統計学(基礎)

第5回
対応のあるデータとマクネマーの検定
(対応のあるχ^2検定)

川崎市立看護大学大学院 看護学研究科 博士前期課程

ごめんなさい、積み残しです

2×2より大きいχ^2乗検定

Health Habits（健康習慣）

- データライブラリ 5. Frequencies　Health Hbits
- 1,184人の学生における身体活動量と果物の摂取量
- 変数：
  - Physical Activity
    参加者の身体活動量（Low＝低い、Moderate＝中程度、Vigorous＝高い）
  - Fruit Consumption
    参加者の果物摂取量（Low＝少ない、Medium＝中程度、High＝多い）

3×3のクロス集計(順序) JASP

結果

分割表

分割表

		Fruit Consumption			
Physical Activity		Low	Medium	High	合計
Low	Count	69.00	25.00	14.00	108.0
	Standardized residuals	3.45	−0.97	−2.99	
Moderate	Count	206.00	126.00	111.00	443.0
	Standardized residuals	−0.83	0.80	0.14	
Vigorous	Count	294.00	170.00	169.00	633.0
	Standardized residuals	−1.19	−0.21	1.59	
合計	Count	569.00	321.00	294.00	1,184.0

カイ二乗検定

	値	df	p
Χ²	14.15	4	.007
N	1,184		

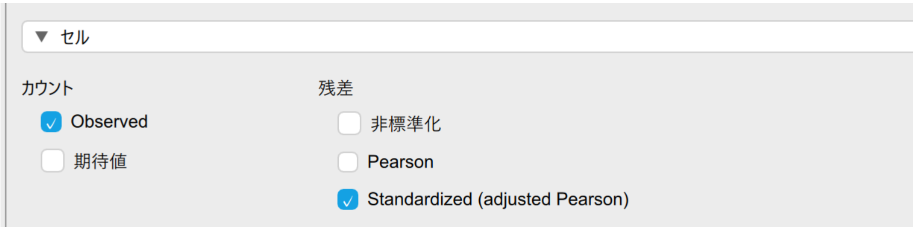
注 Continuity correction is available only for 2x2 tables.

序数ガンマ

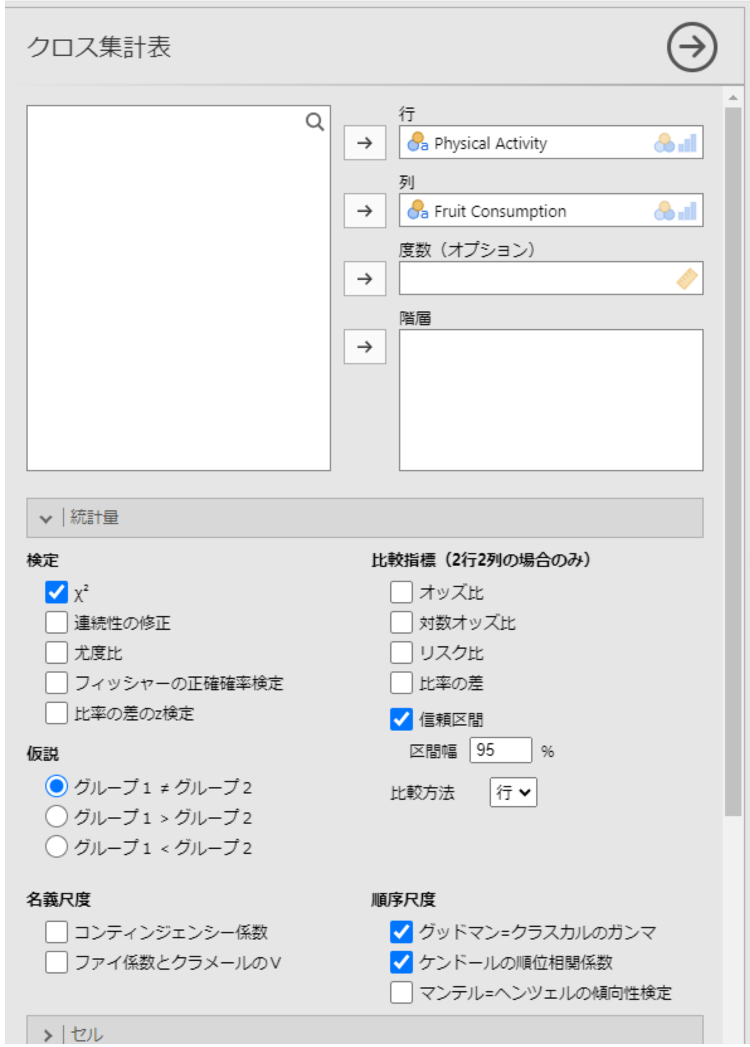
		95% 信頼区間	
ガンマ	標準誤差	下限	上限
0.103	0.044	0.016	0.189

