2. Accelerated Failure Time Model

- Accelerated failure time model (AFTM)
  - failure time T
  - linear regression model on log-transformed T:

$$\log T = -\beta^T Z + \epsilon$$

  - β: regression parameter - ratio of failure time per unit change in covariate
  - ϵ: random error with unspecified distribution
  - directly modeling on failure time

- Alternative form 1 (in survival functions):

$$\begin{aligned} S(t \mid Z) &= \mathrm{Pr}\{T > t \mid Z\} = \mathrm{Pr}\{\log T > \log t \mid Z\} \\ &= \mathrm{Pr}\{-\beta^T Z + \epsilon > \log t \mid Z\} \\ &= \mathrm{Pr}\{\exp(\epsilon) > t \exp(\beta^T Z) \mid Z\} = S_0(t \exp(\beta^T Z)) \end{aligned}$$

- Alternative form 2 (in quantile functions):

$$Q(p \mid Z) = Q_0(p) \exp(-\beta^T Z)$$

- Alternative form 3 (in hazard functions)

$$
\begin{aligned}
S(t \mid Z) &= S_0(t \exp(\beta^T Z)) \\
&\Rightarrow -\frac{d \log S(t \mid Z)}{d\beta} = -\frac{d \log S_0(t \exp(\beta^T Z))}{d\beta} \\
&\Rightarrow \lambda(t \mid Z) = \lambda_0(t \exp(\beta^T Z)) \exp(\beta^T Z)
\end{aligned}
$$

  - cf. proportional hazards model

$$
\lambda(t \mid Z) = \lambda_0(t) \exp(\beta^T Z)
$$

  - the only difference is additional time-scale change in hazard function

- Estimation:
  - data: $\{(X_i, \Delta_i, Z_i), i = 1, 2, \ldots, n\}$
  - assumption: conditional independence
  - how do we estimate β?

- Can we use QPS approach?
  - assume one-dimensional β
  - calculate the score function for β as if $\lambda_0(\cdot)$ were known

- How do we estimate $\lambda_0(\cdot)$?
  - first consider the QPS approach for the proportional hazards model
  - since $E[dM_i(u;\beta_0)] = 0$, let's consider

$$\begin{aligned}\sum_{i=1}^{n} dN_i(u) &= \sum_{i=1}^{n} Y_i(u)\widehat{\lambda}(u \mid Z_i)du \\ &= \sum_{i=1}^{n} Y_i(u)\widehat{\lambda}_0(ue^{\beta_0 Z_i})e^{\beta_0 Z_i}du\end{aligned}$$

  - it seems there is no easy way to factor out $\widehat{\lambda}_0(\cdot)$ for a solution of this equation
  - but how about let's replace u with $ue^{-\beta_0 Z_i}$?

– alternatively, consider

$$\mathscr{F}_t = \sigma\{N_i(se^{-\beta_0 Z_i}), Y_i(se^{-\beta_0 Z_i}), Z_i; s \le t, i = 1, \ldots, n\}$$

– then

$$\begin{aligned}
&E[dN_i(ue^{-\beta_0 Z_i}) \mid \mathscr{F}_{u-}] \\
&\quad = Y_i(ue^{-\beta_0 Z_i}) d\Lambda(ue^{-\beta_0 Z_i} \mid Z_i) \\
&\quad = Y_i(ue^{-\beta_0 Z_i}) \lambda(ue^{-\beta_0 Z_i} \mid Z_i) e^{-\beta_0 Z_i} du \\
&\quad = Y_i(ue^{-\beta_0 Z_i}) \lambda_0(ue^{-\beta_0 Z_i} \cdot e^{\beta_0 Z_i}) \cdot e^{\beta_0 Z_i} e^{-\beta_0 Z_i} du \\
&\quad = Y_i(ue^{-\beta_0 Z_i}) \lambda_0(u) du
\end{aligned}$$

