2019/12/2　北大・医理工統計学⑥

分割表の解析と
ロジスティック回帰

北海道大学　医学統計学
横田　勲

今回の内容

2

- 分割表と効果の指標
- ロジスティック回帰によるオッズ比推定

到達目標

- 分類データに関する推定・検定を知る
- JMPによって一連の解析を実行できる

カテゴリカル(二値,多値)データ

3

- 治療(曝露)の有無、
進行度ステージ(I, II, III, IV)、疾患の有無
- 分割表による要約
  - 人数と曝露群別に求めた割合を表記

治療	進行度ステージ				合計
	I	II	III	IV	
新治療	2 [4%]	5 [10%]	23 [46%]	20 [40%]	50
標準治療	4 [8%]	4 [8%]	24 [48%]	18 [36%]	50

曝露	疾病発生		合計
	あり	なし	
あり	12 [20%]	48 [80%]	60
なし	16 [10%]	144 [90%]	160

割合、率、比

4

- 割合 proportion
  - 全体に占める程度
  - 0から1をとる指標
- 率 rate
  - 単位時間あたりの発生数
  - 0から∞(無限大)をとり、1/単位時間 が単位
- 比 ratio
  - 同じ単位をもつ2指標の相対的な大きさ
  - 0から∞(無限大)をとり、単位はなし

効果の指標

5

- 治療(曝露)効果の方向や大きさの表現

指標	差の指標	比の指標
リスク、割合	リスク差	リスク比
オッズ		オッズ比
率	率差	率比
ハザード		ハザード比

割合に関する効果の指標①

6

- リスク差 risk difference, リスク比 risk ratio
  - 疾病発生割合(リスク)の群間比較
  - 直感的な解釈
  - NNT (Number Needed to Treat): -1/(リスク差)
    - 何名治療すれば、1人の疾病発生を抑えられるか
- 発生オッズ比 incidence odds ratio
  - 疾病発生オッズ(発生数/非発生数)の群間比較
  - 数学的によい性質

2つの治療法を比較する 1800名の観察研究

7

治療法	イベントあり	イベントなし	合計
試験治療	22 (2.2%)	978	1000
標準治療	20 (2.5%)	780	800
合計	42	1758	1800

- リスク差：$\frac{22}{1000} - \frac{20}{800} = -0.003$
- リスク比：$\frac{22}{1000} / \frac{20}{800} = 0.88$
- オッズ比：$\frac{22}{978} / \frac{20}{780} = 0.877.. \approx 0.88$

リスク差に注目

8

治療法	イベントあり	イベントなし	合計
試験治療	22 (2.2%)	978	1000
標準治療	20 (2.5%)	780	800
合計	42	1758	1800

- 試験治療はイベント発生リスクを下げるか？
  - 試験治療での脳卒中リスクは2.2%
  - 標準治療での脳卒中リスクは2.5%
    - 偶然の違いか否か

検定

9

- 観察された差が、偶然によるものか、系統的な違いがあるかを判断
- 背理法の論理
  - 「群間に差がない」と仮定（帰無仮説）
  - 観察値はどの程度まれな現象であるか、を確率で表現
    - p値：観察値とそれよりまれな現象が起こる確率
  - p値が小さければ(例えば5%)、帰無仮説が誤っていたと判断し、有意差あり

群間差が従う正規分布

10

- 帰無仮説が正しい下で、群間差について
  - 平均は0
  - 分散は$\frac{t}{N} \cdot \frac{N-t}{N} \cdot \left(\frac{1}{n} + \frac{1}{m}\right)$

比較群	疾病発生 あり	疾病発生 なし	合計
試験群	a	b	n
対照群	c	d	m
合計	t	$N-t$	N

割合の差の検定

11

- 帰無仮説の下で、以下の検定統計量Zが標準正規分布に従う

$$Z = \frac{(\widehat{p_1} - \widehat{p_2}) - 0}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n} + \frac{1}{m}\right)}} \sim N(0, 1^2)$$

ただし、$\widehat{p_1} = \frac{a}{n}, \widehat{p_2} = \frac{c}{m}, \hat{p} = \frac{t}{N}$

カイ二乗検定

12

- 検定統計量Zを2乗して整理すると

$$Z^2 = \chi^2 = \frac{N(ad-bc)^2}{nmt(N-t)} \sim \chi_1^2$$

  - 自由度1のカイ二乗分布
  - 割合の差の検定と全く同じ
- 2×2分割表以外の分割表にも拡張可能

カイ二乗検定

13

- 治療間で等しい確率で脳卒中が発生する
  - 試験治療を1000人、標準治療を800人
    サンプリングしたら、42人が脳卒中を起こしていた

	イベントあり	イベントなし	合計
試験治療	23.33	976.67	1000
標準治療	18.67	781.33	800
合計	42 (2.33%)	1758	1800