Kendall's Tau

ケンドールのタウb	Z	p
0.061	2.296	.022



3×3のクロス集計(順序) jamovi



クロス集計表

	Fruit Consumption			
Physical Activity	Low	Medium	High	全体
Low	69	25	14	108
Moderate	206	126	111	443
Vigorous	294	170	169	633
全体	569	321	294	1184

χ^2 検定

	値	自由度	p
χ^2	14.2	4	0.007
N	1184		

グッドマン=クラスカルのガンマ

		95%信頼区間	
グッドマン=クラスカルのガンマ	標準誤差	下限	上限
0.103	0.0441	0.0162	0.189

[3]

ケンドールの順位相関係数

ケンドールの順位相関係数	t	p
0.0612	2.30	0.022

事後検定(標準化残差)

	Fruit Consumption		
Physical Activity	Low	Medium	High
Low	3.454	−0.972	−2.995
Moderate	−0.829	0.797	0.139
Vigorous	−1.190	−0.212	1.594

標準化残差（standardized residual）

- 各セルの「観測度数 − 期待度数」がどの程度大きいかを標準偏差単位で示した値
- これをさらに分割表全体の分散構造を考慮して補正したのが「調整済み標準化残差（adjusted standardized residual）」
  - 「調整済みピアソン」とも言う

標準化残差(standardized residual)

絶対値	解釈	備考
約1以下	偶然の範囲	特に偏りなし
約1.96以上	5%水準で有意	有意な偏り
約2.58以上	1%水準で有意	強い偏り

分割表 ▼

Physical Activity		Fruit Consumption Low	Fruit Consumption Medium	Fruit Consumption High	合計
Low	Count	69.00	25.00	14.00	108.0
	Standardized residuals	3.45	−0.97	−2.99	
Moderate	Count	206.00	126.00	111.00	443.0
	Standardized residuals	−0.83	0.80	0.14	
Vigorous	Count	294.00	170.00	169.00	633.0
	Standardized residuals	−1.19	−0.21	1.59	
合計	Count	569.00	321.00	294.00	1,184.0

グッドマン＝クラスカルのガンマ（Goodman–Kruskal's γ）

- 順序カテゴリ間の関係（クロス表）を評価するための指標
- ２つの順序変数の一致（C：concordant pairs ）・不一致の組（D：discordant pairs）の差に基づく。

$\gamma = 1$ → 完全に一致（すべてのペアが同じ方向）

$\gamma = -1$ → 完全に逆方向

$\gamma = 0$ → 一致と不一致が同じくらい（関係なし）

ケンドールの順位相関係数(Kendall’s τ)

- グッドマン＝クラスカルのガンマと似ているが同順位も考慮
  - より厳密になる
  - グッドマン＝クラスカルの方が値が大きめ
  - −1～1 の範囲
    - ±0.3 くらい：弱い関係
    - ±0.5 前後：中程度
    - ±0.7 以上：強い関係

