

データ解析

生存時間データ解析 (生存関数の推測, Cox回帰分析)



創域理工学部

Faculty of Science and Technology

東京理科大学
創域理工学部情報計算科学科
安藤宗司

2023年12月7日

Contents

- 生存関数の推定
 - カプランマイヤー推定値
- 生存関数の検定
 - ログランク検定
 - 重み付きログランク検定

諸前提

- 生存時間を表す確率変数 T
- サンプル数 n
- 打ち切りに関する指示変数 C
- n 個の独立な観測値 (t_i, c_i) , $i = 1, \dots, n$ の標本
- $r \leq n$ 個のイベント, $n - r$ 個の中途打ち切りが存在
- 順序付けられた生存時間 $t_{(1)} < t_{(2)} < \dots < t_{(m)}$
- 時点 $t_{(s)}$ のリスク集合 n_s
- 時点 $t_{(s)}$ のイベント数 d_s
- $r = \sum_{s=1}^m d_s$

離散確率変数の場合

□ 離散確率変数 T

■ 離散的に $0(=t_0) < t_1 < t_2 < \dots < t_j$ の値をとり、各時点において確率関数 $p(t_j) = \Pr(T = t_j)$ をもつとする

□ 生存関数 $S(t) = \Pr(T > t) = \sum_{l=j}^J p(t_l) \quad (t_{j-1} \leq t < t_j)$

□ ハザード関数 $h(t)$

$$h(t) = \Pr(T = t_j \mid T > t_{j-1}) = \frac{p(t_j)}{S(t_{j-1})} = \frac{p(t_j)}{\sum_{k=j}^J p(t_k)}$$

□ 累積ハザード関数 $H(t) = \sum_{l=1}^{j-1} h(t_l) \quad (t_{j-1} \leq t < t_j)$

生存関数とハザード関数の関係

$$h(t_j) = \frac{p(t_j)}{S(t_{j-1})} = \frac{S(t_{j-1}) - S(t_j)}{S(t_{j-1})} = 1 - \frac{S(t_j)}{S(t_{j-1})}$$

したがって

$$\begin{aligned} S(t_j) &= S(t_{j-1})(1 - h(t_j)) \\ &= S(t_{j-2})(1 - h(t_j))(1 - h(t_{j-1})) \\ &= \cdots = S(t_0) \prod_{k=1}^j (1 - h(t_k)) = \prod_{k=1}^j (1 - h(t_k)) \end{aligned}$$

カプランマイヤー推定量の形式との関係性を考えるための良い示唆を与える

生存関数の推定

□ カプランマイヤー法

- 生存関数の推定で最も用いられる手法
- ノンパラメトリックに推定
- 生存関数の推定値（カプランマイヤー推定値）

$t_{(s)} \leq \forall t < t_{(s+1)}$ に対して

$$\hat{S}(t) = \prod_{\{s:t_{(s)} \leq t\}} \frac{n_s - d_s}{n_s} = \prod_{\{s:t_{(s)} \leq t\}} \left(1 - \frac{d_s}{n_s} \right)$$

カプランマイヤー推定値の導出 (1)

順序付けられた生存時間 $t_{(1)} < t_{(2)} < \dots < t_{(m)}$ が与えられた
という条件のもとで, $t_{(s)} \leq \forall t < t_{(s+1)}$ とする。

$$\begin{aligned} S(t) &= \Pr(T > t) \\ &= \Pr(T > t, T > t_{(s)}) \\ &= \Pr(T > t_{(s)}) \Pr(T > t \mid T > t_{(s)}) \\ &= \Pr(T > t_{(s)}, T > t_{(s-1)}) \Pr(T > t \mid T > t_{(s)}) \\ &= \Pr(T > t_{(s-1)}) \Pr(T > t_{(s)} \mid T > t_{(s-1)}) \Pr(T > t \mid T > t_{(s)}) \\ &= \dots = \Pr(T > t_{(1)}) \Pr(T > t_{(2)} \mid T > t_{(1)}) \dots \Pr(T > t \mid T > t_{(s)}) \end{aligned}$$

Kaplan-Meier推定値の導出 (2)

ここで, $t_{(0)} = 0$ とすると

$$\begin{aligned}\Pr(T > t_{(1)}) &= \Pr(T > t_{(1)}, T > t_{(0)}) \\ &= \Pr(T > t_{(0)})\Pr(T > t_{(1)} \mid T > t_{(0)}) \\ &= 1 - \Pr(T \leq t_{(1)} \mid T > t_{(0)})\end{aligned}$$

また,

$$\Pr(T > t_{(s)} \mid T > t_{(s-1)}) = 1 - \Pr(T \leq t_{(s)} \mid T > t_{(s-1)})$$

であることから,

$$S(t) = \left(\prod_{i=1}^s [1 - \Pr(T \leq t_{(i)} \mid T > t_{(i-1)})] \right) [1 - \Pr(T \leq t \mid T > t_{(s)})]$$

Kaplan-Meier推定値の導出 (3)

$$S(t) = \left(\prod_{i=1}^s [1 - \Pr(T \leq t_{(i)} \mid T > t_{(i-1)})] \right) [1 - \Pr(T \leq t \mid T > t_{(s)})]$$

\parallel
 $\Pr(T = t_{(i)} \mid T > t_{(i-1)}) = \lambda_i$
 \parallel
 $\Pr(T \leq t_{(s)} \mid T > t_{(s)}) = 0$

T は $t_{(1)}, t_{(2)}, \dots, t_{(m)}$ の値しか
とらないという条件のもと
での条件付ハザード関数

したがって,

$$S(t) = \prod_{i=1}^s (1 - \lambda_i)$$

カプランマイヤー推定値の導出 (4)

$$\begin{aligned}
 \Pr(T = t_{(i)}) &= \lim_{\delta t \rightarrow 0} \Pr(t_{(i)} - \delta t < T \leq t_{(i)}) \\
 &= \lim_{\delta t \rightarrow 0} [\Pr(T \leq t_{(i)}) - \Pr(T \leq t_{(i)} - \delta t)] \\
 &= \lim_{\delta t \rightarrow 0} [\Pr(T > t_{(i)} - \delta t) - \Pr(T > t_{(i)})] \\
 &= S(t_{(i)} - 0) - S(t_{(i)})
 \end{aligned}$$

$t_{(1)}, t_{(2)}, \dots, t_{(m)}$ が与えられたとき, $S = S(t_{(i)})$ の関数とみたときの尤度は

$$L(S) = \prod_{i=0}^m [\Pr(T = t_{(i)})^{d_i} \Pr(T > t_{(i,1)}) \cdots \Pr(T > t_{(i,u_i)})]$$

区間 $[t_{(i)}, t_{(i+1)})$ で u_i 個が中途打ち切りとなり, その時間を

$t_{(i1)} < \cdots < t_{(iu_i)}$ とする。ただし, $t_{(0)} = 0, t_{(m+1)} = \infty$ とする。

カプランマイヤー推定値の導出 (5)

$$\begin{aligned}
 L(S) &= \prod_{i=0}^m \left[\Pr(T = t_{(i)})^{d_i} \Pr(T > t_{(i,1)}) \cdots \Pr(T > t_{(i,u_i)}) \right] \\
 &= \prod_{i=0}^m \left[\underbrace{(S(t_{(i)} - 0) - S(t_{(i)}))^{d_i}}_{\text{イベント}} \underbrace{\prod_{k=1}^{u_i} S(t_{(i,k)})}_{\text{中途打ち切り}} \right]
 \end{aligned}$$