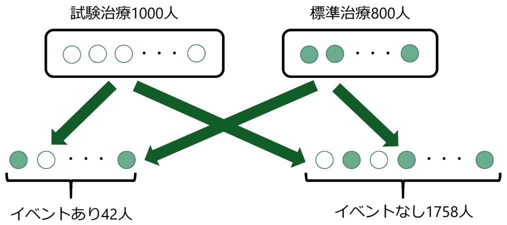
22人 と 20人

  - 観察値とのズレを統計モデルを基に、確率で表現

Fisherの直接確率検定①

14

- 試験治療1000人、標準治療800人と割付けたら
  イベントありが42人含まれていた
  - そのうち試験治療が22人である確率は？



Fisherの直接確率検定②

15

- 試験治療1000人、標準治療800人から
  イベントあり42人が選ばれた
  - そのうち試験治療が0人である確率は？
  - そのうち試験治療が1人である確率は？

    ⋮
  - そのうち試験治療が22人である確率は？
  - そのうち試験治療が23人である確率は？

    ⋮
  - そのうち試験治療が42人である確率は？

この和の2倍が(両側)p値

Fisherの直接確率検定③

16

- $\dfrac{\binom{42}{22}\binom{1758}{978}}{\binom{1800}{1000}}$
  - 1800人を1000人に治療群として割付た場合
    イベント発生の42人中22人が治療群で
    イベントなしの1758人中978人が治療群
    である確率
- $\left\{\dfrac{\binom{42}{22}\binom{1758}{978}}{\binom{1800}{1000}}+\dfrac{\binom{42}{21}\binom{1758}{979}}{\binom{1800}{1000}}+\cdots+\dfrac{\binom{42}{0}\binom{1758}{1000}}{\binom{1800}{1000}}\right\}\times 2$

Fisher検定とカイ二乗検定

17

- 幸い、どちらも同じようなp値
  - Fisher検定 p値=0.79>0.05
  - カイ二乗検定 p値=0.68>0.05
- データの得られ方を考えて
  ふさわしい方を採用すべき
  - 目の前の患者1800人をランダムに割付け
    - Fisherの直接確率検定
  - 大勢の患者から1800人ランダムサンプリング
    - カイ二乗検定

練習①　分割表の検定

18

- 以下の分割表について以下の検定を実行
  - 割合の差の検定（カイ二乗検定）
    - カイ二乗検定統計量が3.84を超えたら
      両側5%水準で有意差あり
  - Fisherの直接確率検定

	イベントあり	イベントなし
新治療	2	8
標準治療	7	3

P値って何だ？

19

以下のうち、正しい説明はどれでしょう？

1. P値は帰無仮説が正しい確率である
2. P=0.05の場合、
   対立仮説が正しい確率が95%である
3. P値は帰無仮説が正しい場合に
   観察された結果が得られる順位である
4. P値が小さいほど、より関連が強い
   （よりリスク因子の影響が強い）
5. P値が1に近ければ、差はない

P値が有意水準を下回った場合

20

- 帰無仮説が誤っていたことは言える
- 「治療によりアウトカムに差があった」と言っていい？
  - 比較可能性があるならば、言ってもよい
    - 治療以外のリスク因子の分布が群間で同じ
      - 年齢、重症度、喫煙歴、、、
      - 測定していない、測定できない因子もすべて
    - 群間差は治療だけから生じる内容、とお膳立て

ランダム化

21

- 様々なリスク因子の分布が
  完全に同一になることはないが、
  グループの人数が増えるほど
  平均的に同一に近づくはず
- 「測定可能な特徴」にもとづいていない
  - 「測定できない特徴」だって
    平均的に同一になっているに違いない！
  - 比較可能性を担保できる最強の方法

観察研究データであれば

22

- 比較可能性を担保することはできない
- “比較可能”な状態に少しでも近づける
  - 交絡バイアスをデザインや解析で制御
    - 治療にもアウトカムにも影響を与える交絡因子によって引き起こされることがある
    - 回帰分析による制御
- とはいえ、比較可能でない以上、
  検定を行う意味がない

症例数設計

23

- ランダム化研究では、
  期待通りの治療効果がある場合、
  80%や90%の確率で有意差がつくよう設計
- 観察研究（特にレトロ研究）では
  症例数設計はされないことがほとんど
  - 症例数をたくさん集めてしまえば、
    わずかな治療効果でも有意差がつく

