- Semiparametric proportional hazards model

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

  - Estimation of β: $\widehat{\beta}_n$ solved in partial score equation

$$S_n(\widehat{\beta}_n) = \sum_{i=1}^{n} \int_0^\infty \left\{ Z_i - \bar{Z}(u; \widehat{\beta}_n) \right\} dN_i(u) = 0$$

  - Estimation of $\Lambda_0(\cdot)$: Breslow estimator

$$\widehat{\Lambda}_0(t; \widehat{\beta}_n) = \int_0^t \frac{\sum_{i=1}^n dN_i(u)}{\sum_{i=1}^n Y_i(u)\exp(\widehat{\beta}_n^T Z_i)}$$

- An application: Kalbfleisch & Prentice (2002, pp. 119-128)

- Model-based prediction
  - predict survival function for a given covariate Z_0
  - model-based prediction

$$\widehat{\Lambda}(t \mid Z_0) = \widehat{\Lambda}_0(t; \widehat{\beta}_n) \exp(\widehat{\beta}_n^T Z_0)$$

  - confidence interval or confidence band

- direct method

$$
\begin{aligned}
& n^{1/2}\{\widehat{\Lambda}(t \mid Z_0) - \Lambda_0(t)\exp(\beta_0^T Z_0)\} \\
&= n^{1/2}\{\widehat{\Lambda}_0(t;\widehat{\beta}_n)\exp(\widehat{\beta}_n^T Z_0) - \widehat{\Lambda}_0(t;\beta_0)\exp(\beta_0^T Z_0)\} \\
&\quad + n^{1/2}\{\widehat{\Lambda}_0(t;\beta_0)\exp(\beta_0^T Z_0) - \Lambda_0(t)\exp(\beta_0^T Z_0)\} \\
&\doteq \Lambda_0(t)\exp(\beta_0^T Z_0) Z_0^T \times n^{1/2}(\widehat{\beta}_n - \beta_0) \\
&\quad + \exp(\beta_0^T Z_0) \times n^{1/2}\{\widehat{\Lambda}_0(t;\beta_0) - \Lambda_0(t)\} \\
&= \text{Term I} + \text{Term II}
\end{aligned}
$$

- simulation method

- bootstrap method

- log-transformation to avoid negative confidence interval/band on $\widehat{\Lambda}(t)$

- confidence interval/band for survival function estimate

$$\widehat{S}(t \mid Z_0) = \exp\{-\widehat{\Lambda}(t \mid Z_0)\}$$

- Reference: Lin, Fleming & Wei (1994, Bmka)

- Model-based estimation of median survival time
  - given covariate Z_0: $S(m \mid Z_0) = 1/2$
  - estimated median $\widehat{m}$: $\widehat{S}(\widehat{m} \mid Z_0) = 1/2$
  - confidence interval for m:
    * find the asymptotic distribution of

$$
\begin{aligned}
& n^{1/2}\{\log \widehat{S}(m \mid Z_0) - \log S(m \mid Z_0)\} \\
= & - n^{1/2}\{\widehat{\Lambda}_0(m) \exp(\widehat{\beta}^T Z_0) - \Lambda_0(m) \exp(\beta^T Z_0)\} \\
\to_D & \mathcal{N}(0, \sigma^2(m))
\end{aligned}
$$

    * 95% confidence interval should the collection of m s.t.

$$
\left| \frac{\log \widehat{S}(m \mid Z_0) - \log(1/2)}{\sigma(m)/\sqrt{n}} \right| < 1.96
$$

- Stratified proportional hazards model
  - strata divide subjects into m disjoint groups
  - allows for multiple strata and each strata has individual baseline hazard function $\lambda_{i0}(t)$
    - multicenter trials; block design
  - model specification: Z_{ij} jth covariate in ith stratum

$$\lambda_i(t \mid Z_{ij}) = \lambda_{i0}(t) \exp(\beta^T Z_{ij})$$

    - most general adjustment of confounding variables
    - no direct estimate of strata effect; interaction; interpretation
  - estimation: $\sum_{i=1}^{m} S_i(\beta) = 0$