– hence we can use

$$\sum_{i=1}^{n} dN_i(ue^{-\beta Z_i}) = \sum_{i=1}^{n} Y_i(ue^{-\beta Z_i}) \widehat{\lambda}_0(u) du$$

to solve for $\Lambda_0(\cdot)$:

$$\widehat{\Lambda}_0(t; \beta) = \int_0^t \frac{\sum_i dN_i(ue^{-\beta Z_i})}{\sum_i Y_i(ue^{-\beta Z_i})}$$

– about $\widehat{\Lambda}_0(\cdot;\beta_0)$:

$$\int_0^t \frac{\sum_i dN_i(ue^{-\beta_0 Z_i})}{\sum_i Y_i(ue^{-\beta_0 Z_i})} = \int_0^t \frac{n^{-1}\sum_i dN_i(ue^{-\beta_0 Z_i})}{n^{-1}\sum_i Y_i(ue^{-\beta_0 Z_i})}$$

$$\to_D \int_0^t \frac{EdN(ue^{-\beta_0 Z})}{EY(ue^{-\beta_0 Z})} = \Lambda_0(t)$$

- What equation shall we use to estimate β?

– now let's consider the score for β

$$l'(\beta) = \sum_{i=1}^{n} \int_0^\infty \frac{\partial \log \lambda(u \mid Z_i;\beta)}{\partial \beta} dM_i(u;\beta)$$

– what is $\partial \log \lambda(u \mid Z_i; \beta)/\partial\beta$?

$$\begin{aligned}
\frac{\partial \log\{\lambda_0(ue^{\beta Z_i})e^{\beta Z_i}\}}{\partial\beta} &= \frac{\partial}{\partial\beta}\{\log \lambda_0(ue^{\beta Z_i}) + \beta Z_i\} \\
&= \frac{\lambda_0'(ue^{\beta Z_i})}{\lambda_0(ue^{\beta Z_i})} ue^{\beta Z_i} Z_i + Z_i \\
&= \left[\frac{\lambda_0'(ue^{\beta Z_i})ue^{\beta Z_i}}{\lambda_0(ue^{\beta Z_i})} + 1\right] Z_i \\
&= W(ue^{\beta Z_i})Z_i \\
&\propto Z_i
\end{aligned}$$

where $W(\cdot)$ is considered as weight function

– for simplicity, first consider $W(u) \equiv 1$ and plug in $\widehat{\Lambda}_0(\cdot;\beta)$

$$
\begin{aligned}
&\sum_{i=1}^{n}\int_0^\infty Z_i\{dN_i(u) - Y_i(u)\widehat{\lambda}(u \mid Z_i)du\} \\
&= \sum_{i=1}^{n}\int_0^\infty Z_i\{dN_i(ue^{-\beta Z_i}) - Y_i(ue^{-\beta Z_i})\widehat{\lambda}(ue^{-\beta Z_i} \mid Z_i)d(ue^{-\beta Z_i})\} \\
&= \sum_{i=1}^{n}\int_0^\infty Z_i\{dN_i(ue^{-\beta Z_i}) - Y_i(ue^{-\beta Z_i})\widehat{\lambda}_0(ue^{-\beta Z_i}\cdot e^{\beta Z_i})e^{\beta Z_i}d(ue^{-\beta Z_i})\} \\
&= \sum_{i=1}^{n}\int_0^\infty Z_i\{dN_i(ue^{-\beta Z_i}) - Y_i(ue^{-\beta Z_i})\widehat{\lambda}_0(u)du\} \\
&= \sum_{i=1}^{n}\int_0^\infty \left[Z_i dN_i(ue^{-\beta Z_i}) - Z_i Y_i(ue^{-\beta Z_i})\frac{\sum_i dN_i(ue^{-\beta Z_i})}{\sum_i Y_i(ue^{-\beta Z_i})}\right] \\
&= \sum_{i=1}^{n}\int_0^\infty \left[Z_i - \bar{Z}(u;\beta)\right] dN_i(ue^{-\beta Z_i})
\end{aligned}
$$

where

$$
\bar{Z}(u;\beta) = \frac{\sum_j Z_j Y_j(ue^{-\beta Z_j})}{\sum_j Y_j(ue^{-\beta Z_j})}
$$