推定量 $\hat{S}(t)$ が連続型とすると第1項は0となるので, $\hat{S}(t)$ は離散型で求まる

第2項は

$$\prod_{i=0}^m \prod_{k=1}^{u_i} S(t_{(i,k)}) \leq \prod_{i=0}^m S(t_{(i)})^{u_i} = \prod_{i=1}^m S(t_{(i)})^{u_i} \quad \text{となる。}$$

カプランマイヤー推定値の導出 (6)

$L(S)$ の最大値は

$$L^*(S) = \prod_{i=1}^m \left[(S(t_{(i)} - 0) - S(t_{(i)}))^{d_i} S(t_{(i)})^{u_i} \right] \quad \text{で考えればよい}$$

スライド9で示した通り,

$t_{(s)} \leq \forall t < t_{(s+1)}, t_{(s-1)} \leq \forall t - 0 < t_{(s)}$ に注意すると

$$S(t_{(i)}) = \prod_{k=1}^i (1 - \lambda_k), \quad S(t_{(i)} - 0) = \prod_{k=1}^{i-1} (1 - \lambda_k)$$

カプランマイヤー推定値の導出 (7)

$$L^*(S) = \prod_{i=1}^m \left[\left(\prod_{k=1}^{i-1} (1 - \lambda_k) - \prod_{k=1}^i (1 - \lambda_k) \right)^{d_i} \left(\prod_{k=1}^i (1 - \lambda_k) \right)^{u_i} \right]$$

$$= \prod_{i=1}^m \left[\lambda_i^{d_i} \prod_{k=1}^{i-1} (1 - \lambda_k)^{d_i} \prod_{k=1}^i (1 - \lambda_k)^{u_i} \right]$$

$$= \prod_{i=1}^m \lambda_i^{d_i} (1 - \lambda_i)^{u_i} \prod_{i=1}^m \prod_{k=1}^{i-1} (1 - \lambda_k)^{d_i+u_i}$$

$$= \prod_{i=1}^m \lambda_i^{d_i} (1 - \lambda_i)^{u_i} \prod_{k=1}^{m-1} \prod_{i=k+1}^m (1 - \lambda_k)^{d_i+u_i}$$

| | k | | | | |
|----------|----------|----------|----------|-----|---------|
| $i = 1$ | | | | | |
| $i = 2$ | 1 | | | | |
| $i = 3$ | 1 | 2 | | | |
| $i = 4$ | 1 | 2 | 3 | | |
| \vdots | \vdots | \vdots | \vdots | | |
| $i = m$ | 1 | 2 | 3 | ... | $m - 1$ |

カプランマイヤー推定値の導出 (8)

$$\begin{aligned}
 L^*(S) &= \prod_{i=1}^m \lambda_i^{d_i} (1 - \lambda_i)^{u_i} \prod_{k=1}^{m-1} \prod_{i=k+1}^m (1 - \lambda_k)^{d_i + u_i} \\
 &= \prod_{i=1}^m \lambda_i^{d_i} (1 - \lambda_i)^{u_i} \prod_{k=1}^{m-1} (1 - \lambda_k)^{\sum_{i=k+1}^m (d_i + u_i)} \\
 &= \prod_{i=1}^m \lambda_i^{d_i} (1 - \lambda_i)^{u_i} \prod_{k=1}^{m-1} (1 - \lambda_k)^{n_k - (d_k + u_k)} \\
 &= \left(\prod_{i=1}^m \lambda_i^{d_i} (1 - \lambda_i)^{u_i + n_i - (d_i + u_i)} \right) (1 - \lambda_m)^{-n_m + (d_m + u_m)} \\
 &= \prod_{i=1}^m \lambda_i^{d_i} (1 - \lambda_i)^{n_i - d_i} \quad (\because n_m = d_m + u_m)
 \end{aligned}$$

カプランマイヤー推定値の導出 (9)

$$L^*(S) = \prod_{i=1}^m \lambda_i^{d_i} (1 - \lambda_i)^{n_i - d_i}$$

は2項分布 m 個の積と同値なので, λ_i の最尤推定値は

$$\hat{\lambda}_i = \frac{d_i}{n_i} \quad (i = 1, \dots, m)$$

したがって, $t_{(s)} \leq \forall t < t_{(s+1)}$ に対して $S(t)$ の最尤推定値は

$$\hat{S}(t) = \prod_{i=1}^s (1 - \hat{\lambda}_i) = \prod_{i=1}^s \left(1 - \frac{d_i}{n_i}\right) \quad \text{カプランマイヤー推定値}$$

カプランマイヤー法の課題

- 最長の観察データ t^* が中途打ち切りの場合
 - 原著論文では, $t > t^*$ のカプランマイヤー推定値は定義されていない
 - 現在は3つの取り扱いがある

$$\hat{S}(t) = \begin{cases} \text{未定義} & \text{(Kaplan and Meier, 1958)} \\ 0 & \text{(Efron, 1967)} \\ \hat{S}(t^*) & \text{(Gill, 1980)} \end{cases}$$

それぞれの方法でバイアスの方向が変わるので注意が必要

具体例

□ カプランマイヤー推定量の適用結果

表1. 被験者4人分のデータ

| 被験者 番号 | 時間 | イベ ント |
|-----------|----|----------|
| 1 | 9 | 0 |
| 2 | 12 | 1 |
| 3 | 15 | 0 |
| 4 | 9 | 1 |

カプランマイヤー推定量

$$\hat{S}(t) = \prod_{i=1}^s \left(1 - \frac{d_i}{n_i}\right)$$

$$\hat{S}(t_{(1)} = 9) = \frac{4-1}{4} = 0.75$$

$$\hat{S}(t_{(2)} = 12) = \left(\frac{4-1}{4}\right) \left(\frac{2-1}{2}\right) = 0.375$$