- Time-dependent covariates in Cox model
  - examples: $Z_i(t)$
    * cumulative exposure to contamination
    * smoking status
    * heart transplant status: 0=prior to, 1=after heart transplant
    * blood presure
  - covariate history: $\bar{Z}_i(t) = \{Z_i(s); 0 \le s \le t\}$
  - hazard function of $\bar{Z}_i(t)$

$$\lambda(t \mid \bar{Z}_i(t)) = \lim_{h \to 0} \frac{\Pr\{t \le T_i < t + h \mid T_i \ge t, \bar{Z}_i(t)\}}{h}$$

- Extended Cox model

$$\lambda(t \mid \bar{Z}_i(t)) = \lambda_0(t) \exp\{\beta^T g(\bar{Z}_i(t))\}$$

  where $g(\cdot)$ is left-continuous functionals of $\bar{Z}_i(t)$

- For simplicity: g is identity functional

- Data collected: $\{X_i, \Delta_i, \bar{Z}_i(X_i); i = 1, \ldots, n\}$

- Filtration:

$$\mathscr{F}_t = \sigma\{N_i(u), Y_i(u), Z_i(u), u \leq t; i = 1, \ldots, n\}$$

- The identity: $E[dN_i(t) \mid \mathscr{F}_{t-}] = Y_i(t)\lambda_0(t) \exp\{\beta^T Z_i(t)\}dt$

- Assume noninformative censoring

- Types of time-dependent covariates
  - internal covariates: change of covariates over time dependent on the individual's risk progression
  - external covariates: sample paths generated by external force; can be measured over time regardless subject is alive or not

- Estimation in presence of external time-dependent covariates: PS 5, Q 2

- Challenging issues
  - interpretation in survival function:

$$S(t \mid Z_i(t)) = \text{exp}\left[-\int_0^t \lambda_0(u)\,\text{exp}\{\beta^T Z_i(u)\}du\}\right]$$

  - data collection: time-dependent covariates need to be collected at each failure time for those at risk
  - joint modeling on internal time-dependent covariates and risk progression

- An example of time-dependent covariates
  - model-checking on proportional hazards model
  - assume Cox model with time-dependent covariates

    $$\lambda(t \mid \bar{Z}_i) = \lambda_0(t) \exp\{\beta Z_i + \gamma Z_i g(t)\},$$

    where $g(t)$ is some pre-specified function, e.g., $\log(t)$
  - testing $\gamma = 0$ is equivalent to checking model adequacy
  - $Z_i g(t)$ is external time-dependent covariates defined by model user

- Model adequacy checking

  - weighted estimation

$$S_n^w(\beta) = \sum_{i=1}^{n} \int_0^\infty W(u)\{Z_i - \bar{Z}(u;\beta)\} dN_i(u)$$

    where $W(u)$ is predictable weight function

  - weighted estimator: $\widehat{\beta}_w$: $S_n^w(\widehat{\beta}_w) = 0$

$$n^{1/2}(\widehat{\beta}_w - \beta_0) \rightarrow_D \mathcal{N}(0, \Sigma_w)$$

  - when model is correct, $\widehat{\beta}_w$ and unweighted $\widehat{\beta}$ should be close to each other

$$n^{1/2}(\widehat{\beta}_w - \widehat{\beta}) \rightarrow \mathcal{N}(0, \Sigma_d)$$

  - test statistic: $TS = n(\widehat{\beta}_w - \widehat{\beta})^T \widehat{\Sigma}_d(\widehat{\beta})^{-1}(\widehat{\beta}_w - \widehat{\beta}) \sim \chi_p^2$

  - Reference: Lin (1991, JASA)

- Efficiency regarding partial likelihood estimator $\widehat{\beta}$

  - Under the semiparametric proportional hazards model with baseline hazard function unspecified, are there any other estimation procedure that yield ‘better’ estimators with smaller asymptotic variances?