どっちの分割表がP値は小さい？

24

治療法	治癒	不変	合計
A	12 (60%)	8	20
B	6 (30%)	14	20

治療法	治癒	不変	合計
A	6200 (31%)	13800	20000
B	6000 (30%)	14000	20000

P値のつかいどころ

25

- どちらの治療法（薬）を採用するかの意思決定
  - 比較可能になるようランダム化を行い、適切な症例数を集めるというお膳立てをして、有意差ありとなったら、新治療を採用するということを事前に定めておく場面
- 観察研究においてはP値の計算は無意味
  - Epidemiology（雑誌）では、P値の使用禁止
    - Statistical testingに集められている論文参照
      https://journals.lww.com/epidem/pages/collectiondetails.aspx?TopicalCollectionId=4

どの程度の効果か

26

治療法	イベントあり	イベントなし	合計
試験治療	22 (2.2%)	978	1000
標準治療	20 (2.5%)	780	800
合計	42	1758	1800

- 試験治療はイベント発生リスクをどの程度下げるか？
  - 試験治療での脳卒中リスクは2.2%
  - 標準治療での脳卒中リスクは2.5%
    - 適切な指標で比べてみよう

効果をはかる指標

27

- リスク差：$\frac{A}{A+B}-\frac{C}{C+D}$
- リスク比：$\frac{A}{A+B}\Big/\frac{C}{C+D}$
  - 脳卒中があった割合の群間差、群間比
- オッズ比：$\frac{A}{B}\Big/\frac{C}{D}$

治療法	イベントあり	イベントなし	合計
試験治療	A	B	A+B
標準治療	C	D	C+D
合計			N

効果をはかる指標

28

- リスク差：$\frac{22}{1000}-\frac{20}{800}=-0.003$
- リスク比：$\frac{22}{1000}\Big/\frac{20}{800}=0.88$
  - 脳卒中があった割合の群間差、群間比
- オッズ比：$\frac{22}{978}\Big/\frac{20}{780}=0.877\approx 0.88$

治療法	イベントあり	イベントなし	合計
試験治療	22 (2.2%)	978	1000
標準治療	20 (2.5%)	780	800
合計	42	1758	1800

リスク差/比とオッズ比の違い

29

- リスク差、リスク比
  - 疾病発生割合を比べたものとして解釈
- オッズ比
  - 直接的な解釈は難しい
  - リスク比を近似したものといえる場合がある
  - もともとは「ケースコントロール研究」にて開発された

全員追跡してもオッズ比？

30

- コホート研究ではオッズ比をあえて出す意味はあまりない
  - ロジスティック回帰で推定される
    - 数学的に近似がうまくいきやすい
- 解析ソフトが出してくれる・・
- 疾患分野ごとの慣習・・・・

効果指標の確からしさ

31

- 同じリスク比、オッズ比
  - リスク比：0.88

	イベントあり	イベントなし	合計
試験治療	2200	97800	100000
標準治療	2000	78000	80000
合計	4200	175800	180000

  - このデータから推定したリスク比のほうが確からしそう
    - それらしいリスク比の範囲を示す

1800人集めてリスク比は0.88

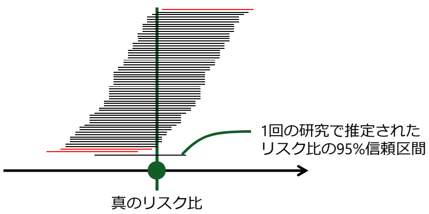
32

- 同じ研究をもう一度行ったら・・？
  - リスク比は0.80かもしれない
  - リスク比は1.00かもしれない
- 研究ごとに推定されるリスク比に、適当な幅をつけてみよう
  - リスク比の95%信頼区間：0.48-1.60

95%信頼区間(Confidence Interval)

33

- 研究ごとに95%信頼区間を計算すれば、100回繰り返したうち、95回は“真の”リスク比を含む



リスク比：0.88 (95%CI 0.48-1.60)

34

- 真のリスク比は0.88だ！
  - そうかもしれない
- 真のリスク比は0.50だ！
  - そうかもしれない
- 真のリスク比は2.00だ！
  - それは違うのでは

信頼区間と検定

35

- 検定で有意とならない値の範囲が信頼区間
- 「P<0.05である」
  ⇔「リスク比の95%CIが1をまたがない」

	あり	なし
試験治療	22	978
標準治療	20	780

95%CI ：0.48-1.60
p=0.68

	あり	なし
試験治療	2200	97800
標準治療	2000	78000

95%CI ：0.83-0.93
p<0.01

練習②　効果の指標

36

- JMPを用いて、以下の分割表についてリスク差、リスク比、オッズ比を求めよ

治療法	イベントあり	イベントなし	合計
試験治療	22 (2.2%)	978	1000
標準治療	20 (2.5%)	780	800
合計	42	1758	1800

