

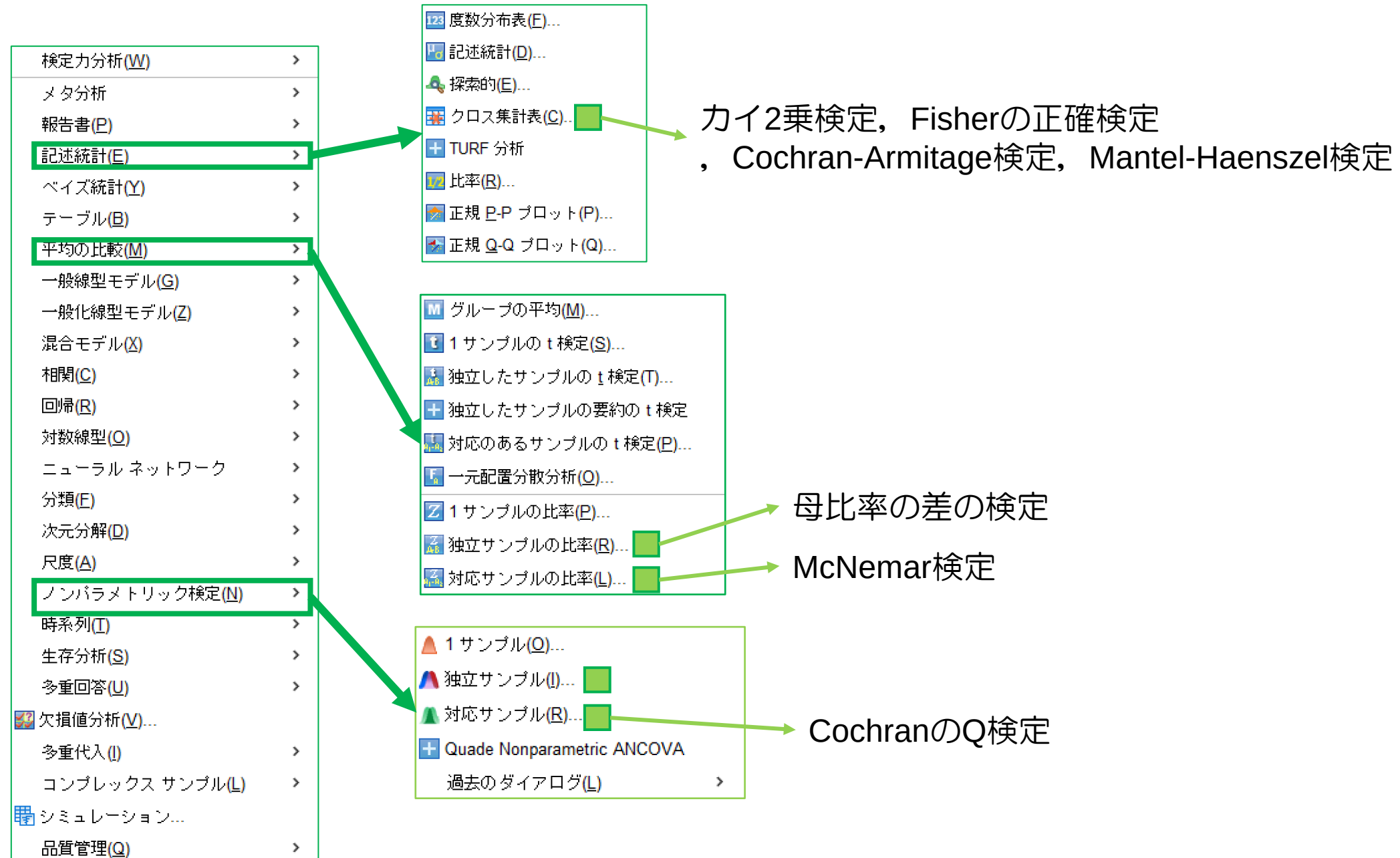
看護統計セミナー²⁰²²

第2回目：質的データの解析

下川敏雄

和歌山県立医科大学附属病院 臨床研究センター

本日のおはなし：SPSSメニューを例に

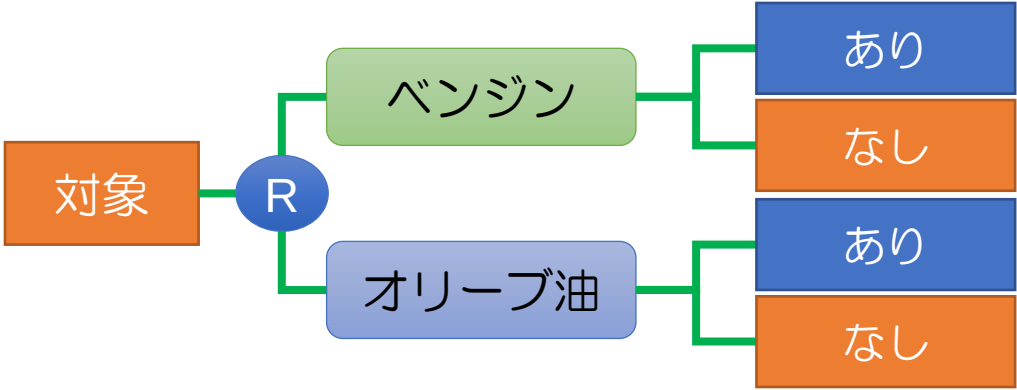


カイ2乗検定およびFisherの正確検定の動機



Clinical Question

テープ剥離に用いるベンジンが皮膚かぶれの原因と考え、剥離材にオリーブ油を利用した。このとき、剥離剤(ベンジン・オリーブ油)と皮膚かぶれの発現の有無には違いがあるだろうか。



	かぶれあり	かぶれなし	計
オリーブ油	24 (32.4%)	100 (60.6%)	124
ベンジン	50 (43.5%)	65 (56.5%)	115
計	74 (31.0%)	165 (69.0%)	239

統計的方法

検定：カイ2乗検定(標本サイズ多数)
Fisherの正確検定(標本サイズ少数)

要約：各群の所見有りの割合
オッズ比及び95%信頼区間

カイ2乗検定およびFisherの正確検定の意味

考え方(1)：因果関係の観点から見る場合

帰無仮説 H_0 ：剥離剤の種類によって かぶれの有無 に違いがない。

対立仮説 H_1 ：剥離剤の種類によって かぶれの有無 に違いがない。

原因

結果

	かぶれあり	かぶれなし
オリーブ油	24 (32.4%)	100 (60.6%)
ベンジン	50 (43.5%)	65 (56.5%)

仮説は若干異なるものの、
検定方法は同じ



考え方(2)：関連性(独立性)の観点から見る場合

帰無仮説 H_0 ：剥離剤の種類 と かぶれの有無 に関連性がない(独立である)。

対立仮説 H_1 ：剥離剤の種類 と かぶれの有無 に関連性がある(独立でない)。

変数1

変数2

カイ2乗検定の意味

● 帰無仮説 H_0 が正しいときのクロス集計表とデータから得られたクロス集計表の乖離度を応用する。

✖ p値を計算する際に近似を用いるため、症例数が少ない等の理由で近似精度が悪くなる可能性がある。

Fisherの正確検定の意味

● すべてのクロス集計表を考え、そのクロス集計表が得られる確率を数学的(確率的)に計算する。

✖ すべてのパターンのクロス集計表を計算するため、計算負荷がかかるため、症例数が多くなるとPCがフリーズする可能性がある。

カイ2乗検定とFisherの正確検定の使い分け：まとめ

	有効	無効	計
治療A	12 (54.5%)	10 (45.5%)	22
治療B	3 (20.0%)	12 (80.0%)	15
計	15 (40.5%)	22 (59.5%)	37

カイ2乗検定

p値 = 0.0783

有意差が認められない

Fisherの正確検定

p値 = 0.0471

有意差が認められる

カイ2乗検定では、p値を**近似**で計算する。そのため、

(1) 標本サイズが少ない場合 (50未満)

(2) どこかのセルの度数が小さい場合 (5未満)

には、近似精度が悪くなる。そのため、Fisherの正確検定を用いるほうが良い。一方で、標本サイズが大きくなると、Fisherの正確検定は計算できなくなる(パソコンがフリーズする)ので、カイ2乗検定を推奨する。なお、標本サイズが増えると、カイ2乗検定のp値とFisherの正確検定のp値は一致する。

カイ2乗検定を用いるときの注意点：2×2クロス集計表では連続性は補正すること

クロス集計表では、症例数のほかに列数および行数が少ない場合、とくに2×2クロス集計表の場合には近似精度が悪くなる (p値が小さくなりますが正確でない可能性がある)。そのため、Yatesの補正を行ったときのp値を採用すること。

	かぶれあり	かぶれなし
オリーブ油	44	100
ベンジン	50	67

❌ カイ2乗検定(補正なし) p-value= 0.041

● カイ2乗検定(補正あり) p-value= 0.056 (n.s.)

● Fisherの正確検定 p-value = 0.052 (n.s.)

