生存時間解析

北海道大学　医学統計学
横田　勲

1

今回の内容

2

- Kaplan-Meier法
- ログランク検定
- Cox比例ハザードモデル
  - ハザード比一定という仮定
- 無情報な打ち切り
- 時間依存性共変量

2

例：Gehanの白血病データ

3

- プラセボ群の再発までの時間(week)
  - 1, 1, 2, 2, 3, 4, 4, 5, 5, 8, 8, 8, 8, 11, 11, 12, 12, 15, 17, 22, 23　(n=21)
- 6-MP群の再発までの時間(week)
  - 6*, 6, 6, 6, 7, 9*, 10*, 10, 11*, 13, 16, 17*, 19*, 20*, 22, 23, 25*, 32*, 32*, 34*, 35* (n=21)
    - *付はその時点で追跡不能となった

3

t検定をやってみよう

4

- 時間は連続的だし・・
  - t検定は左右対称な分布のときなら大体OK
  - 時間データはたいてい右すそが重い
- 追跡不能例もいるけど、まあいいや
  - まあよくないよね？
- t=3.24, 自由度40のt分布よりp=0.002！
  - 再発までの時間の平均が遅くなった？
  - 他の違いを検出しただけ？

4

カイ二乗検定をやってみよう

5

- 追跡不能ってことは、再発しなかったとみなしてしまおう
  - もっと追跡したら再発したのでは？
- 2x2の分割表

	再発した	再発しなかった
プラセボ	21	0
6-MP	9	12

  - $\chi^2 = 16.8, p < 0.001$
  - 時間の長短に関する情報を使っていない

5

Time-to-event アウトカム

6

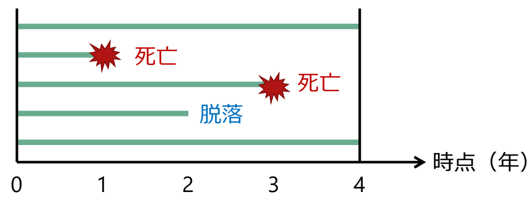
- 連続量、カテゴリカルのほか、医学研究でよく登場するアウトカム
- あらかじめ定義した「イベント」が起こるまでの時間
  - 死亡、再発、入院、治癒、寛解など
    - のぞましくない場合も、のぞましい場合も
  - at risk：まだイベントを起こしていない状態

6

追跡データ

7

- 4人目の正確な死亡時点がわからない
  - （右側）打ち切りデータ

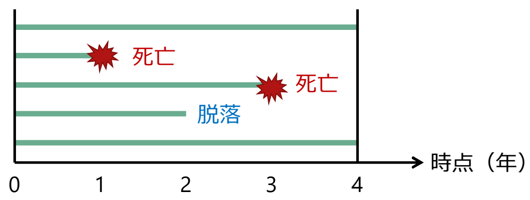


7

発生率

8

- 単位時間あたりの発生
  - のべ観察した人数に対する発生数
    観察人時間
- $\frac{2}{4+1+3+2+4} = 0.14(/年)$



8

打ち切りのあるデータ

9

- ある時点までイベントを起こしていない
- その先で起こるはずのイベントの
  正確な時点が分からない
  - 脱落や研究終了等による
- 適切に考慮する解析方法が生存時間解析
  - 単に除外すると有病率を過大評価しがち
  - イベントなしとすると有病率を過小評価
  - 無情報な打ち切りの仮定

9

生存時間データの要約

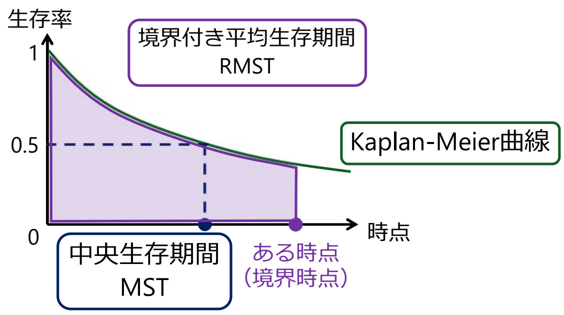
12

- 生存率 survival rate
  - その時点でイベント未発生者の割合（確率）
  - 多くはKaplan-Meier法を用いて推定
    - 生存率の時点に対するプロットをKaplan-Meier曲線
- 中央生存期間 Median Survival Time, MST
  - 生存率が50%となった時点
  - 半分の方がイベントを起こす時点
- 境界付き平均生存期間 Restricted Mean Survival Time
  - 境界時点までの生存期間の平均値