2つの治療法を比較する 1800名の観察研究

37

治療法	イベントあり	イベントなし	合計
試験治療	22 (2.2%)	978	1000
標準治療	20 (2.5%)	780	800
合計	42	1758	1800

- 試験治療により、イベントはやや減った
  - リスク比：0.88 (95%CI: 0.48-1.60)
  - オッズ比：0.88 (95%CI: 0.48-1.62)
- 治療効果はもう少し大きいはず・・・？

年齢でサブグループ化

38

- 75歳以上

治療法	イベントあり	イベントなし	合計
試験治療	18 (3.0%)	582	600
標準治療	6 (6.0%)	94	100

- 75歳未満

治療法	イベントあり	イベントなし	合計
試験治療	4 (1.0%)	396	400
標準治療	14 (2.0%)	686	700

- どちらのサブグループでもリスク比0.5

交絡 (confounding)

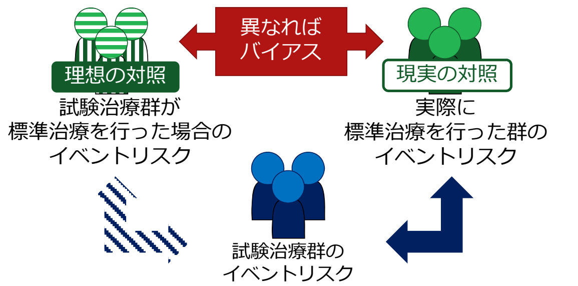
39

- 全体での単純な解析結果と サブグループ解析結果が異なる現象
  - 全体でのリスク比：0.88
  - サブグループでのリスク比
    - 75歳以上：0.50
    - 75歳未満：0.50
  - 全体でのリスク比は、 サブグループ解析の間に入るはずなのに...

治療効果の適切な比較

40

- 試験治療の治療効果を知りたい
  - 比較対照は標準治療



原因を考えてみる

41

- 標準治療(対照)でのイベントリスクが 高齢なほど高い
  - 75歳以上： 6 / 100 = 6.0%
  - 75歳未満： 14 / 700 = 2.0%
- 高齢なほど試験治療を受けがち
  - 試験治療での75歳以上：600/1000=60.0%
  - 標準治療での75歳以上：100/800=12.5%

オッズ比の計算

42

治療法	イベントあり	イベントなし	合計
試験治療	22 (2.2%)	978	1000
標準治療	20 (2.5%)	780	800
合計	42	1758	1800

$$\frac{\text{イベントあり}_{in}\text{試験治療} / \text{イベントなし}_{in}\text{試験治療}}{\text{イベントあり}_{in}\text{標準治療} / \text{イベントなし}_{in}\text{標準治療}}$$

$$= \frac{22/978}{20/780} \approx 0.88$$

ロジスティック回帰モデル

43

$$\log\frac{(\text{イベントあり})}{(\text{イベントなし})} = \beta_0 + \beta_1 \times (\text{試験治療})$$

オッズ

試験治療であれば1,
標準治療なら0

- 試験治療での対数オッズ：$\beta_0 + \beta_1$
- 標準治療での対数オッズ：β_0

$$\text{オッズ比}：\frac{\exp(\beta_0+\beta_1)}{\exp(\beta_0)} = \exp(\beta_1)$$

年齢調整ロジスティック回帰

44

オッズ

- $\mathrm{logit}(p_i) = \log\left(\frac{p_i}{1-p_i}\right) = \beta_0 + x_{1i}\beta_1 + x_{2i}\beta_2$
  - x_{1i}：試験治療なら1、標準治療なら0
  - x_{2i}：75歳以上なら1、75歳未満なら0
- 交互作用項は含めていないので、
年齢によらず治療効果は同じという仮定

年齢を調整した治療のオッズ比

45

$\log(\text{オッズ}) = \beta_0 + \beta_1 \times (\text{試験治療}) + \beta_2 \times (75\text{歳以上})$

オッズ	75歳未満	75歳以上
試験治療	$\exp(\beta_0+\beta_1)$	$\exp(\beta_0+\beta_1+\beta_2)$
標準治療	$\exp(\beta_0)$	$\exp(\beta_0+\beta_2)$

75歳未満でのオッズ比

$$\frac{\exp(\beta_0+\beta_1)}{\exp(\beta_0)} = \exp(\boldsymbol{\beta_1})$$

75歳以上でのオッズ比

$$\frac{\exp(\beta_0+\beta_1+\beta_2)}{\exp(\beta_0+\beta_2)} = \exp(\boldsymbol{\beta_1})$$