4) 統計解析

数量化データの分析は Mann-Whitney U 検定, カテゴリーデータの分析は χ^2 検定を用いた。

このp値はYatesの補正をおこなっていない。さらに症例数が少ないため、本来はFisherの正確検定で行わなければならない。

本当のp値は、0.314, 1.000
見栄えが良いからといって、補正なしは選ばないこと

表2 対象者の属性

	安定群 (n=27)	再入院群 (n=13)	p
(前略)			
退院後の生活状況			
同居の援助者 (有：無, 人)	16 : 11	5 : 8	.217
所得 (有：無, 人)	24 : 3	11 : 2	.702

(後略)

SPSSにおける出力結果

「分析」→「記述統計」→「クロス集計表」➡

	a テーブ離脱	a 皮膚かぶれ
1	ベンジン	なし
2	ベンジン	あり
3	ベンジン	なし
4	オリーブ油	なし
5	ベンジン	なし
6	オリーブ油	なし

- 「正確確率」において「正確」をチェック
- 「統計量」において「カイ2乗」をチェック
- 「セル」の「パーセンテージ」において「行」をチェック

カイ 2 乗検定

	値	自由度	漸近有意確率 (両側)	正確な有意確率 (両側)	正確有意確率 (片側)
Pearson のカイ 2 乗	16.243 ^a	1	<.001	<.001	<.001
連続修正 ^b	15.135	1	<.001		
尤度比	16.475	1	<.001	<.001	<.001
Fisher の直接法				<.001	<.001
有効なケースの数	239				

a. 0 セル (0.0%) は期待度数が 5 未満です。最小期待度数は 35.61 です。

b. 2x2 表に対してのみ計算

- SPSSでは、2×2クロス集計表の場合に、「連続修正」の結果が現れる。カイ2乗検定におけるp値の計算には近似式が用いられる。連続修正は、その近似精度を高めるための補正を行っている(こちらのp値のほうが正確である)。
- 尤度比とは異なる接近法で計算されたものであり(厳密には対数線形モデルのもとで計算されたもの)、カイ2乗検定とほぼ同じになる。
- 今回の場合には、全体で100例を超えないので、Fisherの正確検定(SPSSではFisherの直接法)が最も信頼できる。

クロス集計表

行(O):
a 皮膚かぶれ

列(C):
a テーブ離脱

層 1 / 1
前(V) 次(N)

正確確率(X)...

統計量(S)...

セル(E)...

書式(E)...

スタイル(L)...

ブートストラップ(A)...

☐ クラスター棒グラフの表示(B)

☐ クロス集計表の非表示(I)

OK

貼り付け(P)

戻す(R)

キャンセル

ヘルプ

☒ テーブル層に層変数を表示する(L)

カイ2乗検定と母比率の差の検定

下表はある疾患に対する2種類の治療法の有効率を評価したときのクロス集計表である。治療法によって有効/無効に違いがあるか。

	有効	無効	計
治療A	204 (51.1%)	195 (48.9%)	399
治療B	158 (39.4%)	243 (60.6%)	401
計	362 (45.2%)	438 (54.2%)	800

帰無仮説：治療法によって有効/無効に違いが**ない**。
対立仮説：治療法によって有効/無効に違いが**ある**。



クロス集計表の解析において用いられる検定は、

- ・カイ2乗検定
- ・Fisherの正確検定 (Fisherの直接確率)

である。ただし、 2×2 のクロス集計表では、比率を比較するという観点もある。その場合には、母比率の差の検定というのもある。実は、

母比率の差の検定とカイ2乗検定の結果は一致する

母比率の差に対する統計的推測

2値アウトカムにおける関心のある事象の割合(例：治療成功割合)を比較する方法として、母比率の差の検定がある。

帰無仮説 H_0 ：群1の比率 p_1 = 群2の比率 p_2 ， 対立仮説 H_1 ：群1の比率 $p_1 \neq$ 群2の比率 p_2

SPSSでは、「分析」→「平均の比較」→「独立サンプルの比率」で実行可能

注意：変数の型は「数値」でなければならない

独立したサンプルの比率のグループ統計量

テーブル離脱		成功数	試行	比率	漸近標準誤差
皮膚かぶれ = 1	= 1	24	124	.194	.035
	= 0	50	115	.435	.046

独立したサンプルの比率の信頼区間

区間のタイプ		比率の差	漸近標準誤差	差の 95% 信頼区間	
				下限	上限
皮膚かぶれ = 1	Agresti-Caffo	-.241	.058	-.351	-.124
	Newcombe	-.241	.058	-.350	-.124

独立したサンプルの比率検定

検定の種類		比率の差	漸近標準誤差	Z	有意	
					片側 p 値	両側 p 値
皮膚かぶれ = 1	Wald H0	-.241	.058	-4.030	<.001	<.001

SPSSでは、母比率の差に対する信頼区間として、

- Agresti-Caffo • Brown-Li-Jeffreys • Hauck-Anderson
- Newcombe • Newcombe (連続修正) • Wald
- Wald (連続修正)

が用意されている。WaldおよびWald(連続補正)が教科書で記載されている広く知られた方法である。一方で、Newcombeが最も信頼性がおける方法であることが報告されている。

SPSSでは、母比率の差の検定として、

- Hauck-Anderson • Wald • Wald (連続修正)
- Wald H0 • Wald H0 (連続修正)

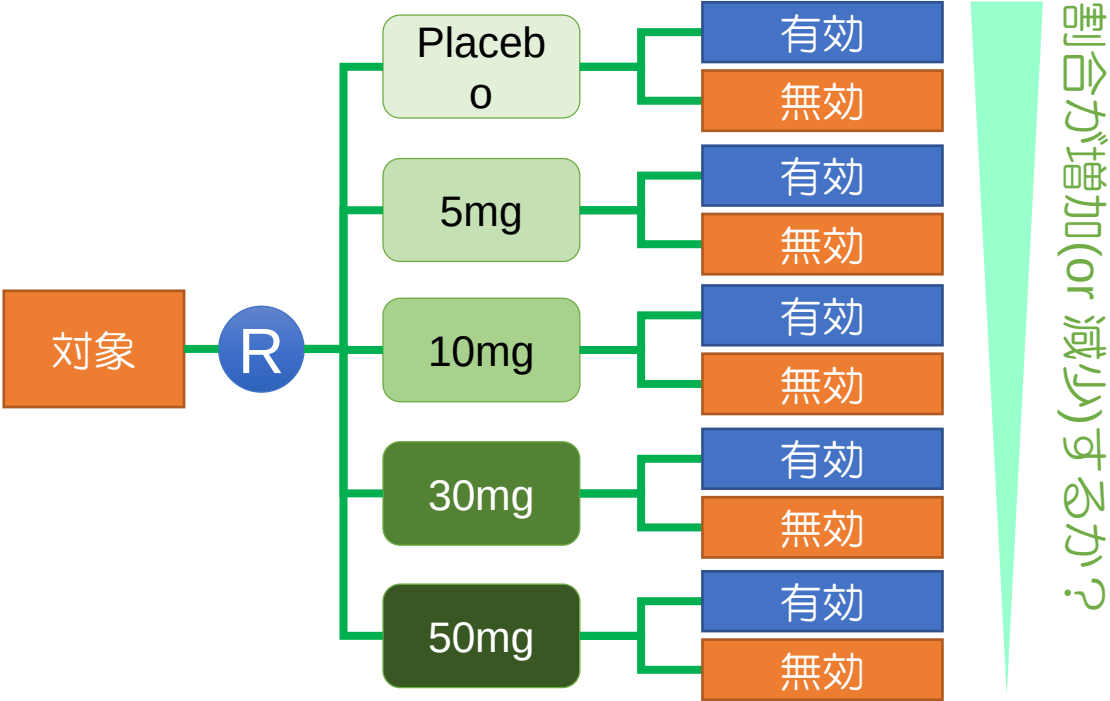
が用意されている。WaldおよびWald(連続補正)が教科書で記載されている広く知られた方法である。一方で、Wald H0およびWald H0(連続修正)は、分散推定に共通の比率を用いた方法である。