12

生存率、MST、RMST

13



13

Kaplan-Meier法①

14

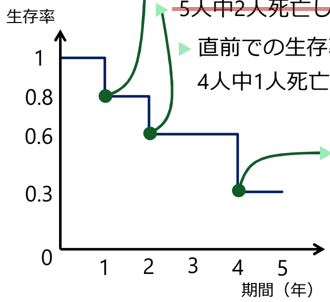
- 直前までat riskである人について、
  イベントを起こさなかった確率を乗じる
  - 生存例は、それまでの間、常に生存してきた
- 以下のデータセットを想定

イベント発生時点（年）	内容
1	死亡
2	死亡
3	脱落（打ち切り）
4	死亡
5	研究終了（打ち切り）

14

Kaplan-Meier法②

15



- 5人中1人死亡したため、生存率は$1-\frac{1}{5}=0.8$
- ~~5人中2人死亡したため、生存率は$1-\frac{2}{5}=0.6$~~
  - 直前での生存率は0.8であり、4人中1人死亡したため、$0.8\times\left(1-\frac{1}{4}\right)=0.6$
- 直前での生存率は0.6であり、2人中1人死亡したため、$0.6\times\left(1-\frac{1}{2}\right)=0.3$

15

打ち切り例の扱い

16

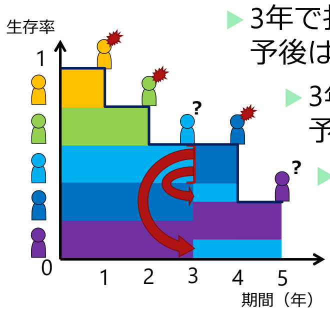
- 3年で打ち切りとなった対象者
  - 1年、2年での生存率を計算する際には、at riskであった人として解析に寄与
  - 4年、5年での生存率計算では分母に入らず
    - 生存率の計算自体には反映されている

16

追跡開始時は5人でスタート

17

- 1人あたり、20%の確率をもつ



- 3年で打ち切りとなった人の予後は分からない
  - 3年でat riskな人の予後で置き換えよう
    - 4年では、＋＝30%だけ生存率が低下

17

Kaplan-Meier推定量

18

$$\hat{S}(t)=\prod_{\{k\};t_k\le t}\left(1-\frac{d_k}{n_k}\right)$$

- d_k：時点t_kにおけるイベント数
  - 打ち切りのみ観察される場合は0
- n_k：時点t_kの直前におけるat risk数
- 縦軸にKaplan-Meier推定した生存率、横軸に追跡時間とした曲線をプロット

18

練習① Gehanの白血病データ

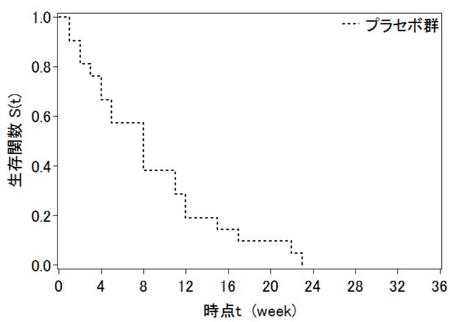
19

- 6-MP群のKaplan-Meier曲線を描こう

19

プラセボ群のみ

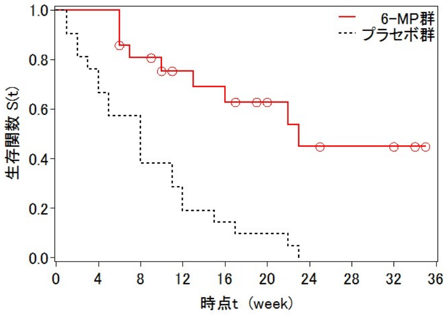
20



20

両群ともプロット

21



21

ログランク検定

22

- 2群の生存関数の間に有意差があるか
  - 帰無仮説は、「2群の生存関数が等しい」
- ログランク検定統計量

$$\frac{\sum_k (O_k - E_k)}{\sqrt{\sum_k V_k}}$$

  - O_k ：時点kでの観察イベント数
  - E_k ：時点kでの期待イベント数
  - V_k ：時点kでの分散

22

期待死亡数

23

- 時点kに対する分割表

	死亡	生存	計
6-MP群	d_{Ak}	$n_{Ak} - d_{Ak}$	n_{Ak}
プラセボ群	d_{Bk}	$n_{Bk} - d_{Bk}$	n_{Bk}
計	d_k	$n_k - d_k$	n_k