ロジスティック回帰分析結果

46

	β	(95%CI)	オッズ比	(95%CI)
切片	-3.89	(-4.38, -3.40)		
x_1；治療	-0.72	(-1.44, 0.01)	0.49	(0.28-1.01)
x_2；年齢	1.13	(0.40, 1.86)	3.10	(1.50-6.49)

- 治療のオッズ比とその信頼区間

$$e^{-0.72} = \exp(-0.72) = 0.49$$

練習③

47

- JMPを用いてロジスティック回帰を行い、1枚前のスライドのような結果となることを確認せよ

指数関数・対数関数

48

- $\log(A \times B) = \log A + \log B$
  - 対数の底はeであり、通常省略
  - 比のモデルは掛け算モデル
対数変換すれば足し算モデルになる
    - 線形モデルで表現できる
- $e^{\log A} = A$
- $e^{a+b} = e^a \times e^b$
  - 以降、e^aのことを$\exp a$と表記する
    - $\exp(a+b) = \exp a \times \exp b$

練習④

49

- 75歳未満で試験治療を受けた人に対する75歳以上で標準治療を受けた人のイベント発生オッズ比は？

一般化線形モデル generalized linear model

50

- アウトカムが指数型分布族に従う
  - 正規分布、二項分布、Poisson分布など
- 分布を表現する正準パラメータθについて
$$g(\theta) = \boldsymbol{X\beta}$$
- $g(\cdot)$　リンク関数
  - リスク差モデルではそのまま$g(\theta) = \theta$

分布とリンク関数の組合せ

51

分布	リンク関数	効果の指標	別名
二項	恒等	リスク差	
二項	対数	リスク比	
二項 / 多項	ロジット	オッズ比	ロジスティック回帰
Poisson	恒等	率差	
Poisson	対数	率比	Poisson回帰
指数	対数	ハザード比	指数回帰
二項	プロビット		プロビット回帰

- 医学研究ではCox回帰を用いたハザード比推定がほとんど
- イベント割合が極端に低い場合に向かない、効果指標がないといった理由でプロビット回帰はあまり使われない

リスク差の回帰モデル①

52

- 一般線形モデルで行ったら？
  - t検定のように回帰係数を平均値の差となるようモデルを作れたし・・
- アウトカムは連続量ではなく、イベントあり／なしの二値
  - アウトカムは二項分布に従うと考えたほうが自然

リスク差の回帰モデル②

53

- $y_i \sim Bin(n_i, p_i)$
  - アウトカムは二項分布に従う
  - 今は$n_i = 1$
  - イベント発生確率がp_i
  - $E(y_i) = n_i p_i$
    - 今はy_iの期待値がイベント発生確率p_i
- $p_i = \beta_0 + x_i \beta_1$
  - x_i：試験治療なら1、標準治療なら0

リスク差の回帰モデル③

54

パラメータ	推定値	(95%信頼区間)
β_0	0.025	(0.014, 0.036)
β_1	-0.003	(-0.017, 0.011)

- 試験治療により、イベント発生が-0.3%(95%CI: -1.7%, 1.1%)だけ減る
  - NNTは$\frac{1}{|-0.003|} \approx 333$
- 割合の差の検定に基づく信頼区間と同じ

練習⑤

55

- JMPを用いて一般化線形モデルをあてはめ、1枚前のスライドのような結果となることを確認せよ
  - リスク比回帰も同様にやってみよう

以降、付録

56

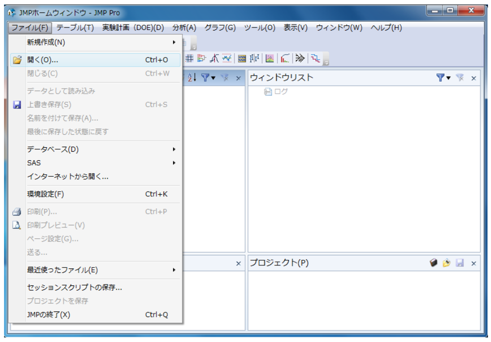
- 分割表の解析
- ロジスティック回帰の実行
- 一般化線形モデルの実行

57

分割表の解析

CSVファイルを読み込む

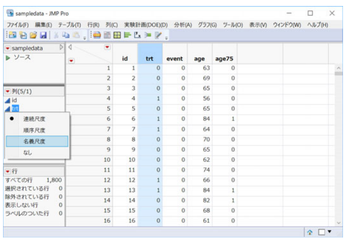
58



正しい尺度に変更

59

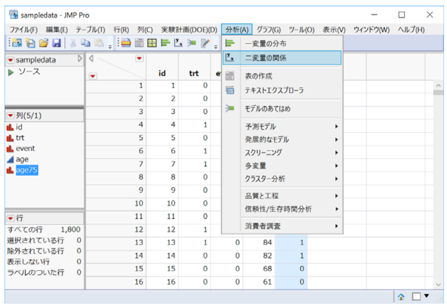
- 連続尺度：年齢、BMI、血圧
- 順序尺度：改善/不変/悪化のような順序性あり
- 名義尺度：あり/なし、術式A/B/C