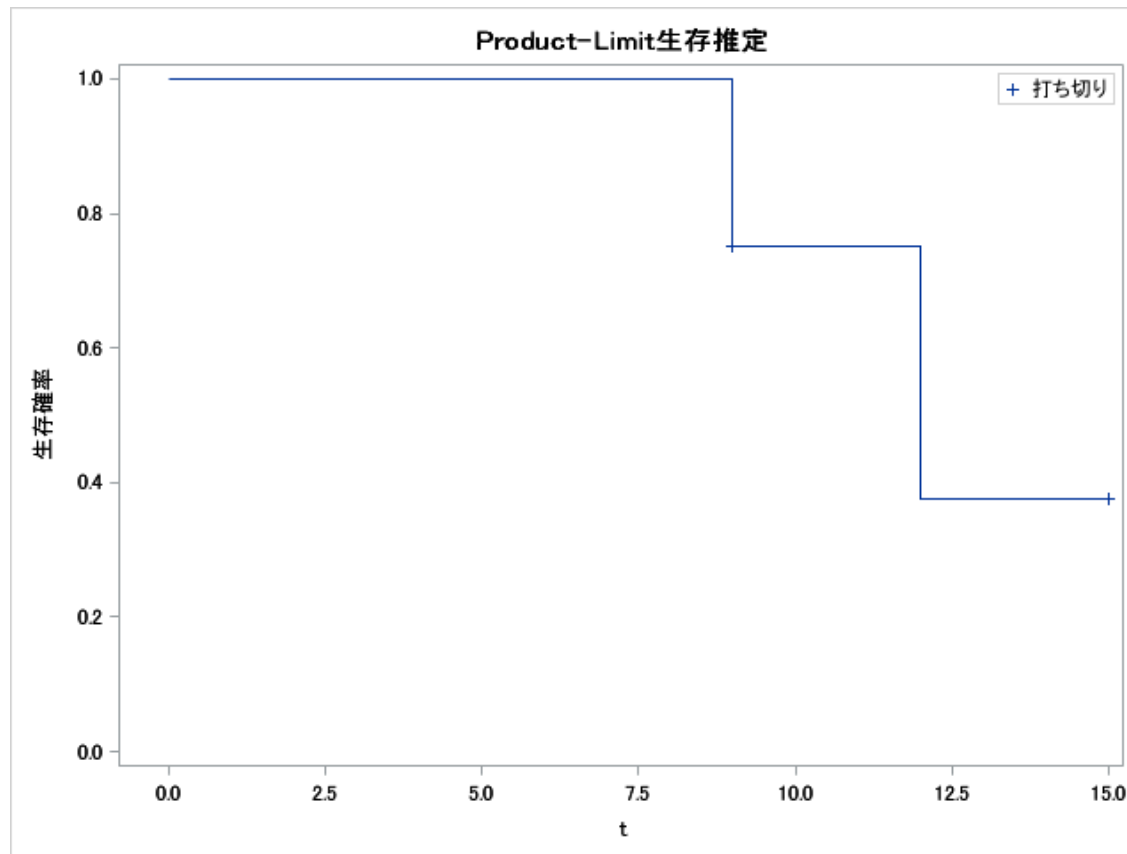
SAS プログラム

```
proc lifetest data = データセット名;  
  time time*c(0);  
run;
```

適用結果

□ 表1のデータに対する適用結果

■ SAS ではLIFETEST プロシジャで実行可能



| 生存関数の推定 | | | |
|---------|---|--------|-----|
| t | | 生存割合 | 生存数 |
| 0 | | 1.0000 | 4 |
| 9 | | 0.7500 | 3 |
| 9 | * | . | 2 |
| 12 | | 0.3750 | 1 |
| 15 | * | 0.3750 | 0 |

カプランマイヤー推定値の分散 (1)

$t_{(s)} \leq \forall t < t_{(s+1)}$ に対して

$$\log \hat{S}(t) = \log \prod_{i=1}^s (1 - \hat{\lambda}_i) = \sum_{i=1}^s \log(1 - \hat{\lambda}_i)$$

$d_i \sim \text{Bin}(n_i, \lambda_i)$ より

$$V(d_i) = n_i \lambda_i (1 - \lambda_i)$$

$$\begin{aligned} V(\log \hat{S}(t)) &= \sum_{i=1}^s V(\log(1 - \hat{\lambda}_i)) \\ &\simeq \sum_{i=1}^s \left(\frac{-1}{1 - \hat{\lambda}_i} \right)^2 V(1 - \hat{\lambda}_i) \\ &= \sum_{i=1}^s \left(\frac{1}{1 - \hat{\lambda}_i} \right)^2 V(\hat{\lambda}_i) \end{aligned} \quad \begin{aligned} &\rightarrow = \sum_{i=1}^s \left(\frac{1}{1 - \frac{d_i}{n_i}} \right)^2 V\left(\frac{d_i}{n_i}\right) \\ &= \sum_{i=1}^s \left(\frac{n_i}{n_i - d_i} \right)^2 \frac{1}{n_i^2} V(d_i) \\ &= \sum_{i=1}^s \left(\frac{1}{n_i - d_i} \right)^2 n_i \lambda_i (1 - \lambda_i) \end{aligned}$$

Kaplan-Meier推定値の分散 (2)

$$\begin{aligned}\hat{V}(\log \hat{S}(t)) &\simeq \sum_{i=1}^s \left(\frac{1}{n_i - d_i} \right)^2 n_i \frac{d_i}{n_i} \left(1 - \frac{d_i}{n_i} \right) \\ &= \sum_{i=1}^s \frac{d_i}{n_i(n_i - d_i)}\end{aligned}$$

さらに,

$$\hat{V}(\log \hat{S}(t)) \simeq \left(\frac{1}{\hat{S}(t)} \right)^2 \hat{V}(\hat{S}(t)) \quad \text{であることから}$$

$$\hat{V}(\hat{S}(t)) \simeq \left(\hat{S}(t) \right)^2 \hat{V}(\log \hat{S}(t))$$

したがって, $t_{(s)} \leq \forall t < t_{(s+1)}$ に対して

$$\hat{V}(\hat{S}(t)) \simeq \left(\hat{S}(t) \right)^2 \sum_{i=1}^s \frac{d_i}{n_i(n_i - d_i)}$$

Greenwoodの公式

生存関数の信頼区間

□ 生存関数の信頼区間

$$\left[\hat{S}(t) - z_{\alpha/2} \sqrt{\hat{V}(\hat{S}(t))}, \hat{S}(t) + z_{\alpha/2} \sqrt{\hat{V}(\hat{S}(t))} \right]$$

$z_{\alpha/2}$: 標準正規分布の上側100($\alpha/2$)% 点

- 標本サイズが小さい場合，正規分布の近似が悪く，区間[0,1]の範囲を逸脱することがある
- 何らかの変換 $g(S(t))$ の尺度上での信頼区間の構成が必要
- $g(\cdot) = \log(\cdot), \log(-\log(\cdot)), \sin^{-1}(\sqrt{\cdot})$ などが提案されている

生存関数の検定

□ 仮説 $\begin{cases} H_0: \text{すべての } t \geq 0 \text{ に対して, } S_1(t) = S_0(t) \\ H_1: \text{ある } t \text{ に対して, } S_1(t) \neq S_0(t) \end{cases}$

$S_1(t)$: 試験群の生存関数, $S_0(t)$: 対照群の生存関数

□ 分割表に基づいて計算

表2. $t_{(j)}$ 時点における生存関数の均一性検定に用いる表

| イベント/群 | 試験群: 1 | 対照群: 0 | 合計 |
|--------|-------------------|-------------------|-------------|
| あり | d_{1j} | d_{0j} | d_j |
| なし | $n_{1j} - d_{1j}$ | $n_{0j} - d_{0j}$ | $n_j - d_j$ |
| リスク集合 | n_{1j} | n_{0j} | n_j |

$t_{(j)}$ 時点における試験群と
対照群のリスク集合: n_{1j}, n_{0j}
各群のイベント数: d_{1j}, d_{0j}

ログランク検定

- 周辺和を $\{d_j, n_j - d_j, n_{1j}, n_{2j}\}$ を与えたとき，帰無仮説のもとで， d_{1j} (or d_{2j}) は超幾何分布に従う

$$\Pr(d_{1j}) = \frac{\binom{d_j}{d_{1j}} \binom{n_j - d_j}{n_{1j} - d_{1j}}}{\binom{n_j}{n_{1j}}} \quad (j = 1, \dots, m)$$

$$E_{1j} = E[d_{1j} \mid H_0] = \frac{n_{1j} d_j}{n_j} \quad V_{1j} = V[d_{1j} \mid H_0] = \frac{n_{1j} d_j (n_j - d_j) (n_j - n_{1j})}{n_j^2 (n_j - 1)}$$