順序関係があるグループにおける傾向性の評価：Cochran-Armitage検定

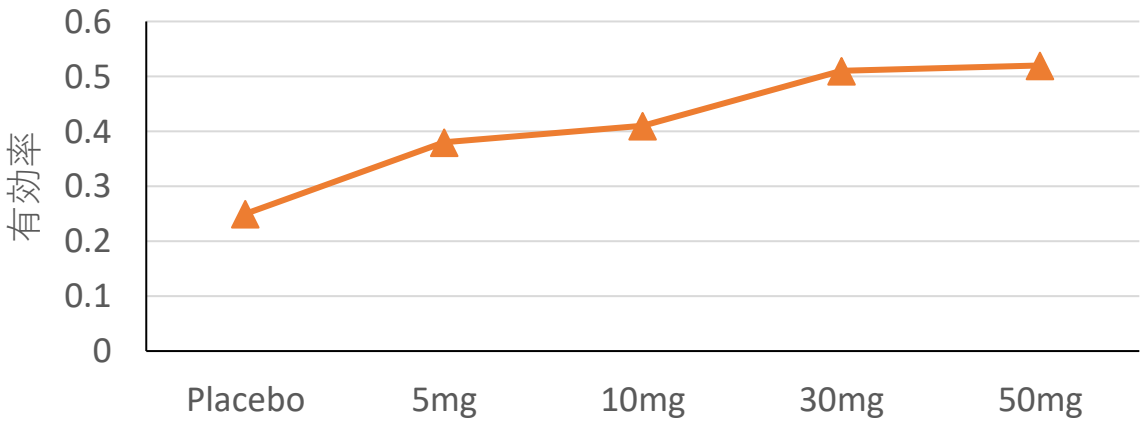


Clinical Question

ある薬剤の用量反応関係を調査するために、Placebo, 5mg, 10mg, 30mg, 50mgの投与群とそのときの効果の有無を調べた。



	Placebo	5mg	10mg	30mg	50mg
有効	79	81	169	318	379
無効	235	132	243	305	204



統計的方法

検定：Cochran-Armitage検定

要約：有効割合

順序関係があるグループにおける傾向性の評価：Cochran-Armitage検定

Cochran-Armitage検定の例示

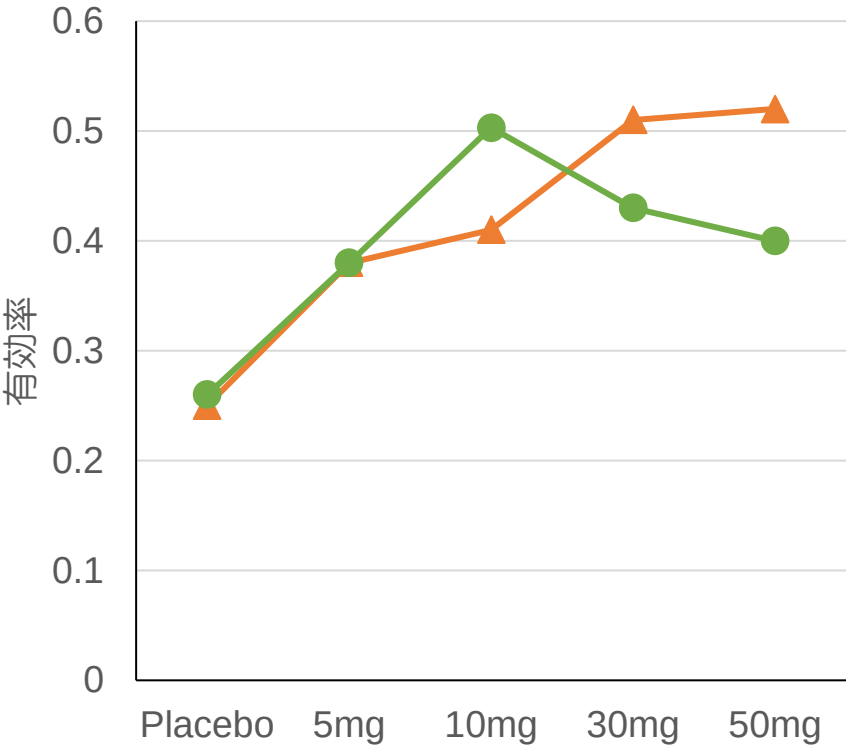
2種類の薬剤の投与量と有効性を評価したときのクロス集計表を下表に示す。

薬剤A

	Placebo	5mg	10mg	30mg	50mg
有効	79	81	169	318	379
無効	235	132	243	305	204

薬剤B

	Placebo	5mg	10mg	30mg	50mg
有効	82	81	220	268	233
無効	232	132	214	355	350



薬剤Aの場合には投与量が増加するにつれて有効性が増加しているものの、薬剤Bでは減少している。



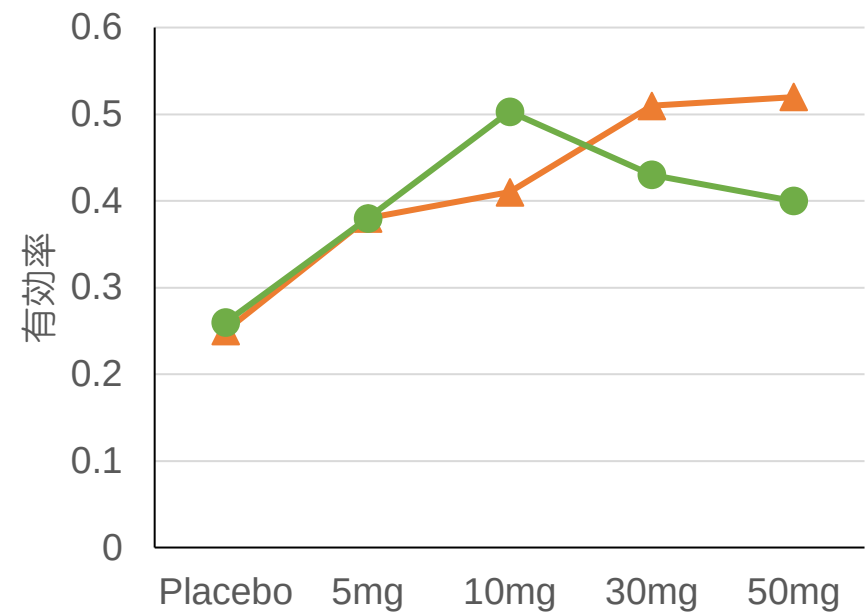
このような研究を計画する理由としては、投与量による有効性の増加傾向(あるいは減少傾向)を評価することにある。

Cochran-Armitage検定

薬剤A **p値<0.001** 投与量によって有効率の上昇傾向が認められた

薬剤B **p値=0.180** 投与量によって有効率の上昇傾向が認められなかった

Cochran-Armitage検定とカイ2乗検定の違い



薬剤A

	Placebo	5mg	10mg	30mg	50mg
有効	79	81	169	318	379
無効	235	132	243	305	204

薬剤B

	Placebo	5mg	10mg	30mg	50mg
有効	82	81	220	268	233
無効	232	132	214	355	350

Cochran-Armitage検定

薬剤A **p値<0.001** 投与量によって有効率の上昇傾向が認められた

薬剤B **p値=0.180** 投与量によって有効率の上昇傾向が認められなかった

カイ2乗検定

薬剤A **p値<0.001** 投与量が薬剤の有効／無効に影響していた。

薬剤B **p値<0.001** 投与量が薬剤の有効／無効に影響していた。

Cochran-Armitage検定とカイ2乗検定では見ている視点が異なる

Cochran-Armitage検定の実例

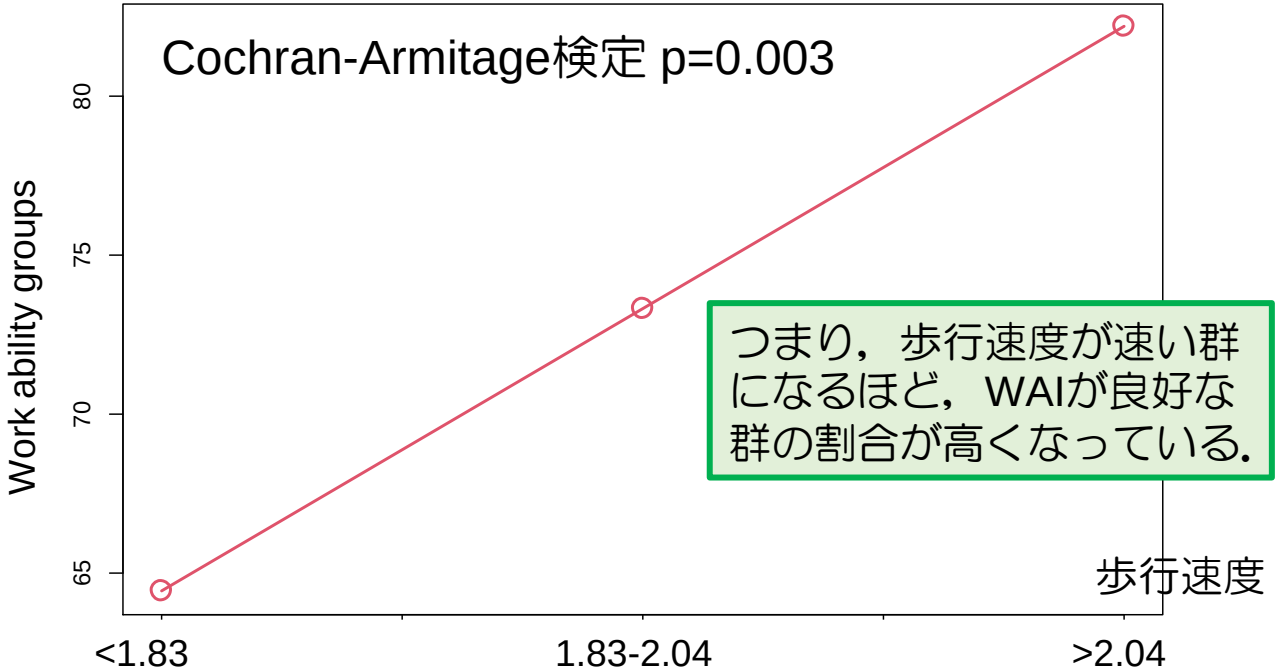
勤務中の病院看護スタッフの最大歩行速度と労働能力の関連性に関する横断研究

Kohl et al. : Journal of Occupational Health, 62(1), 7-14, 2020.

看護師／看護アシスタント312名と対象に、歩行速度と労働適応能力指標(WAI)の項目1(業務能力の最高を10点としたときに、あなたの現在の仕事の能力は難点ですか?)関連性を調査した横断研究である。

Methods: Three-hundred and twelve inpatient nurses and nursing assistants were sampled from a level 1 trauma and teaching hospital from several service lines and acuity levels. Spearman correlation tests were utilized to determine the relationship of gait speed and ratings of item 1 on the Work Ability Index (WAI) as well as **Cochran-Armitage test for linear trend of gait speed**.

	Work ability groups				P-value
	n	%	n	%	
	Moderate to poor		Good to excellent		
	0-7点		8-10点		
Total	83	26.6	229	73.4	
Max gait speed (m/s)					
<1.83	37	35.6	67	64.4	.015 ^a
1.83-2.04	27	26.7	74	73.3	.722
>2.04	19	17.8	88	82.2	.014 ^a



SPSSにおけるCochran-Armitage検定の実行手順

女性の肺癌患者 108 名，対照群 108 名を選出し(0:対照群，1：肺癌患者群)，1日当たりの喫煙量(0：0，1：1～4本，2：5～14本，3：15本～)について調査したケース・コントロール研究の結果がある(丹後, 2013). 肺癌と喫煙量に関連性があるかどうかを検討しなさい.