分割表をかく①

60

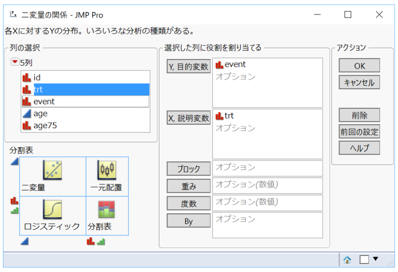
- 分析 ＞ 二変量の関係 を選択



分割表をかく②

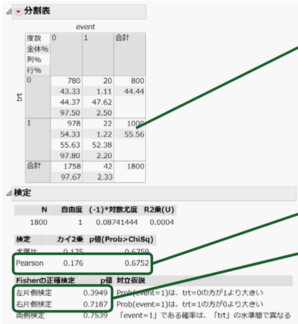
61

- 結果(stroke)を「Y, 目的変数」に
  治療(offp)を「X, 説明変数」に



分割表と検定結果

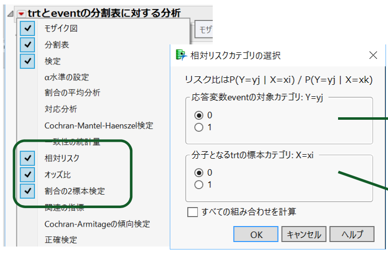
62



- 度数
  列での%
  行での%
  全体での%
  が表示
- カイ二乗p値
- Fisher検定p値
  - 左 or 右片側検定の小さいほうを2倍

リスク差、リスク比、オッズ比

63



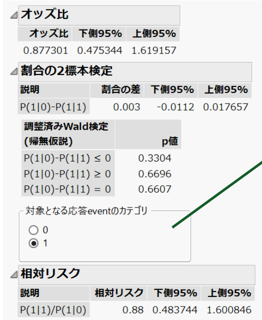
「イベントあり」の入力内容

分子におく治療群

- 相対リスク：リスク比
- 割合の2標本検定：リスク差

リスク差、リスク比、オッズ比

64



- オッズ比
- リスク差
  - 「イベントあり」の入力内容を選ぶ
- リスク比

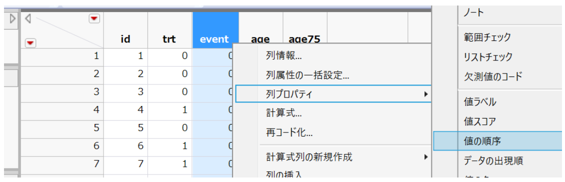
65

ロジスティック回帰の実行

データの準備

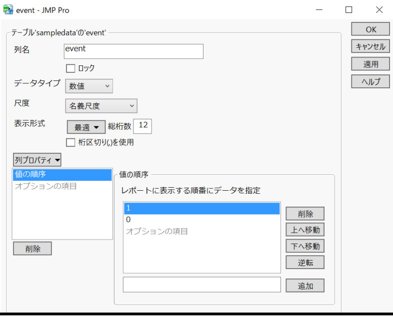
66

- Stroke を右クリック
  - 列プロパティ ＞ 値の順序 をクリック



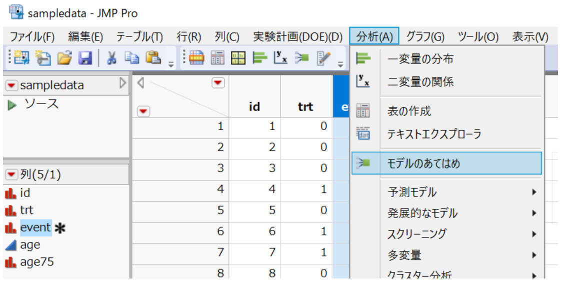
イベントあり を一番上にする
67

- 値の順序 にて 上へ移動 をクリック



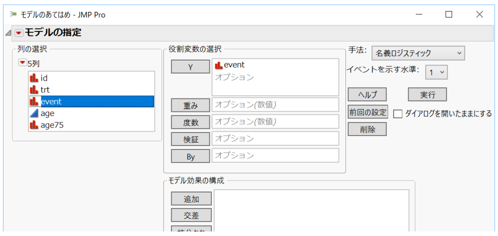
ロジスティック回帰の実行
68

- 分析 ＞ モデルのあてはめ　を選択



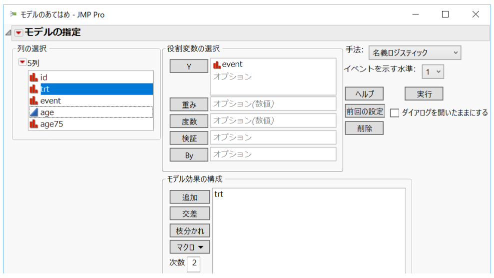
アウトカム (stroke) の選択
69

- 列の選択 にある(アウトカム)をクリック
- 役割変数の選択 にある Y をクリック
  - 手法が 名義ロジスティック になったことを確認



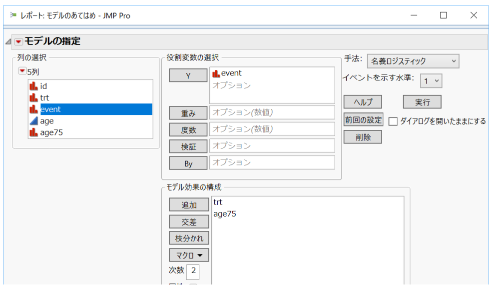
オッズ比を調べたい因子を投入
70

- 列の選択 にある (治療法) を選択
- モデル効果の構成 にある 追加をクリック



調整因子(交絡因子)を投入
71

- 前スライドと同じ
  - 興味のある因子も調整因子もモデル上は同じ扱い