  - For a given parametric family of baseline hazard functions, what is the relative efficiency of $\widehat{\beta}$ to the maximum likelihood estimators?

- Consider a family of parametric submodels

$$\lambda(t \mid Z(t)) = \alpha h_0(t) \exp\{\beta^T Z(t)\}$$

where $h_0(t)$ is known and α is an unknown parameter

- Calculate Fisher information regarding α and β:

$$\mathscr{I} = \begin{pmatrix} I_{\beta\beta} & I_{\beta\alpha} \\ I_{\alpha\beta} & I_{\alpha\alpha} \end{pmatrix}$$

- Information for β by maximizing out α: $I_1 = I_{\beta\beta} - I_{\beta\alpha} I_{\alpha\alpha}^{-1} I_{\alpha\beta}$

- Information for partial score equation:

$$I_2 = \lim_{n\to\infty} n^{-1} E[-\partial S_n(\beta)/\partial\beta]$$

- If $I_2^{-1} I_1$ converges to identity matrix, $\widehat{\beta}$ is fully efficient
  - when $E[Z(t) \mid X \geq t]$ is constant, $\widehat{\beta}$ is fully efficient
  - when $Z(t)$ is time-independent and $\beta = 0$, $\widehat{\beta}$ is fully efficient

- Classcial treatment:
  - Oakes (1977, Bmka)
  - Efron (1977, JASA)
- Semiparametric efficiency calculation
  - Tsiatis (2006)

- Summary on proportional hazards model
  - model specification
  - model interpretation
  - estimation and asymptotics
  - prediction
  - stratified proportional hazards model
  - time-dependent covariates
  - model checking
  - efficiency comparison

- Yet to cover
  - discrete proportional hazards model
  - partial likelihood filtration
  - residuals and model diagnostics
  - model selection
  - use of proportional hazards model in sequential trials