- Mantel-Haenszel統計量

$$Z_{\text{MH}} = \frac{\sum_{j=1}^m (d_{1j} - E_{1j})}{\sqrt{\sum_{j=1}^m V_{1j}}} \approx N(0, 1)$$

Z_{MH} に基づく Mantel-Haenszel 検定は
ログランク検定と呼ばれている

ロ グ ラ ン ク 検 定 の 特 徴

□ 比例ハザード性

- ハザード比が時間に依存せず一定

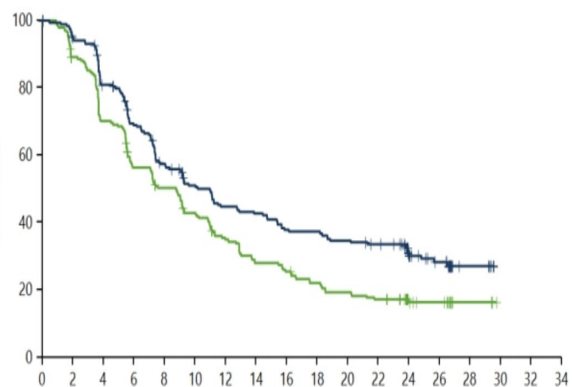
$$HR = \frac{h_1(t)}{h_0(t)} = \text{定数}$$

$h_1(t)$:試験群のハザード関数, $h_0(t)$:対照群のハザード関数

- 比例ハザード性が成り立つ場合, ロ グ ラ ン ク 検 定 は 一 様 最 強 力 検 定 と な る

比例ハザード性の仮定

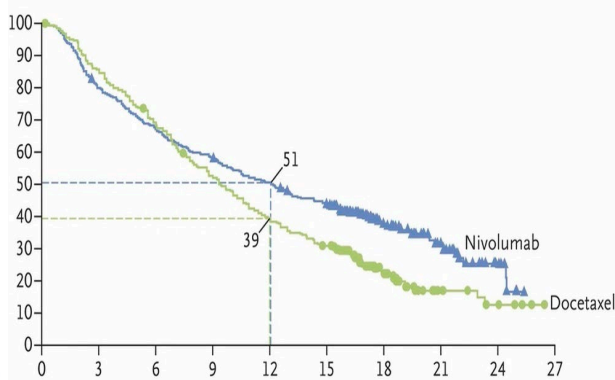
比例ハザード性が
仮定できる事例



Kulkarni et al. (2016)

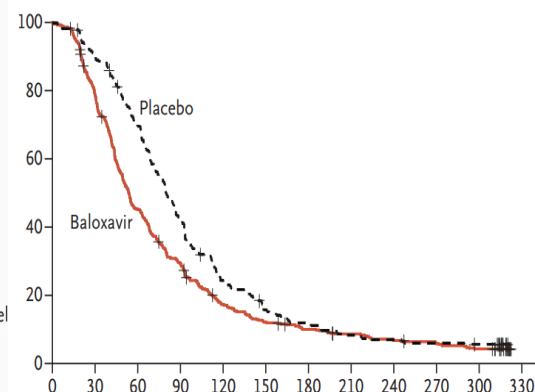
比例ハザード性が仮定できない事例

遅発効果がある場合



Borghaei et al. (2015)

試験早期に
群間差が生じる場合



Hayden et al. (2018)

増加傾向

ログランク検定に代わる生存関数の検定法が必要

重み付きログランク検定

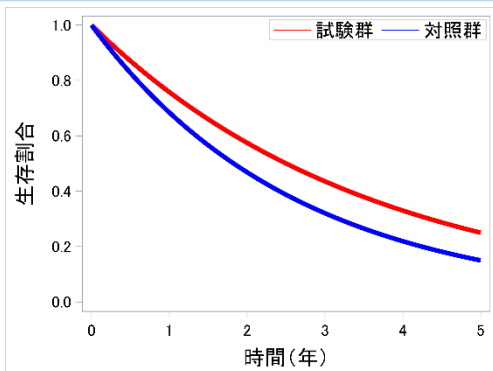
□ 検定統計量

$$Z_{\text{WLR}} = \frac{\sum_{j=1}^m w_j (d_{1j} - E_{1j})}{\sqrt{\sum_{j=1}^m w_j^2 V_{1j}}} \approx N(0,1)$$

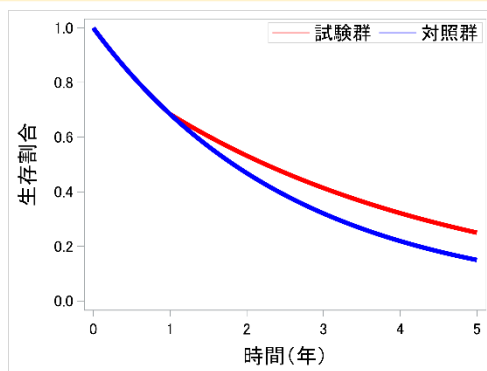
□ $w_j = 1$ と場合， ログランク検定に一致

重みの設定

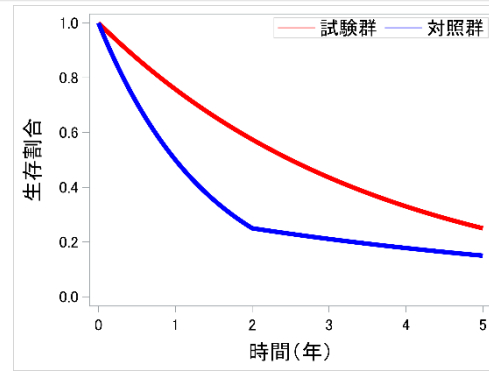
比例ハザード性が
仮定できる場合



試験群に遅発効果がある場合



試験早期から
群間差が生じる場合



検定

ログランク
(LR)

フレミング・
ハリントン(FH)

一般化ウィル
コクソン(GW)

時点の
重み

1

$$S(t)^p (1 - S(t))^q$$

n_j

比例ハザード性の
もとで最強力検定

p, q を任意に設定

試験早期の
群間差を重視

免疫療法に適した統計解析法

Statistics
in Medicine

Research Article

Received 29 March 2016,

Accepted 2 October 2016

Published online 2 November 2016 in Wiley Online Library

(wileyonlinelibrary.com) DOI: 10.1002/sim.7157

Designing therapeutic cancer vaccine trials with delayed treatment effect

Zhenzhen Xu,^{a,*†} Boguang Zhen,^a Yongsoek Park^b and Bin Zhu^c

Arming the immune system against cancer has emerged as a powerful tool in oncology during recent years. Instead of poisoning a tumor or destroying it with radiation, therapeutic cancer vaccine, a type of cancer immunotherapy, unleashes the immune system to combat cancer. This indirect mechanism-of-action of vaccines poses the possibility of a delayed onset of clinical effect, which results in a delayed separation of survival curves between the experimental and control groups in therapeutic cancer vaccine trials with time-to-event endpoints. This violates the proportional hazard assumption. As a result, the conventional study design based on the regular log-rank test ignoring the delayed effect would lead to a loss of power. In this paper, we propose two innovative approaches for sample size and power calculation using the piecewise weighted log-rank test to properly and efficiently incorporate the delayed effect into the study design. Both theoretical derivations and empirical studies demonstrate that the proposed methods, accounting for the delayed effect, can reduce sample size dramatically while achieving the target power relative to a standard practice. Copyright © 2016 John Wiley & Sons, Ltd.