- 分割表の解析の考えから、
6-MP群の期待死亡数は

$$E_k = d_k \times n_{Ak}/n_k$$

23

時点kにおける分散

24

	死亡	生存	計
6-MP群	d_{Ak}	$n_{Ak} - d_{Ak}$	n_{Ak}
プラセボ群	d_{Bk}	$n_{Bk} - d_{Bk}$	n_{Bk}
計	d_k	$n_k - d_k$	n_k

$$V_k = \frac{d_k(n_k - d_k)n_{Ak}n_{Bk}}{n_k^2(n_k - 1)}$$

24

Gehanデータの例；時点1

25

	死亡	生存	計
6-MP群	0	21	21
プラセボ群	2	19	21
計	2	40	42

- 期待死亡数
  - $E_1 = 2 \times \frac{21}{42} = 1$
- 分散
  - $V_1 = \frac{2\times40\times21\times21}{42^2\times(42-1)}$

25

Gehanデータの例；時点2

26

	死亡	生存	計
6-MP群	0	21	21
プラセボ群	2	17	19
計	2	38	40

- 期待死亡数
  - $E_2 = 2 \times \frac{21}{40} = 1.05$
- 分散
  - $V_2 = \frac{2\times38\times21\times19}{40^2\times(40-1)}$

26

Gehanデータの例；全時点

27

- $\sum_k O_k = 9, \sum_k E_k = 19.249$
- $\sum_k V_k = 6.25696$

- 時点を層とした層別解析
  - Mantel-Haenszel検定統計量に同じ

27

重み付きログランク検定

28

$$\frac{\sum_k w_k(O_k - E_k)}{\sqrt{\sum_k w_k^2 V_k}}$$

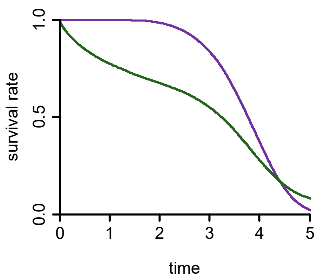
- 重みw_kを導入
  - $w_k = 1$なら通常のログランク検定
  - $w_k = n_k$（at risk数）なら
    Gehan-Breslow流の一般化Wilcoxon検定
  - $w_k = S(t_k)$ なら
    Peto-Prentice流の一般化Wilcoxon検定
  - $w_k = S(t_k)^\rho\{1 - S(t_k)\}^\gamma$ （ρ, γは予め設定）なら
    Fleming-Harringtonの$G(\rho, \gamma)$ class検定