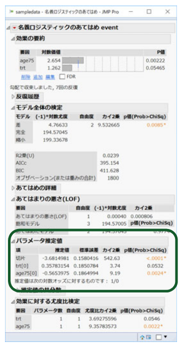
解析結果
72

- パラメータ推定値 に表示
  - ？？？？？

パラメータ推定値

項	推定値	標準誤差	カイ2乗	p値(Prob>ChiSq)
切片	-3.6814981	0.1580416	542.63	<.0001*
trt[0]	0.35783154	0.1850784	3.74	0.0532
age75[0]	-0.5653975	0.1864994	9.19	0.0024*

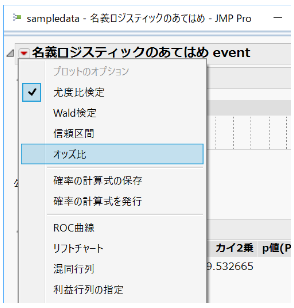
推定値は次の対数オッズに対するものです： 1/0



オッズ比を表示する

73

- 名義ロジスティックのあてはめ 脇にある▼をクリック
- オッズ比をクリック



オッズ比 0.49 (95%CI: 0.24-1.01)

74

- 治療群のどちらが分子・分母かに注意

オッズ比

event: 1対0のオッズ比に対して

trtのオッズ比 → 治療群のオッズ比

水準1	/水準2	オッズ比	p値(Prob>ChiSq)	下側95%	上側95%
1	0	0.4888678	0.0532	0.2366538	1.0098793
0	1	2.0455426	0.0532	0.9902173	4.2255819

age75のオッズ比 → 年齢のオッズ比

水準1	/水準2	オッズ比	p値(Prob>ChiSq)	下側95%	上側95%
1	0	3.0981183	0.0024*	1.4914228	6.4356914
0	1	0.3227766	0.0024*	0.1553835	0.6705007

次の信頼限界にはWald近似が使われています: trt age75
オッズ比の検定と信頼区間は、Wald法に基づいて計算されています。

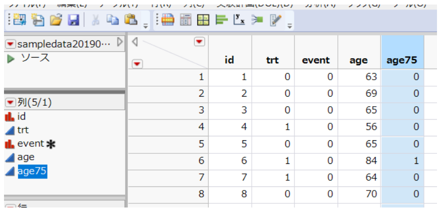
75

一般化線形モデルのあてはめ

尺度を変える

76

- 説明変数に加える変数について、カテゴリカルなものは0/1で入力して連続尺度とする
  - "0"に対する"1"の場合の効果指標が計算

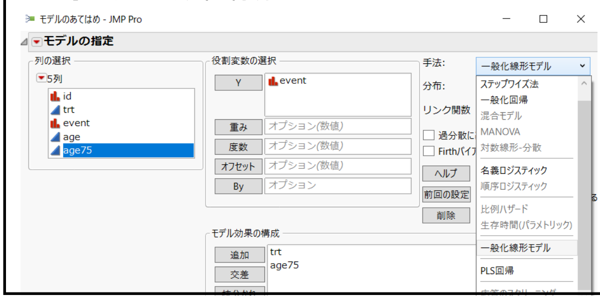


	id	trt	event	age	age75
1	1	0	0	63	0
2	2	0	0	69	0
3	3	0	0	65	0
4	4	1	0	56	0
5	5	0	0	65	0
6	6	1	0	84	1
7	7	1	0	64	0
8	8	0	0	70	0