Keywords: delayed treatment effect; therapeutic cancer vaccine; cancer immunotherapy; sample size and power calculation; non-proportional hazard assumption; cancer clinical trial

区分重み付きロラング検定

□ 仮説

$$H_0: h_1(t) = h_0(t) \quad H_1: \frac{h_1(t)}{h_0(t)} = \begin{cases} \text{HR}_1, & t \leq t^* \\ \text{HR}_2, & t > t^* \end{cases}, \quad \text{HR}_2 < \text{HR}_1 \leq 1$$

t^* : 効果発現までの時間

□ 検定統計量

$$Q_{PWLT} = \frac{\sum_{j \in D_1} w_1 \{d_{1j} - \hat{e}_{1j}\} + \sum_{j \in D_2} w_2 \{d_{1j} - \hat{e}_{1j}\}}{\left[\sum_{j \in D_1} w_1^2 \hat{e}_{1j} \{1 - \hat{e}_{1j}\} + \sum_{j \in D_2} w_2^2 \hat{e}_{1j} \{1 - \hat{e}_{1j}\} \right]^{\frac{1}{2}}}$$

D_1 : t^* 前に発生したイベント

D_2 : t^* 後に発生したイベント

w_1 : D_1 における重み

w_2 : D_2 における重み

■ $hr_1 = 1$ のとき, $w_1 = 0, w_2 = 1$ が最適な重み

ハザード比 (Hazard Ratio)

□ 2群のハザード関数の比

$$\text{HR} = \frac{h_1(t)}{h_0(t)}$$

$h_1(t)$: 試験群のハザード関数

$h_0(t)$: 対照群のハザード関数

□ 比例ハザード性

- ハザード比が時間に依存せず一定

$$\text{HR} = \frac{h_1(t)}{h_0(t)} = \text{定数}$$

比例ハザード回帰モデル

□ 生存時間解析における回帰分析

- 生存時間に関係すると考えられる変数（共変量）が利用可能なときに，生存時間への影響を要約・評価する

□ 比例ハザード回帰モデル

$$\lambda(t|\mathbf{x}) = \lambda_0(t)\exp(\mathbf{x}'\boldsymbol{\beta})$$

$$= \lambda_0(t)\exp(\beta_1 x_1 + \cdots + \beta_p x_p)$$

$\lambda(t|\mathbf{x})$: ハザード関数

$\lambda_0(t)$: 規準ハザード関数

$\mathbf{x} = (x_1, \dots, x_p)'$: 共変量

$\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)'$: 係数パラメータ

比例ハザード回帰モデルのもとでの生存関数

□ 生存関数

$$S(t|\mathbf{x}) = \exp\left(-\int_0^t \lambda(u|\mathbf{x})du\right)$$

$$= \exp\left(-\int_0^t \lambda_0(u)\exp(\mathbf{x}'\boldsymbol{\beta})du\right)$$

$$= \exp\left(-\int_0^t \lambda_0(u)du\right)^{\exp(\mathbf{x}'\boldsymbol{\beta})}$$

$$= S_0(t)^{\exp(\mathbf{x}'\boldsymbol{\beta})}$$

$S_0(t)$: 規準生存関数

比例ハザード回帰モデルのもとでのハザード比

□ 2群のハザード関数の比

■ 比例ハザード回帰モデルのもとで

$$\text{HR} = \frac{\lambda(t|\mathbf{x}_A)}{\lambda(t|\mathbf{x}_B)}$$

\mathbf{x}_A : 群 A の共変量

\mathbf{x}_B : 群 B の共変量

$$= \frac{\lambda_0(t)\exp(\mathbf{x}'_A\boldsymbol{\beta})}{\lambda_0(t)\exp(\mathbf{x}'_B\boldsymbol{\beta})}$$

$$= \exp((\mathbf{x}'_A - \mathbf{x}'_B)\boldsymbol{\beta}) \quad \leftarrow \text{時間 } t \text{ に依存せず一定}$$

比例ハザード回帰モデルは比例ハザード性を仮定している

ハザード比の解釈

□ ハザード比

■ 比例ハザード回帰モデルのもとで

$$\text{HR} = \frac{\lambda(t|\mathbf{x}_A)}{\lambda(t|\mathbf{x}_B)} \quad \begin{array}{l} \mathbf{x}_A: \text{群 } A \text{ の共変量} \\ \mathbf{x}_B: \text{群 } B \text{ の共変量} \end{array}$$

$$\text{HR} = 1 \Leftrightarrow \lambda(t|\mathbf{x}_A) = \lambda(t|\mathbf{x}_B) \Leftrightarrow S(t|\mathbf{x}_A) = S(t|\mathbf{x}_B)$$

$$\text{HR} < 1 \Leftrightarrow \lambda(t|\mathbf{x}_A) < \lambda(t|\mathbf{x}_B) \Leftrightarrow S(t|\mathbf{x}_A) > S(t|\mathbf{x}_B)$$

$$\text{HR} > 1 \Leftrightarrow \lambda(t|\mathbf{x}_A) > \lambda(t|\mathbf{x}_B) \Leftrightarrow S(t|\mathbf{x}_A) < S(t|\mathbf{x}_B)$$

ハザード比の特徴

□ 長所

- 比例ハザード性の仮定のもとで、群間差の妥当な要約指標として頻用されている
- データが豊富に存在する

□ 短所

- 指標として直感的に解釈しづらい
- 比例ハザード性の仮定が成り立たないと解釈が困難になる場合がある

共変量で調整した2群の生存関数の比較 (1)

□ 群 A と B の比較を共変量の影響を調整して行いたい

■ 仮説

$$\begin{cases} H_0: \text{すべての } t \geq 0 \text{ に対して, } S(t|\mathbf{x}_A) = S(t|\mathbf{x}_B) \\ H_1: \text{ある } t \text{ に対して, } S(t|\mathbf{x}_A) \neq S(t|\mathbf{x}_B) \end{cases}$$

■ x_1 : 群を表す変数

$$x_1 = \begin{cases} 1 & (A \text{ 群}) \\ 0 & (B \text{ 群}) \end{cases}$$

■ x_2, \dots, x_p : 共変量 (交絡因子, 予後因子など)

共変量の影響を調整する \Leftrightarrow 群を表す変数 x_1 以外の変数は
同じ値をもつもとで, 群を比較する 36

共変量で調整した2群の生存関数の比較 (2)

□ 群 A と B の比較を共変量の影響を調整して行いたい

共変量の影響を調整する \Leftrightarrow 群を表す変数 x_1 以外の変数は
同じ値をもつもとで、群を比較する

$$\begin{aligned}\text{HR} &= \frac{\lambda(t|\mathbf{x}_A)}{\lambda(t|\mathbf{x}_B)} = \frac{\lambda_0(t)\exp(\mathbf{x}'_A\boldsymbol{\beta})}{\lambda_0(t)\exp(\mathbf{x}'_B\boldsymbol{\beta})} \\ &= \frac{\exp(1 \times \beta_1 + x_2\beta_2 + \cdots + x_p\beta_p)}{\exp(0 \times \beta_1 + x_2\beta_2 + \cdots + x_p\beta_p)} \\ &= \frac{\exp(\beta_1 + c)}{\exp(c)} = \exp(\beta_1)\end{aligned}$$

$$x_1 = \begin{cases} 1 & (A \text{ 群}) \\ 0 & (B \text{ 群}) \end{cases}$$

$$c = x_2\beta_2 + \cdots + x_p\beta_p$$

共変量で調整した2群の生存関数の比較 (3)