グッドマン＝クラスカルのγと
ケンドールの順位相関（τb）

- γ は「クロス表で傾向をざっくり見る」ためのもの
- τ は「順位データの一致度を精密に見る」ためのもの
- γは単純な方向一致率、τは同順位も含めた厳密な一致度を表す

グッドマン＝クラスカルのγと ケンドールの順位相関（τb）

- χ^2乗検定で有意確率がある程度小さい（標準化残差の大きいところがある）けど、γやτが低い

→行列の間に関係性はあるけど、一貫した関係性（順序性や上昇・下降傾向）はない

「分割が多いクロス表」は扱いにくい

- 期待度数が小さくなりやすい
  - セルが増えると、1セルあたりのデータ数が減る
  - その結果、「期待度数＜5」のセルが増えて、χ^2検定の前提(大標本近似)が崩れる
  - フィッシャー検定でも、分割が大きくなると計算が膨大(ほぼ不可能)になる
    - JASPは2×2まで。jamoviはそれ以上でもやるけど途中で止まる

「分割が多いクロス表」は扱いにくい

- 「どこに差があるか」が直感的に見えない
  - 2×2なら「多いか少ないか」がすぐわかる
  - 3×4とかになると、全体で有意でも「どのセルが寄与しているのか」が読みにくい
  - 標準化残差で見るにしても、±1.96以上のセルが点在していて、説明しにくい（順序変数だけど順序性がない場合の説明をどうするか）

クロス表作成上の注意

- あまり大きなクロス表は作らない
- クロス集計表はシンプルなほど関係が見えやすい
  - 直感的に結果が説明できる、2×2クロスがベスト
- 分割が増えると情報が増えても、期待度数が少なくなるので、結果の信頼性が下がる場合がある
- 3×3以上は、基本的には、集約・再分類を検討した方がいい
  - できれば、2×2、せめて2×3くらいに整理して考えるのが基本

対応のあるデータ

対応のあるデータ

- 同じ対象で複数回のデータ
- 前の回と後の回で解答傾向が同じか違うか
- 違っているなら、前と後の間で何かがあったと考える

対応の無いデータ

- クラスとおやつ

No	クラス	おやつ
1	きつね	きのこの里
2	たぬき	きのこの里
3	たぬき	きのこの里
4	きつね	きのこの里
5	きつね	たけのこの山
6	たぬき	きのこの里
7	きつね	たけのこの山
8	たぬき	きのこの里

クラス別の希望するおやつ		おやつ		
		きのこの里	たけのこの山	計
クラス	たぬき	20	10	30
	きつね	12	18	30
	計	32	28	60

対応のあるデータ

- 就職前後で、朝食を食べている、食べていない

番号	就職前	就職後
1	食べている	食べている
2	食べている	食べていない
3	食べている	食べている
4	食べている	食べていない
5	食べていない	食べている
6	食べている	食べている
7	食べている	食べていない
8	食べている	食べていない

		就職後		
		食べている	食べていない	合計
就職前	食べている	35	25	60
	食べていない	15	25	40
	合計	50	50	100

対応のない、ある

- データを見る

No	クラス	おやつ
1	きつね	きのこの里
2	たぬき	きのこの里
3	たぬき	きのこの里
4	きつね	きのこの里
5	きつね	たけのこの山
6	たぬき	きのこの里
7	きつね	たけのこの山
8	たぬき	きのこの里

番号	就職前	就職後
1	食べている	食べている
2	食べている	食べていない
3	食べている	食べている
4	食べている	食べていない
5	食べていない	食べている
6	食べている	食べている
7	食べている	食べていない
8	食べている	食べていない