量的変数		量的変数	
 肺癌		 喫煙	
0	1		
1	0		
1	0		
0	0		
0	0		
0	0		
1	1		
1	1		
1	0		
0	2		
1	2		
1	0		
0	0		
0	0		
1	3		
⋮	⋮		

SPSSの実行方法

- 1. 「分析」→「記述統計」→「クロス集計表」
- 2. 「クロス集計表」のダイアログから「統計量」を選択する.
- 3. 「統計量」ダイアログから「カイ2乗」を選択する.

肺癌 と 喫煙 のクロス表

		喫煙					
		0	1	2	3	合計	
肺癌	0	度数	59	25	18	6	108
		肺癌 の %	54.6%	23.1%	16.7%	5.6%	100.0%
	1	度数	40	16	24	28	108
		肺癌 の %	37.0%	14.8%	22.2%	25.9%	100.0%
合計	度数	99	41	42	34	216	
	肺癌 の %	45.8%	19.0%	19.4%	15.7%	100.0%	

喫煙本数が多いほど，肺がんの罹患率が高くなっている.

カイ 2 乗検定

	値	自由度	漸近有意確率 (両側)
Pearson のカイ 2 乗	20.715 ^a	3	<.001
尤度比	21.967	3	<.001
線型と線型による連関	17.143	1	<.001
有効なケースの数	216		

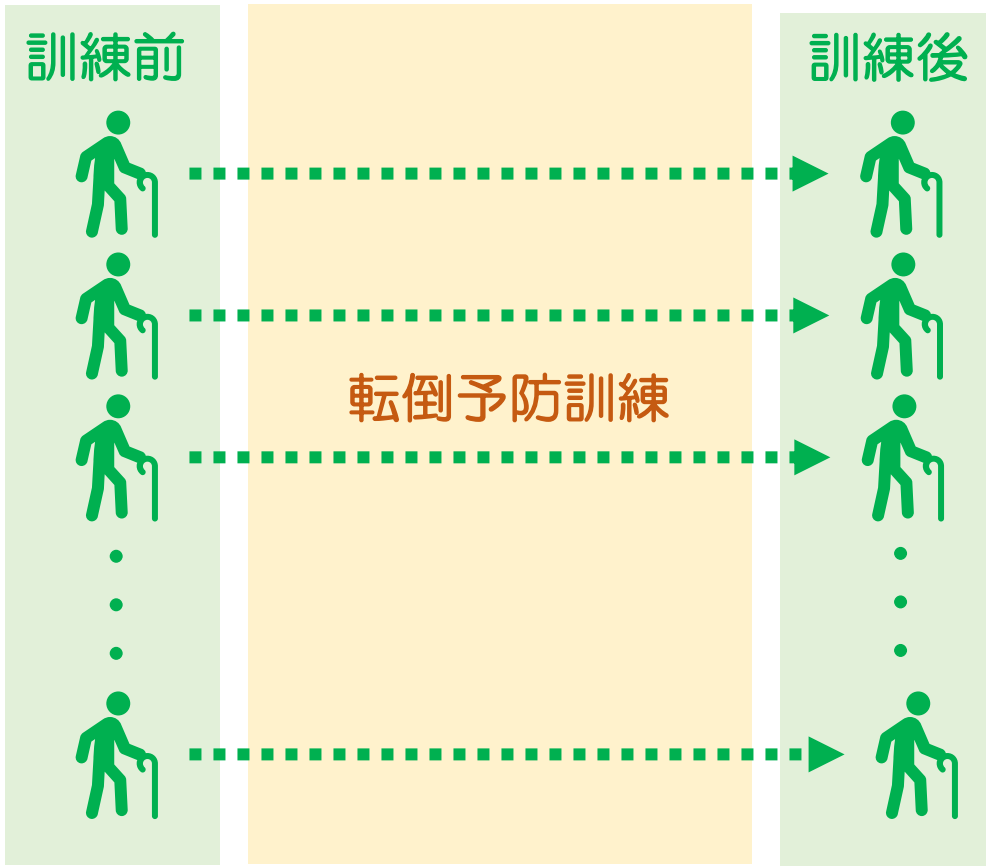
Cochran-Aarmitage検定

a. 0 セル (0.0%) は期待度数が 5 未満です。最小期待度数は 17.00 です。

質的データにも「対応がある場合」が存在します



65歳以上の高齢者を対象に、転倒予防訓練と運動機能の低下の有無に関する研究が実施された。この研究では、200人の被験者に対して、転倒予防訓練前に運動機能検査を行い、3カ月の転倒予防訓練後に同様の検査を実施している。ここでの目標は、転倒予防訓練後に運動機能の低下が改善していることを確認することにある(柳川他, 2011)。



このデータでは、被験者のそれぞれにおいて、訓練前の運動機能の低下の有無、訓練後の運動機能の低下の有無がとられている。

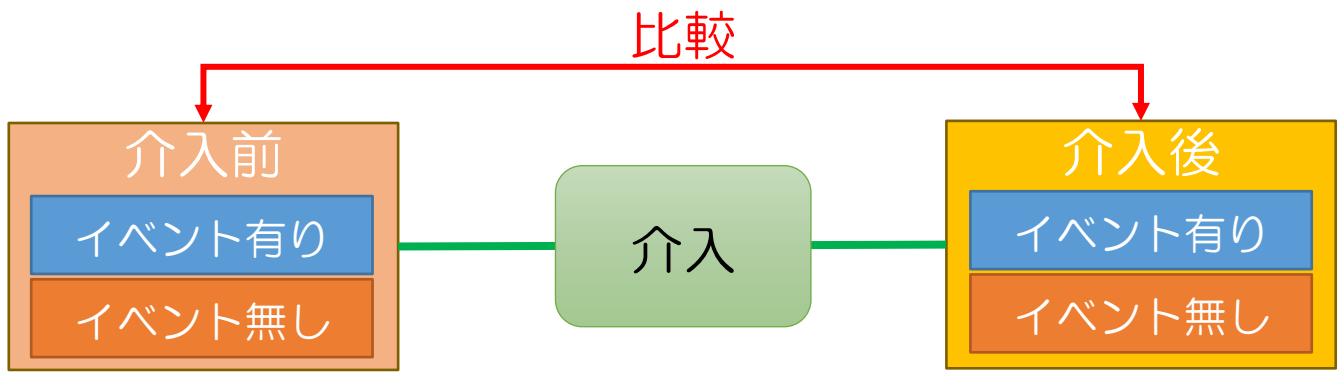
このように、個々の被験者において2個のデータがとられている場合を対応のあるデータという。また、今回の場合には、以下の4パターンが存在する。

- 訓練前に運動機能低下あり，訓練後に運動機能低下あり
- 訓練前に運動機能低下あり，訓練後に運動機能低下なし
- 訓練前に運動機能低下なし，訓練後に運動機能低下あり
- 訓練前に運動機能低下なし，訓練後に運動機能低下なし

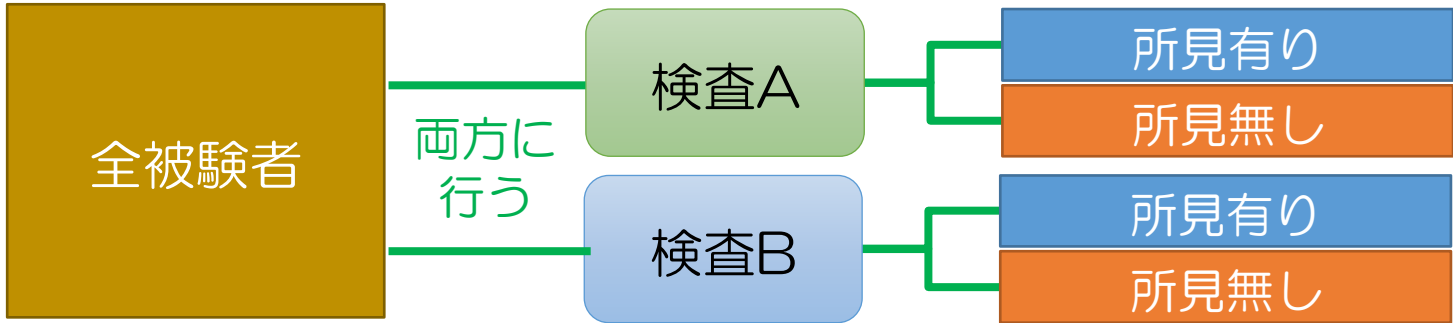
上記のパターンをクロス集計表で表したものを対応のあるクロス集計表という。

対応のあるクロス集計表を用いる幾つかのシチュエーション

Scenario.1
介入前後でのイベントの有無を評価する場合



Scenario.2
同一被験者に検査を実施する場合



Scenario.3
クロスオーバー試験の場合



対応のあるクロス集計表

		訓練後		合計
		低下あり	低下なし	
訓練前	低下あり	40	60	100
	低下なし	30	70	100
合計		70	13	200

• 訓練前に運動機能低下あり，訓練後に運動機能低下あり
 $40/200 = 0.2$ (20%)

■ 訓練前に運動機能低下あり，訓練後に運動機能低下なし
 $60/200 = 0.3$ (30%)

■ 訓練前に運動機能低下なし，訓練後に運動機能低下あり
 $30/200 = 0.15$ (15%)

• 訓練前に運動機能低下なし，訓練後に運動機能低下なし
 $70/200 = 0.35$ (35%)