28

一般化Wilcoxon検定

29

- 追跡早期に差が現れる場合に効率的
  - 感染症の治癒までの期間など

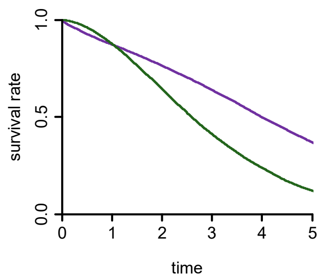


29

F-Hの$G(0,1)$class重み付き検定

30

- 追跡後期に差が現れる場合に効率的

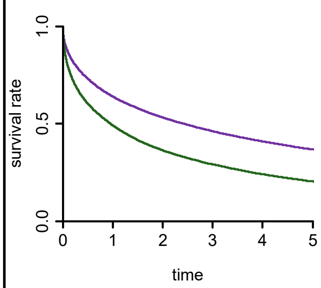


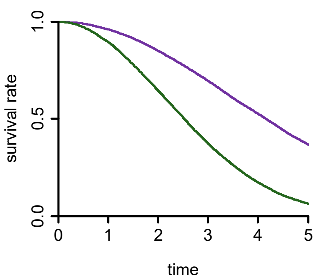
30

通常のログランク検定

31

- 差の現れ方が一定である場合に効率的





31

発生しやすさの表現

32

- 発生リスク、発生オッズ
  - ある期間における発生数を基に計算
  - 打ち切りがない場合に利用可能

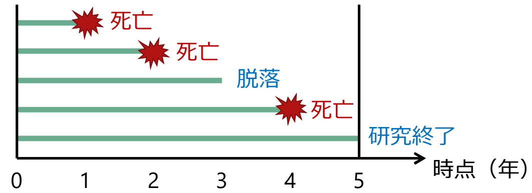
- 発生率
- ハザード
  - 似ているが、ちょっと違う
  - 率比とハザード比の解釈はほぼ同じ

32

発生率 incidence rate

33

- 単位時間あたりのイベント発生
  - 単位は1／(人)時間



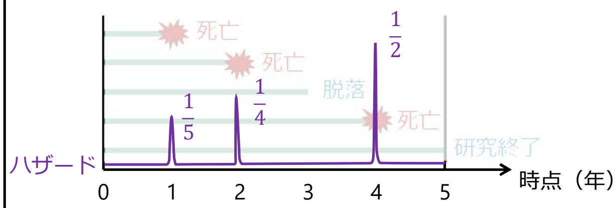
- $\frac{3}{1+2+3+4+5} = 0.2$（／人年）

33

ハザード hazard

34

- 直前まで生存している下で
  微小時間あたりのイベント発生
  - 単位は1／時間



34

時点とともに変化するハザード

35

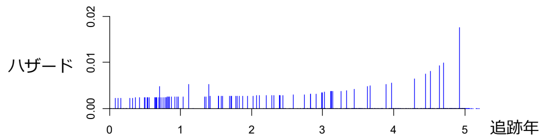
- 前スライドでは、
  1年時点で1/5、2年時点で1/4、
  4年時点で1/2、それ以外の時点では0
  - 要約指標には向かない
- 時が経つにつれ、発生しやすさが
  変化することを柔軟に捉えられる
  - 術後すぐは再発は少ないが、
    しばらくしてから再発が起こりうる
  - ある期間経過後は再発がまれ（治癒する）

35

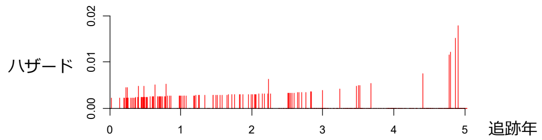
ハザードの比較

36

- 試験群のハザード



- 対照群のハザード

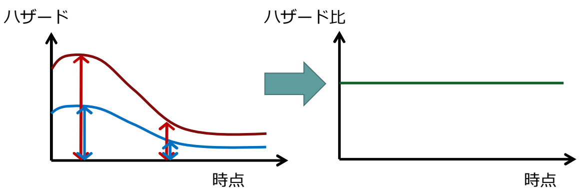


36

ハザード比

37

- ハザード自体はとびとびの値をとるので、
  期間全体を通して、何倍の違いであるか
  - 時点によらずハザード比は一定、という仮定
    （比例ハザード性）
  - ハザードは時間経過とともに変化してもよい



37

生存関数とハザード関数

38

- 生存期間を表す確率変数T

$$S(t) = \Pr(T \geq t)$$

- ハザード関数$\lambda(t)$

$$\lambda(t) = \lim_{\Delta t \to 0} \frac{\Pr(t \leq T < t + \Delta t | T \geq t)}{\Delta t}$$

  - tまではat riskであるものの（$T \geq t$）、
    その直後$t + \Delta t$までにイベント発生する確率
  - 解くと、

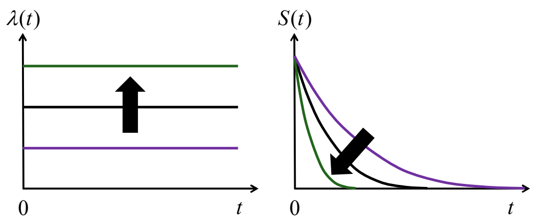
$$\lambda(t) = -\frac{d}{dt}\log_e S(t)$$