対応のない、ある

- クロス集計では一見わかりにくい

クラス別の希望するおやつ		おやつ		
		きのこの里	たけのこの山	計
クラス	たぬき	20	10	30
	きつね	12	18	30
	計	32	28	60

		就職後		
		食べている	食べていない	合計
就職前	食べている	35	25	60
	食べていない	15	25	40
	合計	50	50	100

マクネマーの検定ともいいます

対応のあるχ^2検定

対応のあるデータ

- 就職前後で、朝食を食べている、食べていない

番号	就職前	就職後
1	食べている	食べている
2	食べている	食べていない
3	食べている	食べている
4	食べている	食べていない
5	食べていない	食べている
6	食べている	食べている
7	食べている	食べていない
8	食べている	食べていない

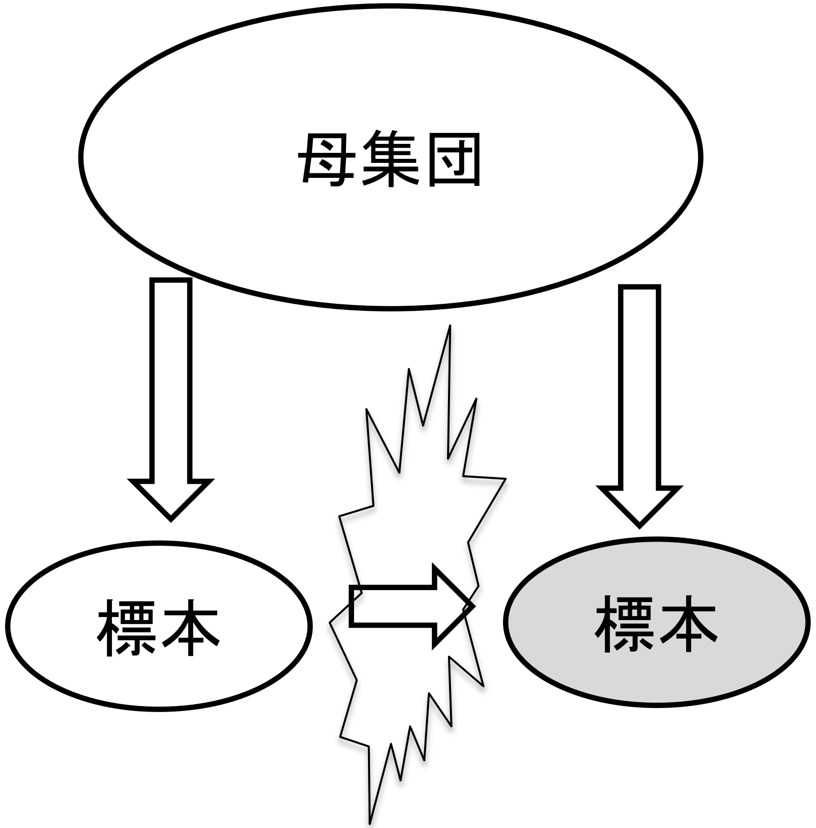
		就職後		
		食べている	食べていない	合計
就職前	食べている	35	25	60
	食べていない	15	25	40
	合計	50	50	100

対応のある検定とは

- データ間に対応がある場合は別の計算方法を使う
- 対応のあるデータ（繰り返しのあるデータ）
  - 同じ対象に対して複数回データを取っている
- 統計値の計算方法は異なるが、結果の分布は同じ
- 有意差がある場合は、その間に何かがあったと考える