評価の方法 = McNemar検定

McNemar検定では，■の割合(30%)と■の割合(15%)が比較される。

その結果，p値は0.002であった。したがって，歩行訓練前後で運動機能に変化が認められた。また，歩行機能低下の患者が低下なしになった割合のほうが高いことから，歩行訓練の効果が認められた。

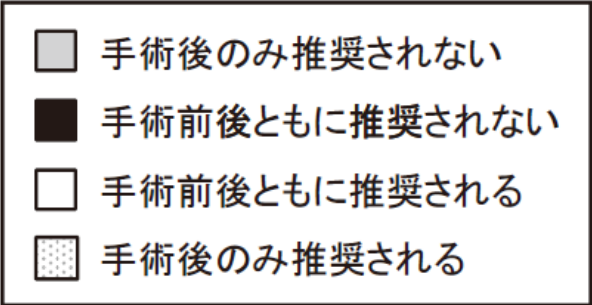
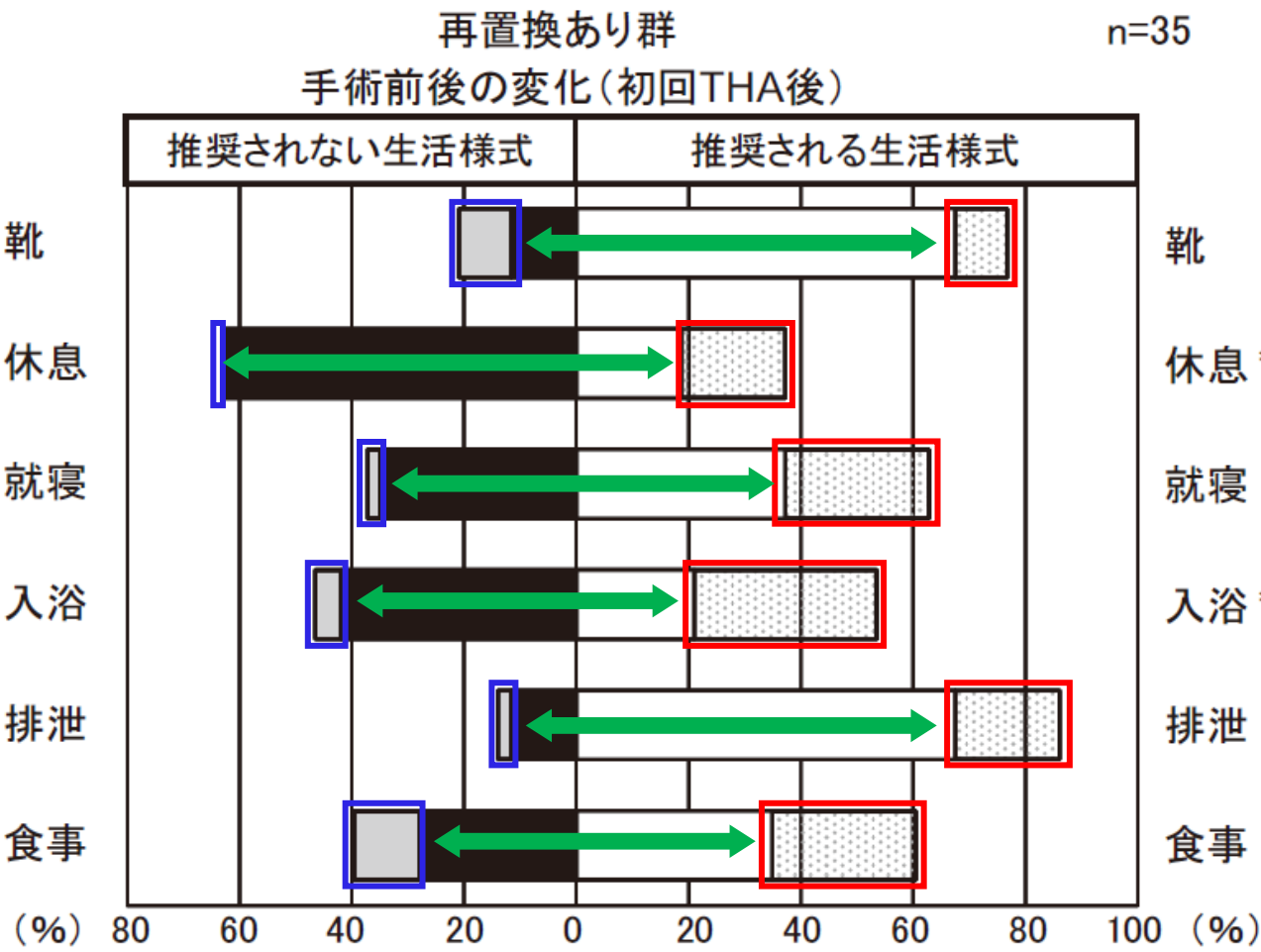
ちなみに，誤ってカイ2乗検定を用いた場合のp値が0.182なので，検定法を誤ることで，本来は有効であった歩行訓練を有効でないと誤認してしまうので注意。

McNemar検定の実例

人工股関節再置換術後患者の再配置に関する住生活環境要因の研究

佐藤：日本看護研究会雑誌, 32(2), 65-74, 2009

この研究は、人工股関節再配置術後患者の手術前後での住生活環境の変化をMcNemar検定により明らかにしている。以下は一部の抜粋である。



このMcNemar検定では、対応のあるクロス集計表

		手術後	
		推奨	非推奨
手術前	推奨	a	b
	非推奨	c	d

のもとで、手術前後で生活環境が変化したかを表している。

つまり、青色で囲った部分と赤色で囲った部分の割合を比較している。

SPSSによるMcNemar検定の実行

65 歳以上の高齢者を対象に、転倒予防訓練と運動機能の低下の有無に関する研究が実施された。この研究では、200 人の被験者に対して、転倒予防訓練前に運動機能検査を行い、3 カ月の転倒予防訓練後に同様の検査を実施している。

 訓練前	 訓練後
低下なし	低下あり
低下あり	低下なし
低下あり	低下なし
低下なし	低下あり
低下あり	低下なし
低下あり	低下なし
低下あり	低下なし
低下なし	低下なし
低下なし	低下なし
低下あり	低下なし
低下あり	低下あり
低下なし	低下なし
低下なし	低下なし
低下あり	低下あり
低下なし	低下なし

・
・
・

・
・
・

SPSSの実行方法

1. 「分析」→「記述統計」→「クロス集計表」
2. 「クロス集計表」のダイアログから「統計量」を選択する。
3. 「統計量」ダイアログから「McNemar」を選択する。

訓練前 と 訓練後 のクロス表

		訓練後		合計
		低下あり	低下なし	
訓練前	低下あり	度数 40	60	100
		総和の % 20.0%	30.0%	50.0%
	低下なし	度数 30	70	100
		総和の % 15.0%	35.0%	50.0%
合計		度数 70	130	200
		総和の % 35.0%	65.0%	100.0%

全体パーセントを用いる(「クロス集計表」ダイアログから「パーセンテージ」のなかの「合計」を選択)。

カイ 2 乗検定

	値	正確な有意確率 (両側)
McNemar 検定		.002 ^a
有効なケースの数	200	

a. 2 項分布を使用

McNemar検定の結果

SPSSによるMcNemar検定の実行：2×2クロス集計表の場合の別計算

変数の型は数値型でなければならない

訓練前	訓練後
0	1
1	0
1	0
0	1
1	0

SPSSの実行方法

- 「分析」→「平均の比較」→「対応サンプルの比率」
- 「クロス集計表」のダイアログから「統計量」を選択する。
- 「対応サンプルの比率」ダイアログにおいて「訓練前」を「変数1」，「訓練後」を「変数2」にドラック&ドロップ。
- 「成功の定義」で「値」を選び，「1」を入力。

対応のあるサンプルの比率の統計量

		成功数	試行	比率	漸近標準誤差
ペア 1	訓練前 = 1	100	200	.500	.050
	訓練後 = 1	70	200	.350	.057

対応のあるサンプルの比率の信頼区間

			差の 95% 信頼区間		
	区間のタイプ	比率の差	漸近標準誤差	下限	上限
ペア 1: 訓練前 - 訓練後	Bonett-Price	.150	.046	.058	.239
	Newcombe	.150	.046	.058	.238
	Wald	.150	.046	.059	.241

対応のあるサンプルの比率検定

			有意			
	検定の種類	比率の差	漸近標準誤差	Z	片側 p 値	両側 p 値
ペア 1: 訓練前 - 訓練後	McNemar	.150	.046	3.162	<.001	.002
	McNemar (連続修正)	.150	.046	3.057	.001	.002

		後		合計
前	あり	40	60	100
	なし	30	70	100
合計		70	13	200

$$\frac{100}{200} - \frac{70}{200} = \frac{40+60}{200} - \frac{40+30}{200} = \frac{60}{200} - \frac{30}{200}$$

この信頼区間は不一致の部分の割合の差を表す。

SPSSでは、

- Agresti-Min
- Bonett-Price
- Newcombe
- Wald
- Wald (連続修正)

が選べる。教科書などでは、WaldあるいはWald(連続修正)が書かれている。一方で、Newcombeを推奨する文献がある。

SPSSでは、

- 正確2項
- McNemar
- Wald
- Mid-p調整済み2項
- McNemar (連続修正)
- Wald (連続修正)

が選べる。つまり、McNemar検定以外の方法での検定も可能である。ちなみに、先ほどは、正確2項(正確なp値)を用いて計算している。

「違いを見る」ことと「一貫性を見ることの違い」

「一貫性」を見るということ

検査Aと検査Bがあったとする。検査Aは精度の高い検査方法であるが費用がかかり、検査Bは新しい検査方法で簡便に行えるとする。この2つの検査を同じ被験者に行った場合、次のような対応のあるクロス集計表が構成される。