- hence, to estimate β, we shall set

$$S_n(\beta) = \sum_{i=1}^{n} \int_0^\infty [Z_i - \bar{Z}(u;\beta)]\, dN_i(ue^{-\beta Z_i}) = 0$$

for a solution $\widehat{\beta}$

- what is $S_n(\beta)$?

$$\begin{aligned} S_n(\beta) &= \sum_{i=1}^{n} \int_0^\infty \left[Z_i - \bar{Z}(ue^{\beta Z_i};\beta)\right] dN_i(u) \\ &= \sum_{i=1}^{n} \Delta_i \left[Z_i - \frac{\sum_j Z_j Y_j(X_i e^{\beta Z_i} \cdot e^{-\beta Z_j})}{\sum_j Y_j(X_i e^{\beta Z_i} \cdot e^{-\beta Z_j})}\right] \\ &= \sum_{i=1}^{n} \Delta_i \left[Z_i - \frac{\sum_j Z_j I(X_j e^{\beta Z_j} \geq X_i e^{\beta Z_i})}{\sum_j I(X_j e^{\beta Z_j} \geq X_i e^{\beta Z_i})}\right] \\ &= \sum_{i=1}^{n} \Delta_i \left[Z_i - \frac{\sum_j Z_j I\{\epsilon_j(\beta) \geq \epsilon_i(\beta)\}}{\sum_j I\{\epsilon_j(\beta) \geq \epsilon_i(\beta)\}}\right] \end{aligned}$$

- this is linear rank test for $\beta = \beta_0$

- What if $\beta = 0$?

  $$\sum_{i=1}^{n} \Delta_i \left[Z_i - \frac{\sum_j Z_j I(X_j e^{\beta Z_j} \geq X_i e^{\beta Z_i})}{\sum_j I(X_j e^{\beta Z_j} \geq X_i e^{\beta Z_i})} \right] = \sum_{i=1}^{n} \Delta_i \left[Z_i - \frac{\sum_j Z_j Y_j(X_i)}{\sum_j Y_j(X_i)} \right],$$

  which is log-rank statistic

- What do we learn from this seemingly long derivation yet full of coincidences?

  - the $O - E$ routine works

  - time-scale can be changed with appropriate filtration

  - risk set structure is preserved

- Accelerated failure time model:

$$\log T = -\beta^T Z + \epsilon$$

- $\beta-$estimation: solve

$$S_n(\widehat{\beta}) = \sum_{i=1}^{n} \left\{ Z_i - \bar{Z}(u; \widehat{\beta}) \right\} dN_i(ue^{-\widehat{\beta} Z_i}) = 0,$$

where

$$\bar{Z}(u; \beta) = \frac{\sum_j Z_j Y_j(ue^{-\beta Z_j})}{\sum_j Y_j(ue^{-\beta Z_j})}$$

- $S_n(\beta)$:

$$S_n(\beta) = \sum_{i=1}^{n} \Delta_i \left[Z_i - \frac{\sum_j Z_j I\{\epsilon_j(\beta) \geq \epsilon_i(\beta)\}}{\sum_j I\{\epsilon_j(\beta) \geq \epsilon_i(\beta)\}} \right],$$

which is a function of the ranks for residuals, $\epsilon_i(\beta) = \log X_i + \beta Z_i$

- What is the difficulty?
  - $S_n(\beta)$ changes only when ranks are altered for at most $\binom{n}{2}$ times
  - $S_n(\beta)$ is not a continuous function of β for a given n but a step function
  - the usual Taylor series expansion does not work