対応のあるデータの検定の考え方

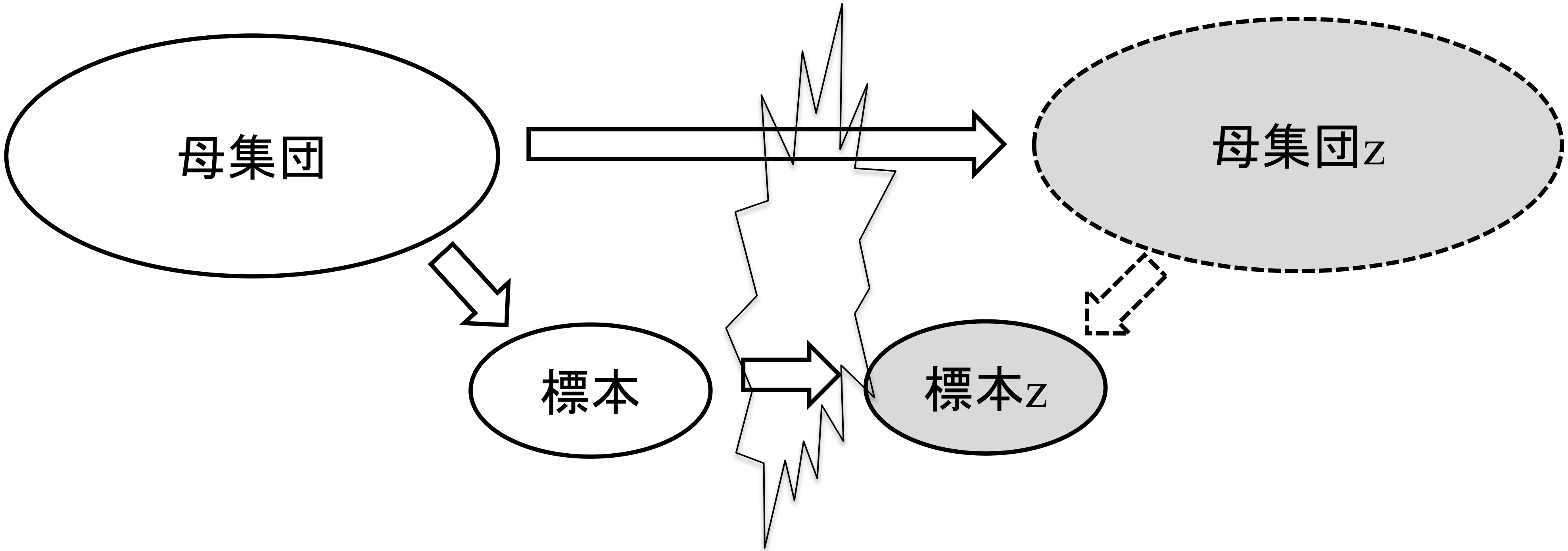
- 前後で違いがあるかないか



違ってない？
前後で違うなら、その間にあったことが影響していると考える

対応のあるデータの検定の考え方

- 前後で違いがあるかないか
  - 標本が違うなら、もう同じ母集団とは言えない



対応がある場合の帰無仮説の考え方

- 棄却されたとき
  - 今回のサンプルからは両群が同じだと言うことは難しい（同じである確率は低い）→変化があった
- 棄却されなかったとき
  - 今回のサンプルからは両群が同じであると推定できる（同じである確率が高い）→変化がなかった

対応のあるχ^2検定

- 同じ人の繰り返しデータ
- 変化があったのは緑色の部分
- その部分の影響がわかればよい

		就職後		
		食べている	食べていない	合計
就職前	食べている	35	25	60
	食べていない	15	25	40
	合計	50	50	100

対応のあるχ^2検定：マクネマーの検定

- $(b-c)$の2乗を$(b+c)$で割った値は自由度1のχ^2分布に従う

$$Z=\frac{(b-c)^2}{(b+c)}$$

  - 計算の仕方は違う（関数を使わない）けど、この値がχ^2分布になる
  - 期待度数表は作らない

		就職後		
		食べている	食べていない	合計
就職前	食べている	a 35	b 25	60
	食べていない	c 15	d 25	40
	合計	50	50	100

χ^2検定の結果の判断(再)

- P値を直接計算できない場合
  - 右のような確率分布表を使う
    - 対応無くてもあっても同じ
  - 自由度と、主な有意確率の時のχ^2値の表
    - 自由度1でχ^2値が3.84のとき有意確率は0.05
    - 自由度1でχ^2値が6.63のとき有意確率は0.01
    - 自由度1でχ^2値が10.83のとき有意確率は0.001