		検査B	
		陽性	陰性
検査A	陽性	(a)	(b)
	陰性	(c)	(d)

この例の場合には、(a) 検査Aおよび検査Bともに陽性、(d) 検査Aおよび検査Bともに陰性の度数(被験者数)が診断結果が検査A、検査Bで診断結果が異なる(b)(c)よりも大きくなることが期待される

評価の方法 = カッパ係数

「違い」を見るということ

手術による不安感に関する調査を実施したとする。この調査では、医師からの手術に関する説明前に、手術に対する不安感(あり、なし)を質問したうえで、説明後に同じ質問を行った。この場合には、次のような対応のあるクロス集計表が作成される。

		説明後	
		不安あり	不安なし
説明前	不安あり	(a)	(b)
	不安なし	(c)	(d)

この例の場合には、説明前に不安ありだった患者が説明後に不安なしに変化することが期待される。そのため、(b)のほうが(c)よりも高くなることが評価される。

評価の方法 = McNemar検定

Kappa係数と重み付けKappa係数

病理診断において，二人の病理医のグレーディングの結果の一致度を評価することを検討している。

Kappa係数

		観察者2			
		G1	G2	G3	G4
観察者1	G1	20	6	3	2
	G2	5	20	2	3
	G3	2	6	10	2
	G4	4	2	3	10

カテゴリ間が一致しているか否かを評価する。例えば，

- 観察者1がGrade1(G1)，観察者2がGrade2(G2)の不一致
- 観察者1がGrade1(G1)，観察者2がGrade4(G4)の不一致

が同じ重み(不一致)として評価される。

SPSSの実行方法

1. 「分析」→「記述統計」→「クロス集計表」
2. 「クロス集計表」のダイアログから「統計量」を選択する。
3. 「統計量」ダイアログから「カッパ」を選択する。

重み付きKappa係数

		観察者2			
		G1	G2	G3	G4
観察者1	G1	20	6	3	2
	G2	5	20	2	3
	G3	2	6	10	2
	G4	4	2	3	10

カテゴリが順序変数の場合には，不一致であっても，不一致の程度が異なる。重み付きKappa係数では，

- 不一致小：観察者1がG1，観察者2がG2の不一致
- 不一致中：観察者1がG1，観察者2がG3の不一致
- 不一致大：観察者1がG1，観察者2がG4の不一致

となるように重みづけを行ったうえでKappa係数を計算する。

SPSSの実行方法

1. 「分析」→「尺度」→「重み付きカッパ」

重み付きKappa係数を計算するには，「尺度」を「順序」に設定しなければならない。

Kappa係数の実際例

転倒アセスメントスコアシートの改訂と看護師の評定者間一致性の検討

森田他：日本看護管理学会誌, 14(1),51-58, 2010

この研究は、転倒アセスメントスコアシートの改訂に伴う信頼性を評価するために、2名の看護師に改訂後のスコアシートを用いて49名の入院患者を評価してもらい、その一致性を検討している。

表 3 「改訂後シート」

分類	項目	スコア	評価
年齢	☐ 70歳以上	1	
転倒経験	☐ 転倒・転落したことがある	1	
活動領域	☐ 足腰の弱り、筋力の低下がある	2	
	☐ 車椅子・杖・歩行器を使用している		
	☐ ふらつきがある(バランスを崩しやすい)		
認識力	☐ 不隠行動がある	3	
	☐ 自立心が強い		
	☐ 理解力・記憶力の低下がある		
	☐ 何でもできると自分を過大評価する		
排泄	☐ 排泄時見守りが必要	2	
	☐ 排泄介助が必要		
	☐ 夜間トイレに行く		
薬剤使用	☐ 麻薬	5	
	☐ 抗うつ剤	4	
	☐ 浣腸緩下剤	3	
	☐ 睡眠安定剤	1	
	☐ 降圧利尿剤	1	
環境	☐ 転科・転棟・転室をした	4	
	☐ 点滴・酸素吸入をしている	2	
【危険度と評価スコアの合計】		合計	
危険度Ⅰ(0～4点)転倒を起こす可能性がある		危険度	
危険度Ⅱ(5～15点)転倒を起こしやすい			
危険度Ⅲ(16点以上)転倒をよく起こす		サイン	

表 4 アセスメント項目別評定者間一致性

項目	κ 係数
70歳以上	1.00
転倒経験の有無	1.00
麻薬	1.00
抗うつ剤	1.00
転科・転棟・転室	1.00
点滴・酸素吸入	1.00
自分を過大評価する	1.00
車椅子・杖・歩行器の使用	0.95
排泄介助	0.93
夜間トイレに行く	0.82
理解力の低下	0.81
降圧利尿剤	0.78
足腰の弱り・筋力低下	0.75
睡眠安定剤	0.73
不穏	0.65
自立心が強い	0.65
ふらつき(バランスを崩しやすい)	0.54
浣腸緩下剤	0.48
排泄見守り	0.38

一致性は、項目毎に行われており、例えば、認識力の各項目の場合には0～3のスコアがとられているため、次のような対応のあるクロス集計表

		看護師 2			
		0	1	2	3
看護師 1	0				
	1				
	2				
	3				

を作ったうえでkappa係数を用いて評価されている。

ただし、スコアの評価なので、本来は、**重み付きKappa係数のほうが適切であったと**考えられる。また、**信頼区間 and/or p値があったほうが良い。**

SPSSの結果

Kappa係数

対称性による類似度

		値	漸近標準誤差 ^a	近似 t 値 ^b	近似有意確率
一致の測定方法	カッパ	.455	.066	7.710	<.001
有効なケースの数		100			

a. 帰無仮説を仮定しません。

b. 帰無仮説を仮定して漸近標準誤差を使用

不一致に対する重みを考慮したことにより、
一貫性に関する結果に差異が認められる。

重み付きKappa係数

Cohen の重み付きカッパ

評価 ^a	重み付きカッパ ^b	標準誤差 ^c	漸近 Z ^d	有意確率	95% 漸近信頼区間	
Obs1 - Obs2	.479	.074	6.726	<.001	下限	上限
					.335	.623

$0.8 \leq \kappa$	非常に強い関連	連関関連がある
$0.6 \leq \kappa < 0.8$	やや強い関連	
$0.4 \leq \kappa < 0.6$	やや弱い関連	
$\kappa < 0.4$	非常に弱い関連	連関関連がない

Kappa係数は0.479であることから、やや弱い連関関係である。

(おまけ) 独立なデータの場合の関連性を見るには？

看護師の勤務年数	病院のタイプ				合計
	1	2	3	4	
10年以上	56	32	20	14	122
5年～9年	13	47	45	34	99
5年未満	27	29	33	45	134
合計	96	108	98	93	355

カイ2乗検定のp値は0.001未満であり、有意だった。すなわち、病院のタイプと看護師の勤務年数に関連性が認められた。



関連性の強さの程度を表す測度に**Cramer係数**がある(ファイ係数というのもあるが、2×2クロス集計表の場合には、ファイ係数は上限がおよび下限が変わるため、ここではCramer係数Vを用いる)。

$0.8 \leq V$	非常に強い関連	関連がある
$0.5 \leq V < 0.8$	やや強い関連	
$0.25 \leq V < 0.5$	やや弱い関連	
$V < 0.25$	非常に弱い関連	関連がない

SPSSの実行方法

1. 「分析」→「記述統計」→「クロス集計表」
2. 「クロス集計表」のダイアログから「統計量」を選択する。
3. 「統計量」ダイアログから「PhiおよびCramer V」を選択する。

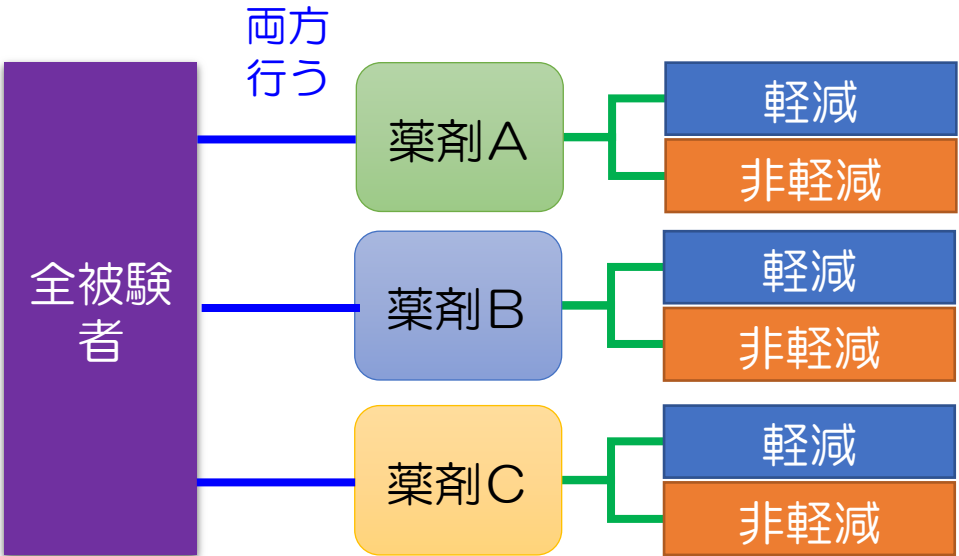
このデータにおけるCramer係数は、0.293であった。したがって、やや弱い関連が認められた。