38

$S(t)$と$\lambda(t)$の対応関係

39

- $S(t) = \exp\left\{-\int_0^t \lambda(u)du\right\}$より、



39

ハザードが一定の場合

40

- 発生率とハザードは同等
  - Poisson回帰、指数回帰のいずれの方法とも同じ率比、ハザード比を推定可能
    - 生存関数に分布を仮定するので、パラメトリックなモデルとよぶ
    - ハザードが一定なので、当然ハザード比も一定

40

Cox比例ハザードモデル

41

- ハザード比一定を仮定した下で、次のような回帰モデルを考える

$$\lambda(t|\boldsymbol{X}) = \lambda_0(t)\exp(\boldsymbol{X\beta})$$

  - $\lambda_0(t)$ ：ベースラインハザード関数
    - ハザードは経時変化してよい
  - $\boldsymbol{X}$ ：説明変数（共変量）
    - 試験群なら$x = 1$、対照群なら$x = 0$とコード化
    - 試験群のハザード関数：$\lambda_0\exp(\beta)$
    - 対照群のハザード関数：λ_0
    - 対照群に対する試験群のハザード比：$\exp(\beta)$

41

Cox回帰の推定結果

42

- ハザード比を推定する回帰分析

最尤推定量の分析

パラメータ	自由度	パラメータ推定値	標準誤差	カイ 2 乗	Pr > ChiSq	ハザード比	95% ハザード比信頼限界	
g	1	-1.59787	0.42162	14.3630	0.0002	0.202	0.089	0.462

対数ハザード比

$e^{-1.598} \approx 0.202$

ログランク検定のp値と同じ
（設定等によってちょっと違うこともある）

42

練習②　率比を求める

43

- Gehanデータについて、率比を求めよう
  - 人時間法によって各群で発生率を計算
  - 群間で発生率の比をとる
- Cox回帰に基づき求められたハザード比と比べてみよう

43

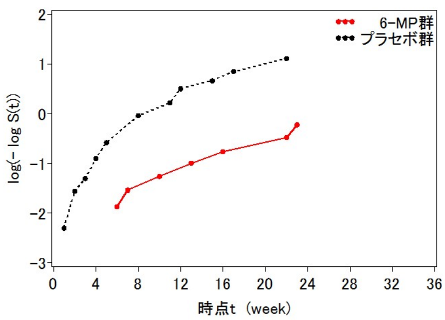
比例ハザード性のチェック

44

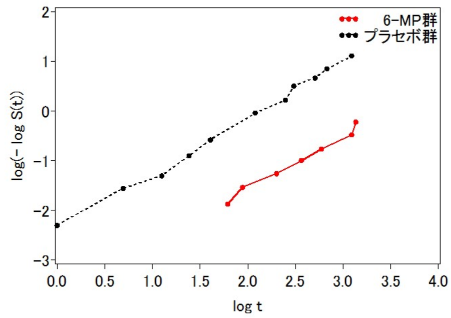
- 時点を含む共変量を追加
  - $\log t$や$\log(1+t)$
  - 回帰係数が0に近い、P値が大きいことを確認
- 二重対数プロット
  - 横軸に時点 or 時点の対数
  - 縦軸に生存関数の推定値の二重対数 $\log\{-\log\hat{S}(t)\}$
  - プロットが平行であるかを確認

44

二重対数プロット①



二重対数プロット②

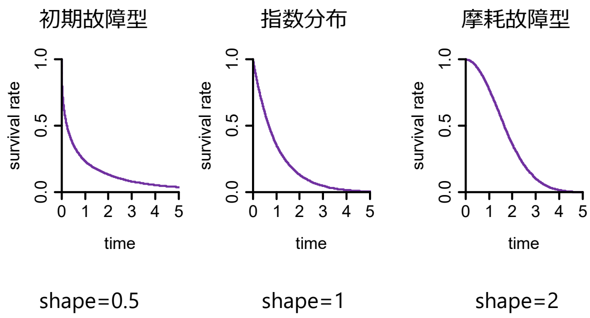


ハザード比の解釈

- ハザードも一定であれば、率比に同じ
- イベントが少なければ、リスク比に近い

- MST（中央生存期間）の比
  - 同じshapeパラメータをもつWeibull分布に従う場合（例えば、指数分布）

Weibull分布



Cox回帰におかれる仮定

- 比例ハザード性のほかにも・・・

$$\lambda(t|\boldsymbol{X}) = \lambda_0(t)\exp(\boldsymbol{X\beta})$$

- 全員共通するベースラインハザード$\lambda_0(t)$
  - 同じ群である対象者のイベント発生しやすさは同じという前提
- 効果はすべて$\exp(\boldsymbol{X\beta})$で説明
  - 治療効果は全員共通