		有意確率			
		0.10	0.05	0.01	0.001
自由度	1	2.71	3.84	6.63	10.83
	2	4.61	5.99	9.21	13.82
	3	6.25	7.81	11.34	16.27
	4	7.78	9.49	13.28	18.47
	5	9.24	11.07	15.09	20.52
	6	10.64	12.59	16.81	22.46
	7	12.02	14.07	18.48	24.32
	8	13.36	15.51	20.09	26.12
	9	14.68	16.92	21.67	27.88
	10	15.99	18.31	23.21	29.59

Excelで行う対応のあるχ^2検定

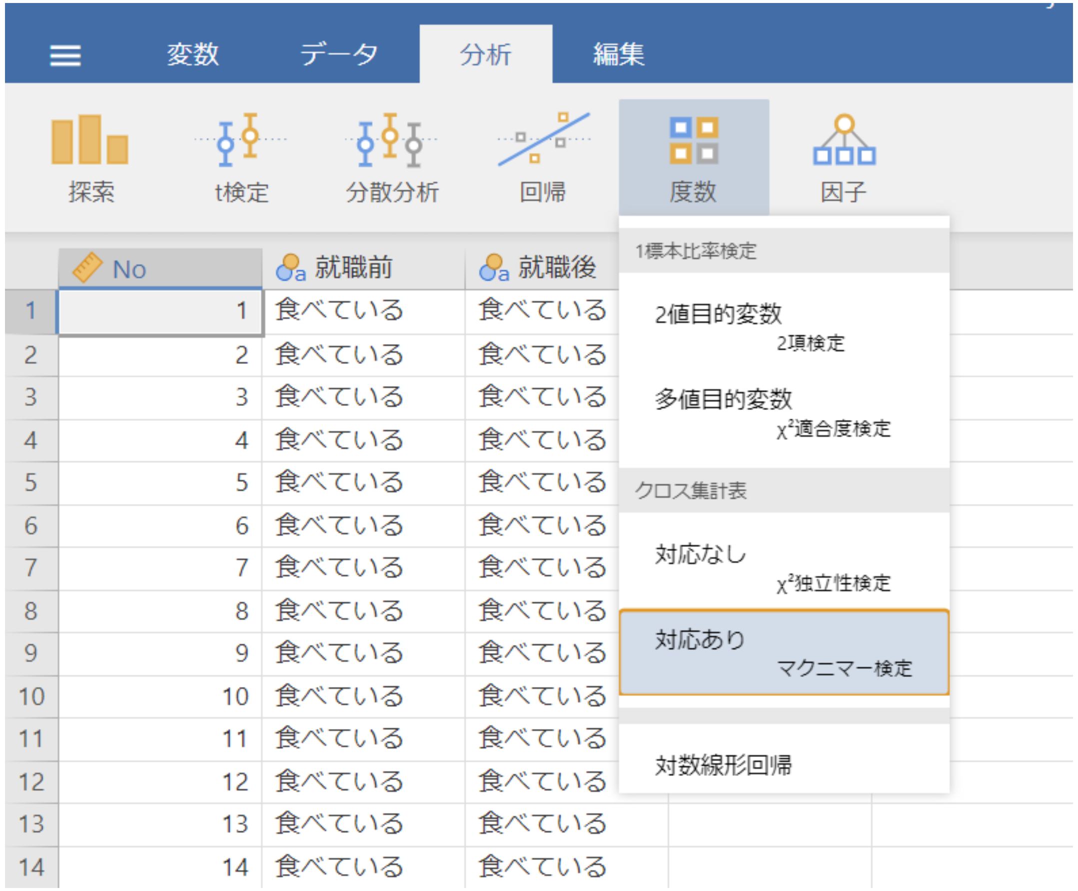
- χ^2値
  - 式の計算で出せる
  - 期待度数表はいらない
  - CHISQ.TEST関数では出せない
- p値
  =CHISQ.DIST.RT（χ^2値,自由度）
  - χ^2値と自由度から、そのχ^2値に該当する確率値を算出
  - 自由度はクロス集計表から求めるのでこの場合は1

