- 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

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

- 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 $\tilde{S}_n(\beta_*) = 0$: $n^{1/2}(\beta_* - \beta_0) \to_D \mathcal{N}(0, \sigma_*^2)$
    4. then show: $n^{1/2}(\widehat{\beta} - \beta_*) \to_P 0$

- 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

$$n^{1/2}\beta_* \to_D \mathcal{N}(0, \sigma_*^2)$$

– 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 $\epsilon_i \le d_i$ 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:

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

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

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

  - equivalently, $T_0 = \int_0^T \exp\{\beta Z(u)\} du$
    - when $Z(u) \equiv Z$, $\log T = -\beta Z + \log T_0$
    - interpretation of β: time expansion/contraction

- Parameter estimation
  - Parametric approach: Cox & Oakes (1984, p.66)
  - Rank-type approach: Robins & Tsiatis (1992, Bmka)
  - QPS approach: Lin & Ying (1995, JSPI)

- Buckley-James estimation
  - log-likelihood function

$$l(\beta) = \sum_{i=1}^{n} \{\delta_i \log f(x_i; \beta) + (1 - \delta_i) \log S(x_i; \beta)\}$$

  - score function:

$$l'(\theta) = \sum_{i=1}^{n} \left\{\delta_i \frac{\partial \log f(x_i; \beta)}{\partial \beta} + (1 - \delta_i) \frac{\partial \log S(x_i; \beta)}{\partial \beta}\right\}$$

  -

$$\frac{\partial \log S(x_i; \beta)}{\partial \beta} = E\left[\left.\frac{\partial \log f(T_i; \beta)}{\partial \beta}\right| T_i > x_i\right]$$

- in general, we can substitute $\frac{\partial \log f(x_i;\beta)}{\partial\beta}$ with any mean-zero quantity $\alpha(u;\beta)$, because

$$\begin{aligned} &E[\delta_i\alpha(X_i;\beta) + (1-\delta_i)E\{\alpha(T_i;\beta) \mid T_i > X_i\}] \\ =&E[\alpha(T_i;\beta)] = 0 \end{aligned}$$

- Buckley-James (1979, Bmka): $\alpha(u;\beta) = \log u + \beta Z_i$ assuming that $E(\epsilon_i) = 0$

- to estimate β, consider estimating equations

$$\sum_{i=1}^{n} \delta_i\alpha(X_i;\beta) + (1-\delta_i)E\{\alpha(T_i;\beta) \mid T_i > X_i\} = 0,$$

where $E\{\alpha(T_i;\beta) \mid T_i > X_i\}$ needs to be appropriately estimated

- Buckely-James approach:

$$\widehat{E}\{\alpha(T;\beta) \mid T > x_i\} = \frac{1}{\widehat{S}(x_i)} \int_{x_i}^{\infty} \alpha(u) d\widehat{F}(u)$$

  where $\widehat{S}(\cdot)$ is the Kaplan-Meier estimator

- therefore, we can use estimating equations

$$S_n(\beta) = \sum_{i=1}^{n} \delta_i \alpha(X_i;\beta) + (1-\delta_i)\widehat{E}\{\alpha(T_i;\beta) \mid T_i > X_i\} = 0$$

  to solve for $\widehat{\beta}$

- Additional thoughts on Buckley-James:

  - self-consistency representation of $\widehat{S}$ (Efron, 1967, Proc. 5th Berk. Symp)

    $$d\widehat{F}(u) = n^{-1} \sum_{i=1}^{n} \{\delta_i dI(u \geq x_i) + (1 - \delta_i) d\widehat{F}(u) I(u > x_i)/\widehat{S}(x_i)\}$$

    considering $\widehat{S}(u-) - \widehat{S}(u)$

  - therefore

    $$\widehat{E}\{\alpha(T)\} = \int_{-\infty}^{\infty} \alpha(u) d\widehat{F}(u) = S_n(\beta)/n$$

  - solving $S_n(\beta) = 0$ is equivalent to solving

    $$\int_{-\infty}^{\infty} \alpha(u;\beta) d\widehat{F}(u) = 0$$

- Functional $\int u d\widehat{F}(u)$
  - major difficulty is in the right-tail with censored data
  - Susarla & Van Ryzin (1980, Ann. Stat.): integration with upper limit that goes to infinity at appropriate rate
  - Reid (1981, Ann. Stat.): various functional forms
  - Gill (1983, Ann. Stat.): regularity conditions on censoring distributions

- Summary on AFT Model
  - model setup and interpretation
  - estimation
    - extending QPS approach
    - local linear approximation
  - Buckley-James estimation

3. Alternative regression models

- Two major classes:
  - hazard models based on rates: proportional hazards model
  - failure time models based on actual failure times: accelerated failure time model