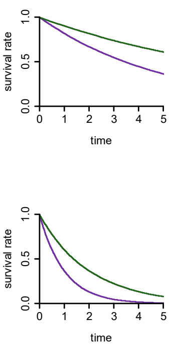
重症な人と軽症な人が混合

軽症
(100人)

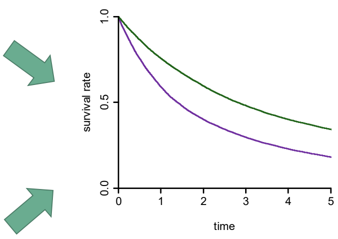
HR=0.5

重症
(100人)

HR=0.5



集団全体(200人)

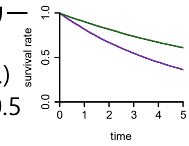


HR=0.6

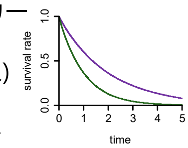
治療効果に異質性

51

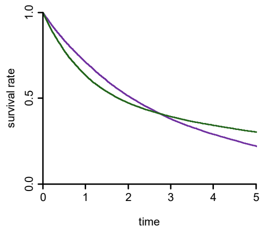
マーカー（＋）(100人) HR=0.5



マーカー（－）(100人) HR=2



集団全体(200人)



HR=0.79
そもそも・・？

51

無情報な打ち切り noninformative censoring

52

- Kaplan-Meier法、ログランク検定、Cox回帰で置かれる仮定
  - ランダムな打ち切り、とも
- 打ち切りとイベント発生が無関係
  - 研究終了時の生存
  - 偶然の事故による追跡不能
- 打ち切り例の予後を、at risk例で置き換えるため

52

情報のある打ち切り

53

- 例えば、
  - 重症患者の転院による追跡不能
  - 死因不明の死亡
- 対処法
  - イベントの定義を変更
    - 「がんによる死亡」から「あらゆる死亡」に
    - 競合リスクとみなす
  - 統計モデルによる対処
    - IPCW(inverse probability of censoring weighted)法
    - 打ち切り理由を丁寧に測定しておく必要

53

複合エンドポイントの例

54

- Composite endpoints
- 心血管イベント
  - 脳卒中、TIA
  - 急性心筋梗塞
  - 狭心症、心不全での入院
  - 解離性大動脈瘤での入院
  - 下肢動脈閉塞
  - 血栓症
  - 透析
  - Cr倍化

Sawada T, et al. Eur Heart J 2009: 30; 2461-9. Retracted.

54

競合リスク解析

55

- Competing risks
- イベントが複数あり、いずれか一つが発生すると、他のイベント発生を観察できなくなる

骨髄移植 → GvHD
骨髄移植 → 再発

糖尿病 → 原病悪化
糖尿病 → がん死

55

競合リスク版”3(4)種の神器”

56

- 累積発生割合 cumulative incidence function
  - Kaplan-Meier推定の代わり
- Gray の検定
  - ログランク検定の代わり
- 原因別ハザード回帰、Fine-Grayモデル
  - Cox比例ハザードモデルの代わり
  - 回帰係数の解釈に多くの議論

56

練習③　同じハザード比

57

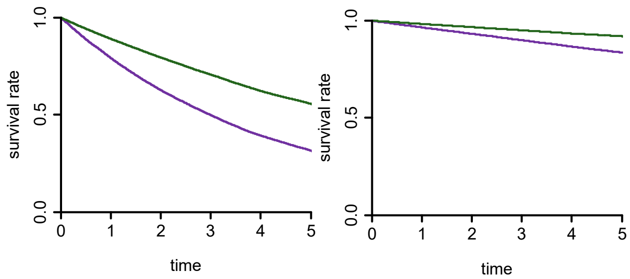
- ハザード比0.5倍もの治療効果！
  - 生存曲線はきれいな指数分布に従っている
- 以下の場合、
  最も近い試験治療群での3年生存率は？
  - 標準治療群の3年生存率：50%
    - 1) 25%　2) 50%　3) 60%　4) 70%　5) 75%　6) 80%
    - 試験治療群でのMSTは？
  - 標準治療群の3年生存率：90%
    - 1) 50%　2) 80%　3) 90%　4) 95%　5) 99%

