</u>, Vol 18, p 354–72.

General References to AFT Models

- Wei, Ying, Lin, 1990, JASA, Vol 79, p 649–652.
- Wei, 1992, Stat in Med, Vol 11, p 1871–1879.

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.