- How do we show the asymptotics of $\widehat{\beta}$?
  - asymptotic linear approximation
    1. find a neighborhood of β_0
    2. approximate $S_n(\beta)$ by a close enough $\tilde{S}_n(\beta)$ linear in β
    3. solve $\tilde{S}_n(\beta_*) = 0$: $n^{1/2}(\beta_* - \beta_0) \rightarrow_D \mathcal{N}(0, \sigma_*^2)$
    4. then show: $n^{1/2}(\widehat{\beta} - \beta_*) \rightarrow_P 0$

- Proof outline
  - for simplicity, assume $\beta_0 = 0$: $\lambda_i(t) = \lambda(t \mid Z_i) = \lambda_0(t)$
  - notice that
    * since $\bar{Z}(u;\beta) = \dfrac{\sum_j Z_j Y_j(ue^{-\beta Z_j})}{\sum_j Y_j(ue^{-\beta Z_j})}$
    * therefore

$$
\begin{aligned}
&\sum_{i=1}^{n} \int_0^\infty \{Z_i - \bar{Z}(u;\beta)\}\, \lambda_0(u) Y_i(ue^{-\beta Z_i})du \\
&= \sum_{i=1}^{n} \int_0^\infty \left\{ Z_i - \frac{\sum_j Z_j Y_j(ue^{-\beta Z_j})}{\sum_j Y_j(ue^{-\beta Z_j})} \right\} \lambda_0(u) Y_i(ue^{-\beta Z_i})du \\
&= \sum_{i=1}^{n} \int_0^\infty Z_i \lambda_0(u) Y_i(ue^{-\beta Z_i})du \\
&\quad - \sum_{i=1}^{n} \int_0^\infty \frac{\sum_j Z_j Y_j(ue^{-\beta Z_j})}{\sum_j Y_j(ue^{-\beta Z_j})} \times \lambda_0(u) Y_i(ue^{-\beta Z_i})du = 0
\end{aligned}
$$

– consider arbitrary β in a neighborhood of $U(0)$:

$$
\begin{aligned}
& S_n(\beta) \\
&= \sum_{i=1}^{n} \int_0^\infty \{Z_i - \bar{Z}(u;\beta)\}\, dN_i(ue^{-\beta Z_i}) \\
&= \sum_{i=1}^{n} \int_0^\infty \{Z_i - \bar{Z}(u;\beta)\}\, \{dN_i(ue^{-\beta Z_i}) - Y_i(ue^{-\beta Z_i}) d\Lambda_i(ue^{-\beta Z_i})\} \\
&\quad + \sum_{i=1}^{n} \int_0^\infty \{Z_i - \bar{Z}(u;\beta)\}\, Y_i(ue^{-\beta Z_i})\{\lambda_i(ue^{-\beta Z_i})e^{-\beta Z_i} - \lambda_0(u)\}du \\
&= \sum_{i=1}^{n} \int_0^\infty \{Z_i - \bar{Z}(u;\beta)\}\, dM_i(ue^{-\beta Z_i}) \\
&\quad + \sum_{i=1}^{n} \int_0^\infty \{Z_i - \bar{Z}(u;\beta)\}\, Y_i(ue^{-\beta Z_i})\{\lambda_0(ue^{-\beta Z_i})e^{-\beta Z_i} - \lambda_0(u)\}du \\
&= \text{Term I} + \text{Term II}
\end{aligned}
$$

- Term I:

$$\sum_{i=1}^{n} \int_0^\infty \{Z_i - \bar{Z}(u;\beta)\} dM_i(ue^{-\beta Z_i})$$
$$= \sum_{i=1}^{n} \int_0^\infty \left\{Z_i - \bar{Z}(ue^{\beta Z_i};\beta)\right\} dM_i(u)$$

- $\beta \in U(0) \Rightarrow \bar{Z}(ue^{\beta Z_i};\beta) \approx \bar{Z}(u;\beta)$

- therefore, Term I is approximated by $S_n(0)$, i.e., $S_n(\beta)$ at true $\beta_0 = 0$