モデルダイアログ

77

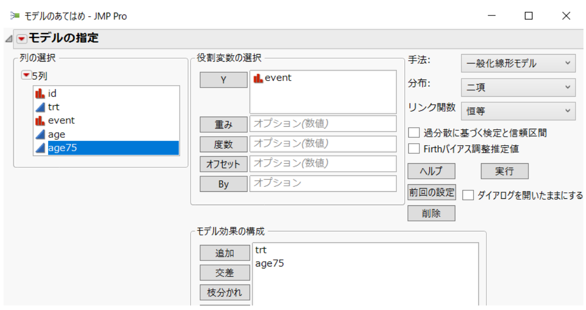
- ロジスティック回帰と同様
- 手法を一般化線形モデルに



分布とリンク関数を指定

78

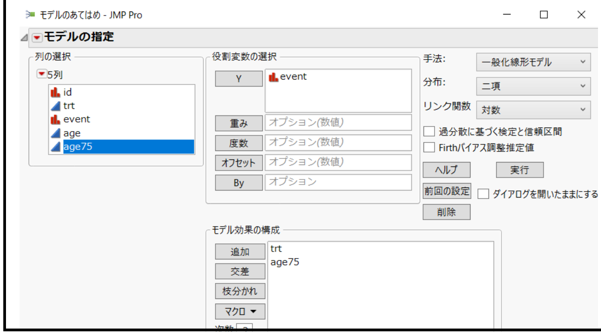
- 例えばリスク差回帰



分布とリンク関数を指定

79

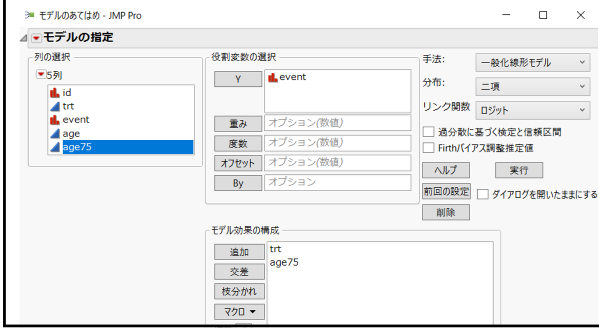
- 例えばリスク比回帰



分布とリンク関数を指定

80

- 例えばロジスティック回帰



パラメータ推定値

81

- ロジスティック回帰の例
- 対数オッズ比が推定値として表示

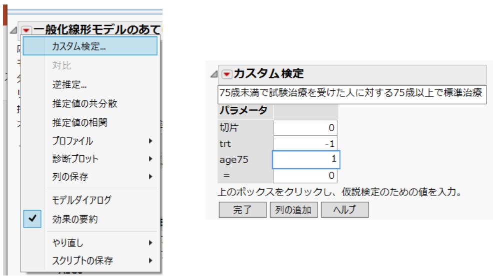
パラメータ推定値

項	推定値	標準誤差	尤度比カイ2乗	p値(Prob>ChiSq)	下側信頼限界	上側信頼限界
切片	-3.889064	0.2516141	238.90216	<.0001*	-4.382551	-3.395577
trt	-0.715663	0.3701567	3.7380566	0.0532	-1.441646	0.0103195
age75	1.1307949	0.3729989	9.1907868	0.0024*	0.3992381	1.8623517

カスタム検定①

82

- 好きな組み合わせを検定



カスタム検定②

83

- 推定値と標準誤差、P値が計算
  - 信頼区間は表示されない
  - 「推定値±1.96×標準誤差」を自分で計算

カスタム検定

75歳未満で試験治療を受けた人に対する75歳以上で標準治療

パラメータ	
切片	0
trt	-1
age75	1
=	0
値	1.8464580095
標準誤差	0.6502815022
カイ2乗	8.1754804831
p値(Prob>ChiSq)	0.004246042
(-1)*対数尤度	198.6581894

$$\exp(1.84) \approx 6.3$$

$$\exp(1.84 - 1.96 \times 0.65) \approx 1.8$$

$$\exp(1.84 + 1.96 \times 0.65) \approx 22.5$$