- General relative risk model (Prentice & Self, 1983, Ann. Stat.)

$$\lambda(t \mid Z) = \lambda_0(t) r(\beta Z)$$

  - exponential form: $r(\beta Z) = \exp(\beta Z)$
  - linear form: $r(\beta Z) = 1 + \beta Z$

- Additive hazards model (Lin & Ying, 1994, Bmka):

$$\lambda(t \mid Z) = \lambda_0(t) + \beta Z$$

  - interpretation: additive covariate effect
  - embedded constraint: $\lambda_0(t) + \beta Z > 0$
  - QPS model estimation

$$E[dN_i(t) - Y_i(t)\beta Z_i dt \mid \mathscr{F}_{t-}] = Y_i(t)\lambda_0(t)dt$$

  - invariant in marginalization
  - extension: additive-multiplicative hazards model (Lin & Ying, 1994, Ann. Stat.)

$$\lambda(t \mid Z_1, Z_2) = \lambda_0(t)\exp(\beta Z_1) + \gamma Z_2$$

- Accelerated hazards model:

$$\lambda(t \mid Z) = \lambda_0\{t \exp(\beta Z)\}$$

  - parameter interpretation: accelerated/decelerated risk progression

  - parameter estimation: QPS approach

$$E[dN_i(te^{-\beta_0 Z_i}) \mid \mathscr{F}_{t-}; \beta_0] = Y_i(te^{-\beta_0 Z_i})e^{-\beta_0 Z_i}\lambda_0(t)dt$$

- alternative approach:
  - for any γ, transform $T_i^* = T_i e^{\gamma Z_i}$. The hazard function for transformed time becomes:

$$\lambda_i^*(t \mid Z_i) = \lambda_i(te^{-\gamma Z_i})e^{-\gamma Z_i} = \lambda_0(te^{(\beta-\gamma)Z_i})e^{-\gamma Z_i}$$

  - when $\gamma = \beta$, then we obtain the proportional hazards model

$$\lambda_i^*(t \mid Z_i) = \lambda_0(t)e^{-\beta Z_i}$$

  - algorithm motivated to find an estimate
- Chen & Wang (2000, JASA)

- General class of hazards model:

$$\lambda(t \mid Z) = \lambda_0\{t \exp(\beta Z)\} \exp(\gamma Z)$$

  - a general class to include PHM, AFTM and AHM as sub classes
  - identifiability: exponential distribution
  - model estimation: QPS approach
  - semiparametric efficiency
  - Chen & Jewell (2001, Bmka)

- Proportional odds model

$$\log \frac{S(t \mid Z)}{1 - S(t \mid Z)} = \log \frac{S_0(t)}{1 - S_0(t)} + \beta Z$$

  - Bennett (1983, Stat. Med.)
  - Murphy (1997, JASA)
  - Yang & Prentice (1999, JASA)
  - Open questions: does QPS work? how does it compare with other approaches? what's the optimal weight function for log-rank statistic when alternative is the proportional odds model?

- Generalized model

$$g\{S(t \mid Z)\} = g_0(t) + \beta Z$$

  - $g(\cdot)$ known decreasing function
  - log-log link: proportional hazards model
  - logit link: proportional odds model
  - generalized odds-rate model

$$g(s) = \log\left(\frac{1}{\lambda}\frac{1 - s^{\lambda}}{s^{\lambda}}\right) I(\lambda > 0) + \log(-\log s) I(\lambda = 0)$$

- Linear transformation model

$$h(T) = -\beta Z + \epsilon$$

  - $h(\cdot)$ is unknown
  - ϵ's distribution is known
    - extreme value distribution: proportional hazards model
    - standard logistic distribution: proportional odds model
  - Cheng & Wei (1995, Bmka)

- Proportional mean residual life model

$$m(t \mid Z) = m_0(t)\,\text{exp}(\beta Z)$$

  - model interpretation
  - model constraint: $m_0(t)\,\text{exp}(\beta Z) + t$ shall be monotonically non-decreasing
  - connection with hazard functions under renewal processes
  - model estimation: QPS approach
  - Oakes (1990, Bmka)
  - Chen & Cheng (2005, Bmka)

- Additive mean residual life model

$$m(t \mid Z) = m_0(t) + \beta Z$$

  - model interpretation: additional life expectancy
  - model constraint: $m_0(t) + \beta Z \geq 0$
  - model estimation: QPS and Buckley-James
  - semiparametric efficiency
  - Chen & Cheng (2006, Bmka)

- Summary
  - alternative regression models developed to address various limitations of the proportional hazards model
  - no perfect model
  - software yet to be developed
  - still an active research area