- Term II:

$$\sum_{i=1}^{n} \int_0^\infty \{Z_i - \bar{Z}(u;\beta)\} Y_i(ue^{-\beta Z_i})\{\lambda_0(ue^{-\beta Z_i})e^{-\beta Z_i} - \lambda_0(u)\}du$$

$$\approx \sum_{i=1}^{n} \int_0^\infty \{Z_i - \bar{Z}(u;\beta)\} Y_i(ue^{-\beta Z_i}) \left.\frac{\partial\{\lambda_0(ue^{-\beta Z_i})e^{-\beta Z_i}\}}{\partial\beta}\right|_{\beta=0} (\beta - 0)du$$

-

$$\left.\frac{\partial\{\lambda_0(ue^{-\beta Z_i})e^{-\beta Z_i}\}}{\partial\beta}\right|_{\beta=0} = -\{\lambda_0'(u)u + \lambda_0(u)\}Z_i = -\{\lambda_0(u)u\}'Z_i$$

- Term II can be approximated by

$$-n\beta \times n^{-1}\sum_{i=1}^{n} \int_0^\infty \{\lambda_0(u)u\}'Z_i\{Z_i - \bar{Z}(u;\beta)\} Y_i(ue^{-\beta Z_i})du$$

$$= -n\beta \times n^{-1}\sum_{i=1}^{n} \int_0^\infty \{\lambda_0(u)u\}'\{Z_i - \bar{Z}(u;\beta)\}^2 Y_i(ue^{-\beta Z_i})du$$

- Term II can be approximated by

$$-n\beta \times n^{-1} \sum_{i=1}^{n} \int_0^\infty \{\lambda_0(u)u\}' \{Z_i - \bar{Z}(u;\beta)\}^2 Y_i(ue^{-\beta Z_i}) du$$

- let $B(u;\beta) = E[\{Z - \bar{Z}(u;\beta)\}^2 Y(ue^{-\beta Z})]$

- then Term II can be further approximated by $-n\beta A(0)$, where

$$A(0) = \int_0^\infty \{\lambda_0(u)u\}' B(u,0) du$$

- in summary of approximation
  - Term I: $S_n(0)$
  - Term II: $-n\beta A(0)$
  - $S_n(\beta)$:
  $$\tilde{S}_n(\beta) = S_n(0) - n\beta A(0)$$
  is a linear function of $\beta \in U(0)$
  - denote $\tilde{S}(\beta_*) = 0$, i.e.
  $$n^{1/2}\beta_* = \frac{n^{-1/2}S_n(0)}{A(0)}$$
  - therefore
  $$n^{1/2}\beta_* \to_D \mathcal{N}(0, \sigma_*^2)$$

- How to show $n^{1/2}(\widehat{\beta} - \beta_*) \to_P 0$?
  - this is basically to show that $\widehat{\beta}$, if exists, shall be consistent
  - such consistency then requires sufficient closeness of $S_n(\beta)$ to $\tilde{S}_n(\beta)$:

$$\sup_{|\beta|<cn^{-1/2}} n^{-1/2}|S_n(\beta) - \tilde{S}_n(\beta)| \to_P 0$$

    for any constant $c > 0$

- to establish this condition

1. pointwise convergence: for any fixed $d > 0$

$$n^{-1/2}|S_n(n^{-1/2}d) - \tilde{S}_n(n^{-1/2}d)| \to_P 0$$

2. for a fixed set of d_i forming a mesh $d_0 < d_1 < \ldots < d_m$:

$$\max_{i=1}^{m} n^{-1/2}|S_n(n^{-1/2}d_i) - \tilde{S}_n(n^{-1/2}d_i)| \to_P 0$$

3. $n^{-1/2}S_n(\beta)$ does not fluctuate within any interval in the mesh

- Since $n^{-1/2}S_n(0) \to_D \mathscr{N}(0, \sigma^2)$ by MCLT,

$$n^{1/2}\widehat{\beta} \to_D \mathscr{N}(0, \sigma^2/A(0)^2)$$

- Ref: Tsiatis (1990, Ann. Stat.)