57

ハザード比は相対指標

58

- 同じハザード比でも、印象は大きく異なる



58

絶対的な効果指標

59

- 中央生存期間
  - 中央生存期間まで達しなかったら・・？
  - 中央値を用いる恣意性
- 〇年生存率
  - 〇年を用いる恣意性
    - 複数時点の生存率をなるべく表示
- 発生率
  - 生存時間の分布に依存
    - 指数分布に従っていそうならいいかも

いくつか示しておくべき

59

時間依存性共変量 $\boldsymbol{X}(t)$

60

- time-dependent covariates
- 時間経過とともに値が変化しうる共変量
  - 大気汚染、気圧
  - 年齢
  - 血圧、血清脂質、炎症マーカー
  - 発作の回数

60

時間依存性共変量Coxモデル

61

$$\lambda(t) = \lambda_0(t)\exp\{\boldsymbol{X}(t)\boldsymbol{\beta}\}$$

- 外生変数 external variable である必要
  - $\boldsymbol{X}(t)$が観察されるかとイベント発生は無関係
  - 外生変数でないものを内生 internal 変数

61

練習④　内生変数はどれ？

62

- 大気汚染
- 気圧
- 年齢
- 血圧
- 血清脂質
- 炎症マーカー
- 発作の回数

62

内生変数の場合

63

- $\boldsymbol{X}(t)$が観察されたということは、対象者がイベント未発生だと確定
  - ハザードλとの対応が不成立

$$\Pr\{T > t \mid \boldsymbol{X}(t)\} \neq \exp\left[-\int_0^t \lambda\{u \mid \boldsymbol{X}(u)\}du\right]$$

- 因果関係を調べるなら慎重に
  - 結果で他の結果を単に条件付けてはならない
  - 予後のよい人を選択することのバイアス

63

関連や予測なら・・

64

- いくつかの対処法
  - ランドマークモデル
  - 同時 joint モデル
  - 多状態 multi-state モデル
  - ・・・

64

ランドマークモデル

65

- ランドマーク時点 s における at risk例で条件付け
- 時点 s までに得られた情報を利用
- 単一の生存時間である場合

(vanHouwelingen 2007, *Scand Stat Theory Appl.*)

$$\boxed{\lambda(t|s, \boldsymbol{X}(t)) = \lambda_0(t|s)\exp\{\boldsymbol{X}(s)\boldsymbol{\beta}(s)\}}$$

  - $\boldsymbol{X}(s)$はこのモデルにおいて外生変数

65

ランドマークモデルの解釈

66

- $t = 0$でなく$t = s$を研究開始時点とみなす
- 標的集団は「時点sにて生存しているような集団」
  - 研究開始時に定義できる集団ではない

66

例：脂質低下と心血管イベント

67

- スタチンによる脂質低下は心血管イベントと因果関係があるか？
  - 脂質低下したからイベントが起こりにくい？
  - イベント発生しにくい人は脂質低下しやすい？
    - 逆因果が生じうる
- 治療開始から半年間で脂質低下した人はその先の予後がよい
  - 関連はいえる
  - 半年後の脂質低下で予後を予測できそう
  - 半年以内にイベントを起こすような人は？
    - 議論を諦める（万能ではない）

67

ランドマークモデル適用方法

68

- ランドマーク時点 s 以前のイベント・打ち切り例は左側切断(truncation)
  - 除外するだけ
- 以上の加工を行ったデータセットに通常のCox回帰を適用するだけ

68

まとめ

69

- 生存時間の3種の神器
  - Kaplan-Meier法、ログランク検定、Cox回帰
- ハザード比の意味
- 無情報打ち切り
  - 競合リスク解析
- 時間依存性共変量の扱い