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

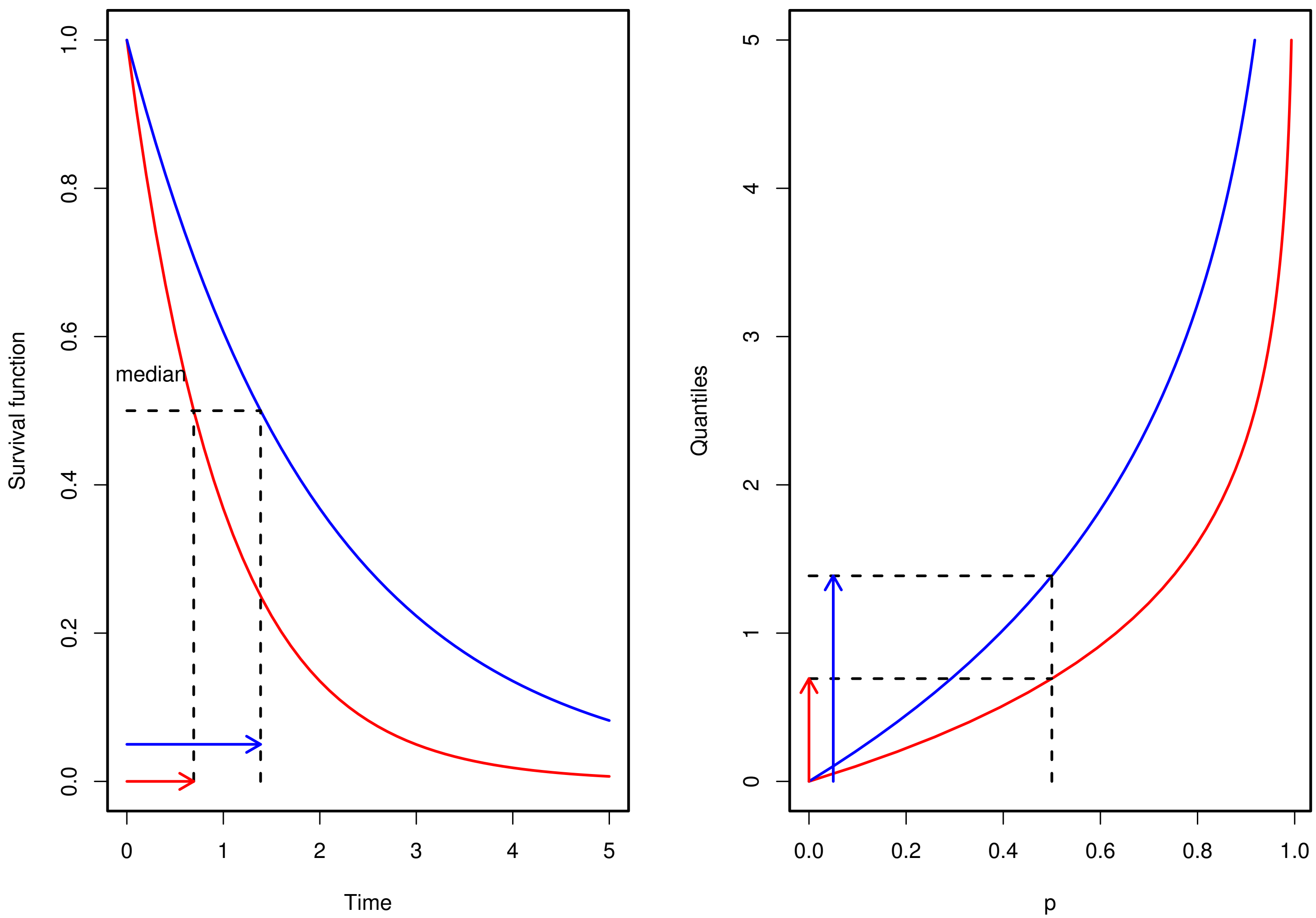
- Alternative form 1 (in survival functions):

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

- Alternative form 2 (in quantile functions):

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

- Two-sample AFTM: $S_1(t) = S_0(te^{\beta})$ vs $Q_1(p) = Q_0(p)e^{-\beta}$



- Alternative form 3 (in hazard functions)

$$S(t \mid Z) = S_0(t \exp(\beta^T Z))$$

$$\Rightarrow - \frac{d \log S(t \mid Z)}{d\beta} = -\frac{d \log S_0(t \exp(\beta^T Z))}{d\beta}$$

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

  - 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

–

$$
\begin{aligned}
\mathscr{L}(\beta) = \prod_{i=1}^{n} \lambda(X_i \mid Z_i)^{\Delta_i} S(X_i \mid Z_i) \\
\Rightarrow l(\beta) &= \sum_{i=1}^{n} \left[\Delta_i \log \lambda(X_i \mid Z_i) - \int_0^{X_i} \lambda(u \mid Z_i) du \right] \\
&= \sum_{i=1}^{n} \int_0^{\infty} [\log \lambda(u \mid Z_i; \beta) dN_i(u) - Y_i(u) \lambda(u \mid Z_i; \beta) du] \\
\Rightarrow l'(\beta) &= \sum_{i=1}^{n} \int_0^{\infty} \frac{\partial \log \lambda(u \mid Z_i; \beta)}{\partial \beta} [dN_i(u) - Y_i(u) \lambda(u \mid Z_i; \beta) du] \\
&= \sum_{i=1}^{n} \int_0^{\infty} \frac{\partial \log \lambda(u \mid Z_i; \beta)}{\partial \beta} dM_i(u; \beta)
\end{aligned}
$$

– we would expect $E[dM_i(u; \beta_0)] = 0$

– $\partial \log \lambda(u \mid Z_i; \beta)/\partial \beta$ is some weight function, which can be replaced by any predictable functions without affecting zero-mean

- 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

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

  - 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