- Variance calculation
  - how to calculate $A(0) = \int_0^\infty \{\lambda_0(u)u\}' B(u, 0) du$, which involves the baseline hazard function and its derivative?
    * direct approach: kernel density estimation recommended to estimate baseline hazard function
    * simulation approach
    * bootstrap approach

– Weighted estimation

$$S_n^W(\beta) = \sum_{i=1}^{n} W_n(u;\beta)\left\{Z_i - \bar{Z}(u;\beta)\right\} dN_i(ue^{-\beta Z_i}) = 0,$$

* $W_n(u;\beta_0)$ is $\mathscr{F}_u$-measurable and converges uniformly in probability to a deterministic function $w(u)$

* $W_n(u) = \sum_j Y_j(u)/n$: Gehan-Wilcoxon

* $W_n(u) = \widehat{S}_{KM}(u-)$: Prentice-Wilcoxon

* $W_n(u) = \{\widehat{S}_{KM}(u-)\}^{\rho}$: G^{ρ}-family

- Asymptotic properties on weighted estimators $\widehat{\beta}_n^W$ for a general β_0

$$n^{1/2}(\widehat{\beta}_n^W - \beta_0) \rightarrow_D \mathcal{N}(0, \sigma^2(w)/A(w)^2)$$

  * $\sigma_W^2 = \int_0^\infty w(u;\beta_0)^2 A(u,\beta_0)\lambda(u \mid Z)du$
  * $A(w) = \int_0^\infty w(u;\beta_0)A(u,\beta_0)\{\lambda(u \mid Z)u\}'du$

- By Cauchy-Schwarz inequality, optimal weight function should be proportional to

$$\begin{aligned}
\frac{\{\lambda(u \mid Z)u\}'}{\lambda(u \mid Z)} &= \frac{\{\lambda_0(ue^{\beta Z})u\}'}{\lambda_0(ue^{\beta Z})} \\
&= \left[\frac{\{\lambda_0'(ue^{\beta Z})ue^{\beta Z}\}}{\lambda_0(ue^{\beta Z})} + 1\right] Z \\
&= \frac{d}{d\beta}\left\{\log \lambda_0(ue^{\beta Z}) + \beta Z\right\} \\
&= \frac{\partial \log \lambda(t \mid Z)}{\partial \beta}
\end{aligned}$$

- Partial likelihood for AFTM
  - data: (X_i, Δ_i, Z_i)
  - $e_i(\beta) = \log X_i + \beta Z_i = \min\{\log T_i - \beta Z_i, \log C_i - \beta Z_i\} = \min(\epsilon_i, d_i)$
  - $\Delta_i = 1$ for $\epsilon_i \leq d_i$ and 0 for $\epsilon_i > d_i$
  - for a known β, data reformulated as $(e_i(\beta), \Delta_i, Z_i)$.
  - when $\beta = \beta_0$, $\log T_i + \beta_0 Z_i$ follows same distribution as if

$$\lambda(t \mid Z_i) = \lambda_0(t) \exp(\gamma Z_i)$$

    for $\gamma = 0$

  - partial likelihood score function for γ:

$$\sum_{i=1}^{n} \Delta_i \left[Z_i - \frac{\sum_j Z_j I\{e_j(\beta_0) \geq e_i(\beta_0)\}}{\sum_j I\{e_j(\beta_0) \geq e_i(\beta_0)\}} \right]$$

- Time-dependent covariates $\bar{Z}(t) = \{Z(s), 0 \leq s < t\}$
  - hazard-based model

$$\lambda(t \mid \bar{Z}(t)) = \lambda_0 \left[\int_0^t \exp\{\beta Z(u)\} du\right] \exp\{\beta Z(t)\}$$

- References:
  - Cox & Oakes (1984, p.66)
  - Robins & Tsiatis (1992, Bmka)
  - Lin & Ying (1995, JSPI)