□ 群 A と B の比較を共変量の影響を調整して行いたい

共変量の影響を調整する \Leftrightarrow 群を表す変数 x_1 以外の変数は
同じ値をもつもとで、群を比較する

$$\frac{S(t|\mathbf{x}_A)}{S(t|\mathbf{x}_B)} = \frac{S_0(t)^{\exp(\mathbf{x}'_A\boldsymbol{\beta})}}{S_0(t)^{\exp(\mathbf{x}'_B\boldsymbol{\beta})}}$$

$$= S_0(t)^{\exp(\mathbf{x}'_A\boldsymbol{\beta} - \mathbf{x}'_B\boldsymbol{\beta})}$$

$$= S_0(t)^{\exp(c)(\exp(\beta_1) - 1)}$$

$$x_1 = \begin{cases} 1 & (A \text{ 群}) \\ 0 & (B \text{ 群}) \end{cases}$$

$$c = x_2\beta_2 + \cdots + x_p\beta_p$$

共変量で調整した2群の生存関数の比較 (4)

□ 比例ハザード回帰モデルのもとで

■ 仮説

$$\begin{cases} H_0: \text{すべての } t \geq 0 \text{ に対して, } S(t|\mathbf{x}_A) = S(t|\mathbf{x}_B) \\ H_1: \text{ある } t \text{ に対して, } S(t|\mathbf{x}_A) \neq S(t|\mathbf{x}_B) \end{cases} \Leftrightarrow \begin{cases} H_0: \beta_1 = 0 \\ H_1: \beta_1 \neq 0 \end{cases}$$

比例ハザード回帰モデルの分類

□ パラメトリックモデル

- 規準ハザード関数 $\lambda_0(t)$ に特定の確率分布を仮定する方法
- 指数分布, **ワイブル分布**

□ セミパラメトリック (Cox) モデル

- 規準ハザード関数 $\lambda_0(t)$ に特定の確率分布を仮定しない方法
- $\lambda_0(t)$ は未知とし, β に関する推測を行う
- 条件付き推測により, 局外母数ともいえる $\lambda_0(t)$ を推定方程式から除く必要がある

Cox比例ハザード回帰モデル (1)

□ 2種類の事象 A, B の列を考える

$$\{A_1, A_2, \dots, A_m\}, \quad \{B_0, B_1, B_2, \dots, B_m\}$$

$$A_j = \{\text{時点 } t_{(j)} \text{ でイベントが発生}\}$$

$$B_j = \{\text{区間 } [t_{(j)}, t_{(j+1)}) \text{ で中途打ち切りが発生}\}$$

□ データ全体の尤度（確率）

$$\Pr(B_0 \cap A_1 \cap B_1 \cap A_2 \cap B_2 \cap \dots \cap A_m \cap B_m)$$

Cox比例ハザード回帰モデル (2)

□ 条件付き確率を順次求める

$$\begin{aligned}
 & \Pr(B_0 \cap A_1 \cap B_1 \cap A_2 \cap B_2 \cap \cdots \cap A_m \cap B_m) \\
 &= \Pr(B_0 \cap A_1) \Pr(B_1 \cap A_2 \cap B_2 \cap \cdots \cap A_m \cap B_m \mid B_0 \cap A_1) \\
 &= \Pr(B_0 \cap A_1) \Pr(B_1 \cap A_2 \mid B_0 \cap A_1) \Pr(B_2 \cap A_3 \cap B_3 \cap \cdots \mid B_0 \cap A_1 \cap B_1 \cap A_2) \\
 &= \cdots = \Pr(B_0 \cap A_1) \left[\prod_{j=2}^m \Pr(B_{j-1} \cap A_j \mid B_0 \cap A_1 \cap B_1 \cap A_2 \cap \cdots \cap B_{i-2} \cap A_{j-1}) \right] \\
 &\quad \times \Pr(B_m \mid B_0 \cap A_1 \cap B_1 \cap A_2 \cap \cdots \cap B_{m-1} \cap A_m)
 \end{aligned}$$

Cox比例ハザード回帰モデル (3)

$$\begin{aligned}
 & \Pr(B_0 \cap A_1 \cap B_1 \cap A_2 \cap B_2 \cap \cdots \cap A_m \cap B_m) \\
 &= \Pr(B_0 \cap A_1) \left[\prod_{j=2}^m \Pr(B_{j-1} \cap A_j \mid B_0 \cap A_1 \cap B_1 \cap A_2 \cap \cdots \cap B_{j-2} \cap A_{j-1}) \right] \\
 & \quad \times \Pr(B_m \mid B_0 \cap A_1 \cap B_1 \cap A_2 \cap \cdots \cap B_{m-1} \cap A_m) \\
 &= \Pr(A_1 \mid B_0) \Pr(B_0) \left[\prod_{j=2}^m \Pr(A_j \mid B_0 \cap A_1 \cap \cdots \cap B_{j-2} \cap A_{j-1} \cap B_{j-1}) \right. \\
 & \quad \left. \times \Pr(B_{j-1} \mid B_0 \cap A_1 \cap \cdots \cap B_{j-2} \cap A_{j-1}) \right] \times \Pr(B_m \mid B_0 \cap A_1 \cap \cdots \cap B_{m-1} \cap A_m)
 \end{aligned}$$

Cox比例ハザード回帰モデル (4)

$$\begin{aligned} & \Pr(B_0 \cap A_1 \cap B_1 \cap A_2 \cap B_2 \cap \cdots \cap A_m \cap B_m) \\ &= \Pr(A_1 | B_0) \left[\prod_{j=2}^m \Pr(A_j | B_0 \cap A_1 \cap \cdots \cap B_{j-2} \cap A_{j-1} \cap B_{j-1}) \right] \\ & \quad \times \Pr(B_0) \left[\prod_{j=2}^{m+1} \Pr(B_{j-1} | B_0 \cap A_1 \cap \cdots \cap B_{j-2} \cap A_{j-1}) \right] \end{aligned}$$

中途打ち切りである事象 B_j は生存関数の推定には
ほとんど情報をもっていない



回帰パラメータ $\boldsymbol{\beta}$ にもほとんど情報がないと
考えても問題はないと考える

回帰パラメータ $\boldsymbol{\beta}$ に関する尤度

□ 尤度全体から事象 A_j に関する部分尤度を考える

$$PL(\boldsymbol{\beta}) = \Pr(A_1|B_0) \left[\prod_{j=2}^m \Pr(A_j \mid B_0 \cap A_1 \cap \cdots \cap B_{j-2} \cap A_{j-1} \cap B_{j-1}) \right]$$