実は

- JASPにはマクネマー検定がありません
  - Rにはあるので、コマンドで実行は可能
  - 手計算でいけるのと、後述の理由でプライオリティが低いと考えられている
- jamoviにはあります

jamoviでやるマクネマー検定

- data05_01



jamoviでやるマクネマー検定

クロス集計表

	就職後		
就職前	食べている	食べていない	全体
食べている	35	25	60
食べていない	15	25	40
全体	50	50	100

マクニマー検定

	値	自由度	p
χ^2	2.50	1	0.114
N	100		

データの大きさの問題

結果

対応ありクロス集計表

クロス集計表

	就職後		
就職前	食べている	食べていない	全体
食べている	35	25	60
食べていない	15	25	40
全体	50	50	100

マクニマー検定

	値	自由度	p
χ^2	2.50	1	0.114
N	100		

結果

対応ありクロス集計表

クロス集計表

	就職後		
就職前	食べている	食べていない	全体
食べている	175	125	300
食べていない	75	125	200
全体	250	250	500

マクニマー検定

	値	自由度	p
χ^2	12.5	1	< .001
N	500		

2×2より大きな繰り返し

- マクネマーではなくバウカー検定
  - Bowker’s test:拡張マクネマー検定
- 3×3が限界かなぁ
- 解釈に順序性が出てくるので、そこがうまく説明できるかどうか
- 順序性を考えるならWilcoxon符号付順位検定
  - ノンパラで説明