3群以上の比較の場合にはどうなるか？：CochranのQ検定



Clinical Question

神経障害性疼痛患者に対して、3種類の除痛薬 (A,B,C)を投与したときの痛みの軽減の有無を調べるためのチェンジオーバー試験を実施した。



		B	
		あり	なし
A	あり	5	1
	なし	4	2

		C	
		あり	なし
A	あり	2	4
	なし	1	5

		C	
		あり	なし
B	あり	3	6
	なし	0	3

統計的方法

検定：CochranのQ検定, McNemar検定の多重比較
要約：対応のあるクロス集計表
全体パーセント

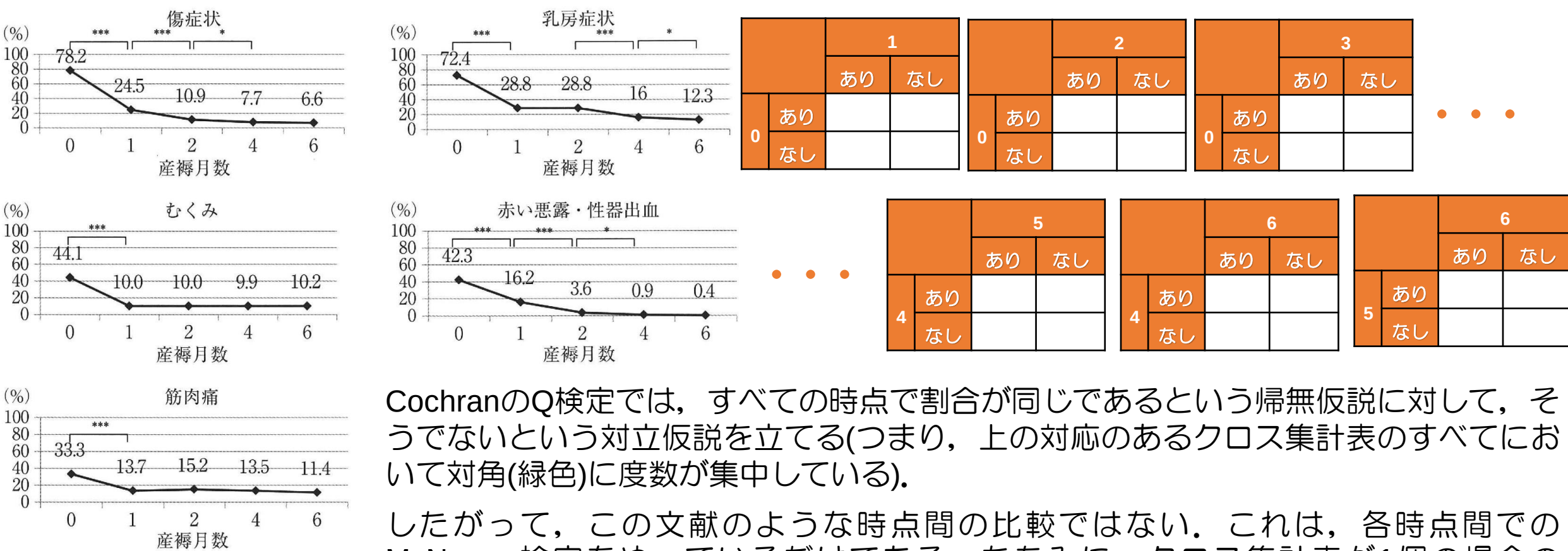
CochranのQ検定の誤った利用

じょく婦が有する身体症状の産後6か月間の推移に関するコホート研究

磐田他：母性衛星, 58, 567-574, 2008

母子ともに異常がない2,709人のじょく婦を対象に産後入院中、1、2、4、6カ月時の5時点において、質問紙調査票を用いて37種類の身体症状の有無を調査した研究である。

分析方法: (前略) 時間の推移による割合の変化に関してはコクランのQ検定を実施した。



CochranのQ検定では、すべての時点で割合が同じであるという帰無仮説に対して、そうでないという対立仮説を立てる(つまり、上の対応のあるクロス集計表のすべてにおいて対角(緑色)に度数が集中している)。

したがって、この文献のような時点間の比較ではない。これは、各時点間でのMcNemar検定をやっているだけである。ちなみに、クロス集計表が1個の場合のCochranのQ検定とMcNemar検定は同じものである。

SPSSによるCochranのQ検定の実行

ここでは、ある疾患患者 17 名に 3 種類の薬剤(Treat：新薬，Control：既存薬，Placebo：プラセボ)のチェンジオーバー試験(すなわち、それぞれの被験者は、ウォッシュアウト期間を通じて、3 種類の薬剤の全てが投与・評価されている)を実施したときの結果(仮想例)である。このとき、アウトカムは、2 値 (0：無効，1：有効)がとられている。

SPSSの実行方法

- 1. 「分析」→「ノンパラメトリック検定」→「過去のダイアログ」→「K個の対応サンプルの検定」
- 2. 「複数の対応サンプルの検定」のダイアログにおいて「CochranのQ」にチェックを入れる。

 Treat	 Control	 Placebo
1	1	0
1	1	0
1	0	1
0	0	0
1	0	0
1	0	0
1	1	0
1	1	0
1	0	0
1	0	1
0	1	0
1	0	0
1	1	0
0	0	0
1	1	1
1	0	0
1	0	0

Cochran の検定

	度数	
	0	1
Treat	3	14
Control	10	7
Placebo	14	3

検定統計量

度数	17
Cochran の Q	13.286 ^a
自由度	2
漸近有意確率	.001

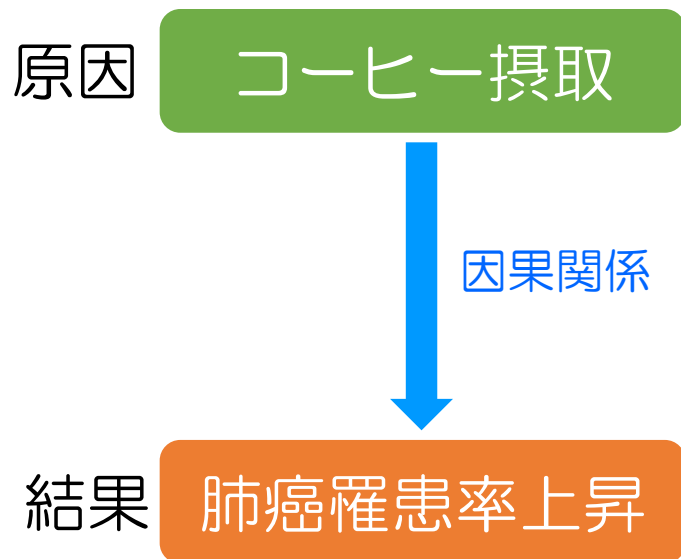
a. 1 は成功したものとして処理されます。

なお、SPSS Exact Testsを導入している場合には、正確なp値を計算することができる。

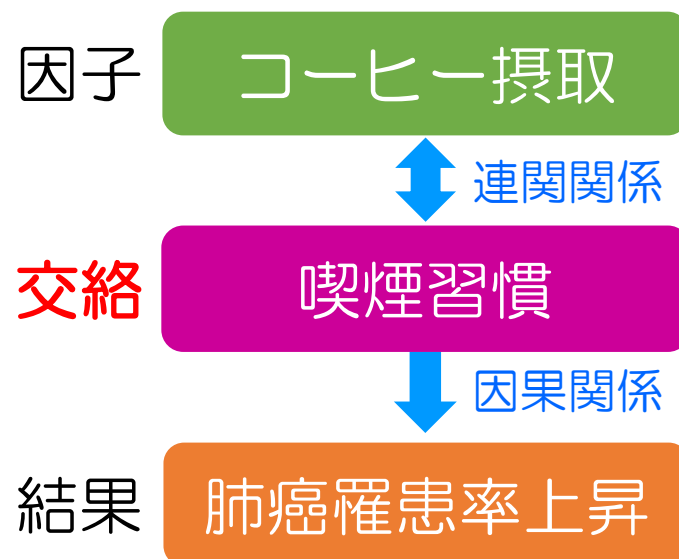
交絡とは

コーヒー摂取が肺癌の罹患率を上昇させるという研究結果が得られたとする。

研究において想定されたこと



実際に考えること



以下の関係がある場合に要因Xは交絡因子と呼ばれる。

1. 要因Xは疾患の既知のリスク要因である (例：喫煙習慣は肺癌のリスク因子である)。
2. 要因Xは因子と関連があるが、因子の結果ではない (例：喫煙はコーヒー摂取と連関関係はあるが、コーヒー摂取の結果ではない)。

交絡因子の調整を行った場合の検定：Cochran-Mantel-Haenszel検定

消化器癌(上部消化器, 下部消化器)に対する外科治療に対する縫合術(真皮縫合術, スキンステープラー)の比較の結果を以下に示す. 評価項目は, 創合併症の発現である(Tsujinaka et al., Lancet, 2013).

部位	縫合の方法	創合併症		計
		あり	なし	
上部	真皮縫合術	29 (0.076)	353 (0.924)	382
	スキンステープラー	39 (0.094)	374 (0.906)	413
	計	68 (0.086)	727 (0.914)	795
下部	真皮縫合術	18 (0.102)	158 (0.898)	176
	スキンステープラー	20 (0.198)	81 (0.802)	101
	計	38 (0.137)	239 (0.863)	277

■ 上部消化器がんにおけるスキンステープラーに対する真皮縫合術のオッズ比

$$OR = \frac{29 \times 374}{353 \times 39} = 0.788$$

■ 下部消化器がんにおけるスキンステープラーに対する真皮縫合術のオッズ比

$$OR = \frac{18 \times 81}{158 \times 20} = 0.461$$

部位別で縫合術の比較をした結果, いずれの部位でも真皮縫合術のほうが創合併症の発現は低く, 下部消化器がんでは半分未満になる.