□ $\boldsymbol{\beta}$ に関する推測

- 生存時間にタイ（同時間）がない場合 ($d_j = 1 \ (j = 1, \dots, m)$)
- 生存時間にタイがある場合

生存時間にタイがない場合

■ 時点 $t_{(j)}$ のイベント数 $d_j = 1$ ($j = 1, \dots, m$) の場合

$$\begin{aligned} & \Pr(A_j \mid B_0 \cap A_1 \cap \dots \cap B_{j-2} \cap A_{j-1} \cap B_{j-1}) \\ &= \Pr(\text{リスク集合 } n_j \text{ という条件のもとで, 1例に時点 } t_{(j)} \text{ でイベント発生}) \end{aligned}$$

$$= \frac{\Pr(T = t_{(j)} \mid T > t_{(j-1)}, \mathbf{x}_{(j)})}{\sum_{k \in R(t_{(j)})} \Pr(T = t_{(j)} \mid T > t_{(j-1)}, \mathbf{x}_k)}$$

$$= \frac{\lambda(t_{(j)} \mid \mathbf{x}_{(j)})}{\sum_{k \in R(t_{(j)})} \lambda(t_{(j)} \mid \mathbf{x}_k)}$$

$$= \frac{\exp(\mathbf{x}'_{(j)} \boldsymbol{\beta})}{\sum_{k \in R(t_{(j)})} \exp(\mathbf{x}'_k \boldsymbol{\beta})}$$

$t_{(j)}$: 順序付けられた生存時間 $t_{(1)} < t_{(2)} < \dots < t_{(m)}$

$R(t_{(j)})$: 時点 $t_{(j)}$ のリスク集合にある参加者の集合

\mathbf{x}_k : 参加者 k の共変量 $\mathbf{x}_k = (x_{1k}, \dots, x_{pk})'$

$\mathbf{x}_{(j)}$: 生存時間が時点 $t_{(j)}$ と観測された参加者の共変量

生存時間にタイがない場合の部分尤度

□ 部分尤度

$$\begin{aligned}
 \text{PL}(\boldsymbol{\beta}) &= \Pr(A_1 | B_0) \left[\prod_{j=2}^m \Pr(A_j | B_0 \cap A_1 \cap \cdots \cap B_{j-2} \cap A_{j-1} \cap B_{j-1}) \right] \\
 &= \prod_{j=1}^m \Pr(A_j | B_0 \cap A_1 \cap \cdots \cap B_{j-2} \cap A_{j-1} \cap B_{j-1}) \\
 &= \prod_{j=1}^m \frac{\exp(\mathbf{x}'_{(j)} \boldsymbol{\beta})}{\sum_{k \in R(t_{(j)})} \exp(\mathbf{x}'_k \boldsymbol{\beta})}
 \end{aligned}$$

生存時間にタイがある場合 (1)

■ 時点 $t_{(j)}$ のイベント数が d_j ($j = 1, \dots, m$) の場合

$$\begin{aligned} & \Pr(A_j \mid B_0 \cap A_1 \cap \dots \cap B_{j-2} \cap A_{j-1} \cap B_{j-1}) \\ &= \Pr(\text{リスク集合 } n_j \text{ という条件のもとで, } d_j \text{ 例に時点 } t_{(j)} \text{ でイベント発生}) \\ &= \frac{\prod_{k=1}^{d_j} \Pr(T = t_{(j)} \mid T > t_{(j-1)}, \mathbf{x}_{(j_k)})}{\sum_{(i_1, \dots, i_{d_j}) \in R(t_{(j)}, d_j)} \prod_{k=1}^{d_j} \Pr(T = t_{(j)} \mid T > t_{(j-1)}, \mathbf{x}_{(i_k)})} \\ &= \frac{\prod_{k=1}^{d_j} \lambda(t_{(j)} \mid \mathbf{x}_{(j_k)})}{\sum_{(i_1, \dots, i_{d_j}) \in R(t_{(j)}, d_j)} \prod_{k=1}^{d_j} \lambda(t_{(j)} \mid \mathbf{x}_{(i_k)})} \end{aligned}$$

$R(t_{(j)}, d_j)$: リスク集合 n_j のなかから d_j 例を選ぶ添字の組み合わせ

$\mathbf{x}_{(j_k)}$: 生存時間が時点 $t_{(j)}$ と観測された参加者 k の共変量

生存時間にタイがある場合 (2)

$$\begin{aligned}
 & \Pr(A_j \mid B_0 \cap A_1 \cap \cdots \cap B_{j-2} \cap A_{j-1} \cap B_{j-1}) \\
 &= \frac{\prod_{k=1}^{d_j} \exp(\mathbf{x}'_{(j_k)} \boldsymbol{\beta})}{\sum_{(i_1, \dots, i_{d_j}) \in R(t_{(j)}, d_j)} \prod_{k=1}^{d_j} \exp(\mathbf{x}'_{(i_k)} \boldsymbol{\beta})} \\
 &= \frac{\exp(\sum_{k=1}^{d_j} \mathbf{x}'_{(j_k)} \boldsymbol{\beta})}{\sum_{(i_1, \dots, i_{d_j}) \in R(t_{(j)}, d_j)} \exp(\sum_{k=1}^{d_j} \mathbf{x}'_{(i_k)} \boldsymbol{\beta})} \\
 &= \frac{\exp(\mathbf{s}'_{(j)} \boldsymbol{\beta})}{\sum_{(i_1, \dots, i_{d_j}) \in R(t_{(j)}, d_j)} \exp(\sum_{k=1}^{d_j} \mathbf{x}'_{(i_k)} \boldsymbol{\beta})}
 \end{aligned}$$

$$\mathbf{s}_{(j)} = \mathbf{x}_{(j_1)} + \mathbf{x}_{(j_2)} + \cdots + \mathbf{x}_{(j_{d_j})}$$

生存時間にタイがある場合の部分尤度

□（正確な）部分尤度

$$\begin{aligned}
 \text{PL}(\boldsymbol{\beta}) &= \Pr(A_1 | B_0) \left[\prod_{j=2}^m \Pr(A_j \mid B_0 \cap A_1 \cap \cdots \cap B_{j-2} \cap A_{j-1} \cap B_{j-1}) \right] \\
 &= \prod_{j=1}^m \Pr(A_j \mid B_0 \cap A_1 \cap \cdots \cap B_{j-2} \cap A_{j-1} \cap B_{j-1}) \\
 &= \prod_{j=1}^m \frac{\exp(\mathbf{s}'_{(j)} \boldsymbol{\beta})}{\sum_{(i_1, \dots, i_{d_j}) \in R(t_{(j)}, d_j)} \exp(\sum_{k=1}^{d_j} \mathbf{x}'_{(i_k)} \boldsymbol{\beta})}
 \end{aligned}$$

部分尤度の計算

- タイの数が多いと部分尤度は、組み合わせ $\binom{n_j}{d_j}$ の計算が必要となるので、計算が困難になる

- 近似の部分尤度が提案されている
 - Breslow (1974) SASのデフォルト
 - Efron (1977) Rのデフォルト
 - Cox (1972)

Breslowの近似部分尤度

□ 正確な部分尤度の最も単純な近似方法

$$PL(\boldsymbol{\beta}) = \prod_{j=1}^m \frac{\exp(\mathbf{s}'_{(j)}\boldsymbol{\beta})}{\sum_{(i_1, \dots, i_{d_j}) \in R(t_{(j)}, d_j)} \exp(\sum_{k=1}^{d_j} \mathbf{x}'_{(i_k)}\boldsymbol{\beta})} \quad \text{正確な部分尤度}$$

$$PL^*(\boldsymbol{\beta}) = \prod_{j=1}^m \frac{\exp(\mathbf{s}'_{(j)}\boldsymbol{\beta})}{\left(\sum_{k \in R(t_{(j)})} \exp(\mathbf{x}'_k\boldsymbol{\beta}) \right)^{d_j}} \quad \text{近似部分尤度}$$