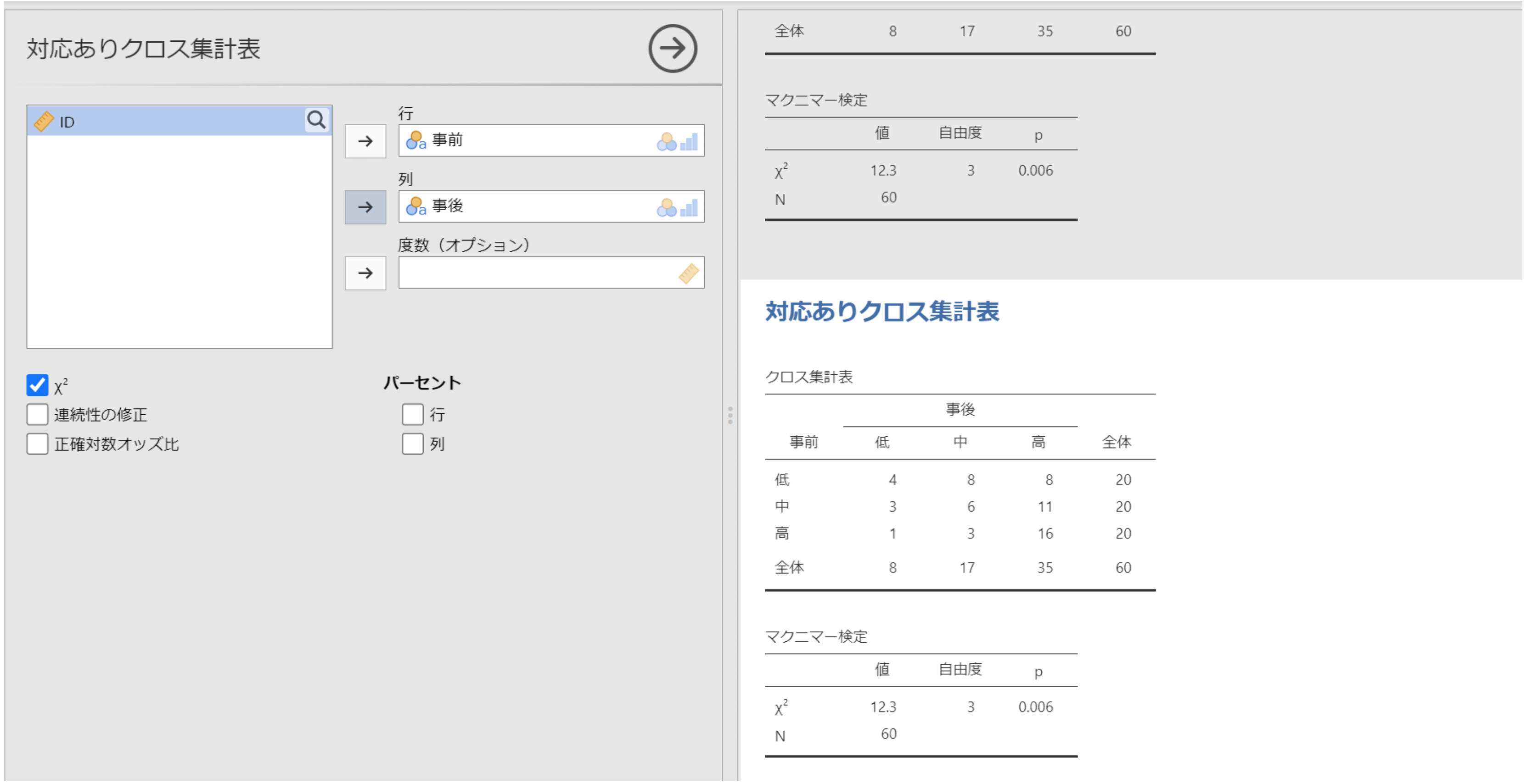
大きな繰り返しのあるデータ

- jamoviだと、対応のあるχ^2検定が、3×3以上になると、自動でバウカー検定になる
  - そもそもマクネマー検定は2×2しかできない

jamoviのバウカー検定 (マクネマーと同じ)

- data05_02

jamoviのバウカー検定



大きな繰り返しのあるデータ

- あまり進めない理由
  - データがたくさん必要
  - バウカー検定は順序性は見ていない
  - カテゴリ数が多い場合、そのカテゴリは順序性があるのか？
    - そもそも前後の差を見るときに、変わったかどうかだけなのか、何らかの順序性を仮定しないのか

    →分析手法では無くて、研究計画の問題

クロス集計では無いのかも

- 基本的にχ^2検定は比率の差の検定（前回説明）
  - 厳密に言うと期待度数との乖離を見ている
  - なので、観測度数が多くなると、差が出る傾向がある
  - 順序性を考慮する場合は、グッドマン＝クラスカルのγやケンドールの順位相関（τb）を見る
- 対応がある場合（マクネマー、バウカー）は、対称性の検定
  - 変化無しを境にして、どちらかの変化が多いかを見ている
  - 方向性までは見ていない

対称性

- 対称軸を境に同じ比率かどうかを見ているだけ
- 順序（方向）は見ていない

		事後			
		低	中	高	合計
事前	低	4	8	8	20
	中	3	6	11	20
	高	1	3	16	20
	合計	8	17	35	60

カテゴリに順序性があるのなら

- クロス集計をしてマクネマーやバウカー検定はしない
  - ウィルコクソン符号付順位検定を選んだ方がいい
    - 「10回目　ノンパラメトリック検定」で説明
    - 対応のある順序変数の検定
- 順序性を考慮しなくてはいけないかどうかは、研究計画時点での問題

カテゴリに順序性がある場合

- 2値（はい・いいえ）の場合、変化＝順序と見なせる
- 3値だと、順序性があるかないかを考える必要がある
  - 大抵は「よい・ふつう・わるい」のように順序性がある場合が多い
  - 同じ質的変数でも、名義尺度か順序尺度かはこういうときに気をつけないといけない
  - どういうデータにするかは研究計画でちゃんと決める