共変量

部位	縫合の方法	創合併症		計
		あり	なし	
上部	真皮縫合術	29 O_{111}	353 O_{112}	382
	スキンステープラー	39 O_{121}	374 O_{122}	413
	計	68	727	795 N_1
下部	真皮縫合術	18 O_{211}	158 O_{212}	176
	スキンステープラー	20 O_{221}	81 O_{222}	101
	計	38	239	277 N_2

Mantel-Haenszel推定値

$$OR_{\text{adj}} = \frac{\frac{O_{111} \times O_{112}}{N_1} + \frac{O_{211} \times O_{212}}{N_2}}{\frac{O_{112} \times O_{121}}{N_1} + \frac{O_{212} \times O_{222}}{N_2}} = \frac{\frac{29 \times 374}{795} + \frac{18 \times 81}{277}}{\frac{353 \times 39}{795} + \frac{158 \times 20}{277}} = 0.658$$

共通オッズ比とは、交絡因子の層別を行ったうえで、それぞれの層に共通なオッズ比を表す。調整なしのオッズ比OR=0.709なので、手術部位で調整することで、オッズ比が小さくなっている。

Cochran-Mantel-Haenszel検定とは、共通オッズ比(Mantel-Haenszel)が1であるという帰無仮説に対して、1でないという対立仮説を検定する。

Cochran-Mantel-Haenszel検定の実例

亜急性期患者の亜急性期病院から急性期病院への緊急病院間転送タイミングに関する研究

Considine et al. : International Journal of Nursing Studies, 91, 77-85, 2019.

この研究は、亜急性期治療を受けた患者が急性期病院に緊急病院間移送が実施された602件の転送(早期転送：54件，遅期転送：548件)において，急性期病院での院内死亡率を比較したコホート研究である。なお，本研究はオーストラリアにおいて実施されている。ここでは，亜急性期治療による入院後1日以内に転送された患者を早期転送群，1日以上経過した群を遅期転送群としている。

Data Analysis: (前略) Cochran-Mantel-Haenszel test was also used to compare health service outcomes following time of emergency transfer until discharge or death between early and late emergency transfers groups. Cochran-Mantel-Haenszel model adjusted odds ratios (AORs) and 95% confidence intervals (95% CI) were reported (交絡因子の言及はなし).

Outcomes	Early transfers		Late transfers		AOR	95% CI	p ^a
	n	%	n	%			
Serious adverse events							
• In-hospital cardiac arrest (院内心停止)	0	0.0	12	2.3	N/A	N/A	N/A
• Unplanned ICU admission (予定外のICU入院)	2	3.7	23	4.3	0.91	0.21 – 4.03	0.907
• Rapid Response Team call (対応チームの緊急呼び出し)	9	16.7	112	21.1	0.79	0.37 – 1.71	0.553
Mortality							
• In-patient mortality for all patients (52 early transfer patients & 497 late transfer patients) (院内死亡率)	2	3.8	80	16.1	0.22	0.05 – 0.96	0.044
• In-patient mortality excluding patients with LOMT (30 early transfer & 193 late transfer patients) (LOMTを除く院内死亡率)	1	3.3	16	8.3	0.42	0.05 – 3.48	0.418

^a CHM = Cochran-Mantel-Haenszel test.

Cochran-Mantel-Haenszel検定

ここでは、頭部外傷データを用いる。このデータは、カナダの 3121 名の軽度頭部外傷患者に対する CT による脳所見(1：有，0：無)の有無に対して、嘔吐の有無(1：有，0：無)が影響を及ぼすか否かを検討している。ここでは、交絡因子として、頭蓋底骨折の有無(1：有，0：無)を用いる。

	嘔吐の有無	頭蓋底骨折	CT脳所見
1	0	0	0
2	0	0	0
3	0	0	0
4	0	0	0
5			
6			
7			
8			
9			
10	0	0	0

Cochran-Mantel-Haenszel検定では、要因およびアウトカムは2値変数でなければならない。

・ ・ ・
・ ・ ・
・ ・ ・

SPSSの実行方法

- 「分析」→「記述統計」→「クロス集計表」
- 「クロス集計表」のダイアログから「行」に「嘔吐の有無」(要因)，「列」に「CT脳所見」，「層」に「頭蓋底骨折」(交絡因子)ドラッグ&ドロップする。
- 「統計量」をクリックし，ダイアログから「CochranとMantel-Haenszelの統計量」にチェックを入れる。

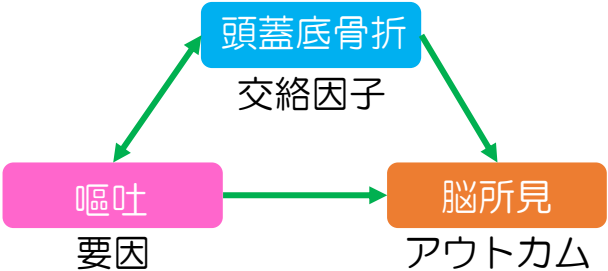
オッズ比の等質性の検定			
	カイ 2 乗	自由度	漸近有意確率 (両側)
Breslow-Day	.008	1	.929
Tarone	.008	1	.930

条件付独立の検定			
	カイ 2 乗	自由度	漸近有意確率 (両側)
Cochran	79.106	1	<.001
Mantel-Haenszel	77.034	1	<.001

条件付き独立性仮定のもとでは、Cochranの統計量は、層の数が固定されている場合にのみ自由度1のカイ 2 乗分布として漸近分布しますが、Mantel-Haenszelの統計量は、常に自由度1のカイ 2 乗分布として漸近分布します。観測値と期待値との差の総和が0である場合に連続性修正がMantel-Haenszelの統計量から取り除かれます。

Mantel-Haenszel の共通オッズ比の推定値				
推定値		4.033		
ln(推定値)		1.395		
ln(推定値)の標準誤差		.167		
漸近有意確率 (両側)		<.001		
漸近 95% 信頼区間	共通オッズ比	下限	2.910	
		上限	5.591	
	ln(共通オッズ比)	下限	1.068	
		上限	1.721	

Mantel-Haenszel の共通オッズ比の推定値は、1.000 の仮定の共通オッズ比のもとで漸近的に正規分布します。したがって、推定値の自然対数です。



Cochran-Mantel-Haenszel検定では、層で分けたときのクロス集計表でのオッズ比，つまり，
頭蓋底骨折有のサブグループにおける嘔吐の有無による脳所見有無のオッズ比
頭蓋底骨折無のサブグループにおける嘔吐の有無による脳所見有無のオッズ比
が同じであることを仮定している。オッズ比の等質性の検定において有意であれば，等質性の仮定を満たさないと判断する。

Mantel-Haenszel検定のほうを見るのが一般的

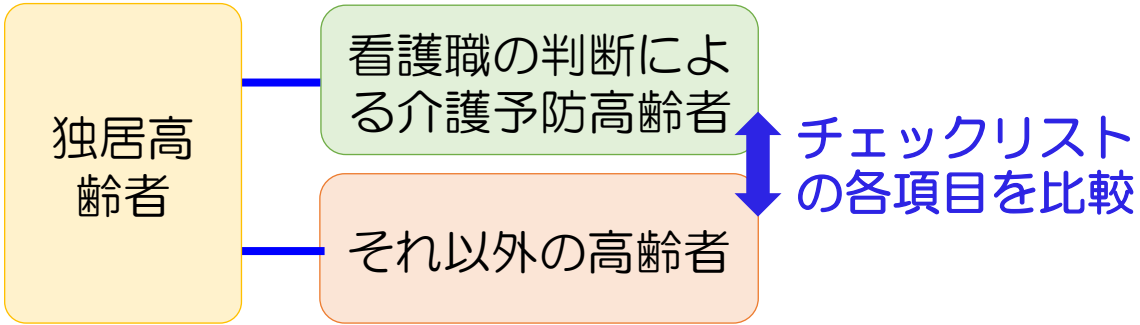
共通オッズ比の推定値である。つまり，ここでの共通オッズ比4.033 [2.910, 5.591]は，頭蓋底骨折の有無について調整したときのオッズ比を表している。

Cochran-Mantel-Haenszel検定の勘違い：Cochran-Armitageとの違い

独居虚弱高齢者における介護予防対象者把握の検討

河野他：日本公衆衛生学会誌，2，83-92，2008

ある地域で介護認定を受けていない独居高齢者を対象とした研究において，介護予防対象者を把握するための基本チェックリストと地域看護職の判断との内容についての検討を行っている。



3. 分析方法

各群の比較検討には，カテゴリー変数は χ^2 検定，順序性のあるカテゴリー変数は Mantel-Haenszel の χ^2 検定，連続変数は t-検定を使用した。

健康状態を順序関係のあるカテゴリとみなせなくもない。つまり，この検定は，Cochran-Armitage検定であり，Cochran-Mantel-Haenszel検定ではない。

				看護職の判断による 介護予防事業対象者 n = 33	それ以外の高齢者 n = 46	P 値
健康状態	順序カテゴリ カルデータ	とても健康である	n (%)	3 (9.1)	9 (19.6)	.0010
		まあ健康である	n (%)	14 (42.4)	32 (69.6)	
		あまり健康でない	n (%)	16 (48.5)	5 (10.9)	
		全く健康ではない	n (%)	0 (0)	0 (0)	



Thank you for your kind attention

shimokaw@wakayama-med.ac.jp



toshibow2000@gmail.com