時点 $t_{(j)}$ の d_j 人のイベントは，区別できるもので順に発生していると仮定している

タイである観測値がそれほど多くなければ，部分尤度の適切な近似となる

Efronの近似部分尤度

■ Breslowよりも精度の高い近似方法

$$PL(\boldsymbol{\beta}) = \prod_{j=1}^m \frac{\exp(\mathbf{s}'_{(j)}\boldsymbol{\beta})}{\sum_{\left(i_1, \dots, i_{d_j}\right) \in R(t_{(j)}, d_j)} \exp\left(\sum_{k=1}^{d_j} \mathbf{x}'_{(i_k)}\boldsymbol{\beta}\right)}$$

$$PL^*(\boldsymbol{\beta}) = \prod_{j=1}^m \frac{\exp(\mathbf{s}'_{(j)}\boldsymbol{\beta})}{\prod_{k=1}^{d_j} \left[\sum_{l \in R(t_{(j)})} \exp(\mathbf{x}'_l \boldsymbol{\beta}) - (k-1)d_j^{-1} \sum_{l \in D(t_{(j)})} \exp(\mathbf{x}'_l \boldsymbol{\beta}) \right]}$$

$D(t_{(j)})$: 時点 $t_{(j)}$ でイベントが発生した参加者の集合

実際には多くの場合でBreslowの近似と似たような結果となる

Coxの近似部分尤度

- 生存時間の観測が連続的ではなく，本質的に離散的である状況に対するモデルに基づく近似方法

$$PL(\boldsymbol{\beta}) = \prod_{j=1}^m \frac{\exp(\mathbf{s}'_{(j)}\boldsymbol{\beta})}{\sum_{(i_1, \dots, i_{d_j}) \in R(t_{(j)}, d_j)} \exp(\sum_{k=1}^{d_j} \mathbf{x}'_{(i_k)}\boldsymbol{\beta})}$$

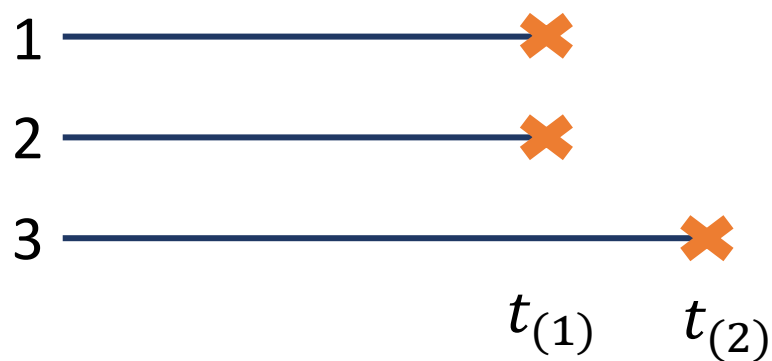
$$PL^*(\boldsymbol{\beta}) = \prod_{j=1}^m \frac{\exp(\mathbf{s}'_{(j)}\boldsymbol{\beta})}{\sum_{k \in R(t_{(j)}, d_j)} \exp(\mathbf{s}'_{(k)}\boldsymbol{\beta})}$$

タイがない場合の部分尤度と同様

比例ハザードモデルではなく，離散ロジスティックモデルに対応する推定法

このモデルのもとでは，タイである観測値は許容される

近似部分尤度の違い (1)



時点 $t_{(1)}$ で個体1と2のイベントが同時に発生している

個体1にイベントが発生，次に個体2にイベントが発生

$$\frac{h_1}{h_1 + h_2 + h_3} \times \frac{h_2}{h_2 + h_3} \times \frac{h_3}{h_3}$$

h_i : 個体 i のハザード

個体2にイベントが発生，次に個体1にイベントが発生

$$\frac{h_2}{h_1 + h_2 + h_3} \times \frac{h_1}{h_1 + h_3} \times \frac{h_3}{h_3}$$

正確な部分尤度

$$PL(\boldsymbol{\beta}) = \left(\frac{h_1}{h_1 + h_2 + h_3} \times \frac{h_2}{h_2 + h_3} + \frac{h_2}{h_1 + h_2 + h_3} \times \frac{h_1}{h_1 + h_3} \right) \times \frac{h_3}{h_3}$$

近似部分尤度の違い (2)

正確な部分尤度

$$\text{PL}(\boldsymbol{\beta}) = \left(\frac{h_1}{h_1 + h_2 + h_3} \times \frac{h_2}{h_2 + h_3} + \frac{h_2}{h_1 + h_2 + h_3} \times \frac{h_1}{h_1 + h_3} \right) \times \frac{h_3}{h_3}$$

Breslowの近似部分尤度

$$\text{PL}^*(\boldsymbol{\beta}) = \frac{h_1 \times h_2}{(h_1 + h_2 + h_3)^2} \times \frac{h_3}{h_3}$$

Coxの近似部分尤度

$$\text{PL}^*(\boldsymbol{\beta}) = \frac{h_1 \times h_2}{h_1 \times h_2 + h_2 \times h_3 + h_3 \times h_1} \times \frac{h_3}{h_3}$$

Efronの近似部分尤度

$$\text{PL}^*(\boldsymbol{\beta}) = \frac{h_1 \times h_2}{(h_1 + h_2 + h_3) \left(\frac{h_1}{2} + \frac{h_2}{2} + h_3 \right)} \times \frac{h_3}{h_3}$$

比例ハザード性が成り立つかを確認する方法

□ 比例ハザード回帰モデルのもとでの生存関数

$$S(t|\mathbf{x}) = S_0(t)^{\exp(\mathbf{x}'\boldsymbol{\beta})} \quad S_0(t): \text{規準生存関数}$$

■ 二重対数変換 $\log(-\log(\cdot))$

$$\begin{aligned} \log(-\log S(t|\mathbf{x})) &= \log(-\log S_0(t)^{\exp(\mathbf{x}'\boldsymbol{\beta})}) \\ &= \log(-\exp(\mathbf{x}'\boldsymbol{\beta})\log S_0(t)) \\ &= \log(\exp(\mathbf{x}'\boldsymbol{\beta})) + \log(-\log S_0(t)) \\ &= \mathbf{x}'\boldsymbol{\beta} + \log(-\log S_0(t)) \end{aligned}$$

比例ハザードモデルのもとでは、生存関数 $S(t|\mathbf{x})$ の $\log(-\log(\cdot))$ 尺度上において、層間でプロットが平行になる

ログランク検定との関係

□ 比例ハザード回帰モデル

$$\lambda(t|x_1) = \lambda_0(t)\exp(x_1\beta_1) \quad x_1 = \begin{cases} 1 & (A \text{ 群}) \\ 0 & (B \text{ 群}) \end{cases}$$

□ 仮説

$$\begin{cases} H_0: \text{すべての } t \geq 0 \text{ に対して, } S_A(t) = S_B(t) \\ H_1: \text{ある } t \text{ に対して, } S_A(t) \neq S_B(t) \end{cases} \Leftrightarrow \begin{cases} H_0: \beta_1 = 0 \\ H_1: \beta_1 \neq 0 \end{cases}$$

この仮説のスコア検定はログランク検定に一致する