

クロス集計表と諸種の検定

臨床研究センター
谷岡健資

2019年度 医学統計セミナー

開催日	講義内容
5月21日 (火)	記述統計・グラフ表示・統計ソフトウェア
6月18日 (火)	単群・2群比較のための諸種の検定
7月25日 (木)	分散分析と多重比較
8月22日 (木)	相関分析と回帰分析
9月26日 (木)	クロス集計表の解釈と諸種の検定
10月17日 (木)	ロジスティック回帰分析
11月14日 (木)	生存曲線の推定と比較
12月 3日 (火)	Cox比例ハザードモデル
1月23日 (木)	感度・特異度・ROC曲線とその比較
2月27日 (木)	繰り返し測定分散分析、混合効果モデル

目次

- はじめに
- 一標本の母比率に関する検定
- カイ自乗検定
- Fisherの正確検定
- 因果関係と連関関係
- オッズ比およびその信頼区間
- 対応のあるクロス集計表について
- K×Iクロス集計表におけるカイ2乗値

1.はじめに

□ 質的データについて (東京大学教養学部統計学教室, 1994)

名義尺度

観測値が単に対象のカテゴリを示しているデータ

順序尺度

観測値が序数としての意味をもち、対象間の順序付けを示しているデータ

間隔尺度

観測値が任意の単位の何倍という形で示され、測定値の差が意味をもつデータ

比例尺度

観測値が任意の単位の何倍という形で示され、さらに原点0が絶対的な意味をもっていて、差や比率も意味を持つデータ

質的データ

量的データ

1.はじめに

□質的データの例（一標本）

介入研究の例

ある癌の臨床試験で，単群の試験治療の奏効率は下記の通りであった．

$$\frac{\text{試験治療によって奏功した被験者の人数}}{\text{試験治療を受けた被験者の人数}} = 0.60$$

観察研究の例

冠動脈疾患(CHD)患者のうち，喫煙者の割合は下記の通り

$$\frac{\text{冠動脈疾患でかつ喫煙者の人数}}{\text{冠動脈疾患患者の人数}} = 0.56$$

1.はじめに

□質的データにおける気を付けないといけない点

割合は単純な指標であるが実は気を付けないといけない点が多々ある

発生率 (incidence rate)

ある一定期間内に**新たに**その疾患にかかった人数

その期間内に存在したリスク保有者の人数

存在率 (prevalence)

特定時点に集団中に存在する対象疾患罹患者の数

その時点でその集団に属する人の総数

⇒ 存在率はリスクの指標にはならないが、発生率はなりうる

同じ割合でも母集団や時点，期間等に留意する必要がある

1.はじめに

□クロス集計表の作り方

2群の比較を行う際、
質的2変量データから
クロス集計表を作成する

表1：質的データ（2変量）

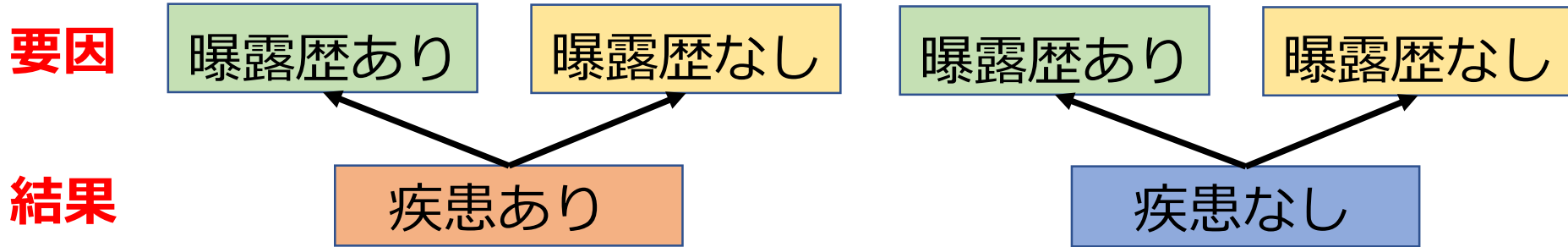
ID	喫煙の有無	ケース
1	喫煙者	冠動脈疾患
2	喫煙者	コントロール
3	非喫煙者	コントロール
4	喫煙者	冠動脈疾患
5	非喫煙者	コントロール
...	...	...
N	非喫煙者	コントロール

表2：クロス集計表(Gordis, 2010)

	冠動脈疾患患者	コントロール	合計
喫煙者	112 (0.56)	176 (0.44)	288
非喫煙者	88 (0.44)	224 (0.56)	312
合計	200 (1.00)	400 (1.00)	600

1.はじめに

□クロス集計表で気を付けないといけない点 ケースコントロール研究



「疾患あり」(ケース)のなかで**曝露歴あり**と**曝露歴なし**を集める
「疾患なし」(コントロール)のなかで**曝露歴あり**と**曝露歴なし**を集める

表2：クロス集計表(Gordis, 2010)

	冠動脈疾患患者	コントロール	合計
喫煙者	112	176	288
非喫煙者	88	224	312
合計	200	400	600



質問：冠動脈疾患の存在率は0.33である。正しいか？

1.はじめに

□クロス集計表で気を付けないといけない点 ケースコントロール研究

表2：クロス集計表(Gordis, 2010)

	冠動脈疾患患者	コントロール	合計
喫煙者	112	176	288
非喫煙者	88	224	312
合計	200	400	600



質問：冠動脈疾患の存在率は0.33である。正しいか？

⇒ 正しくない

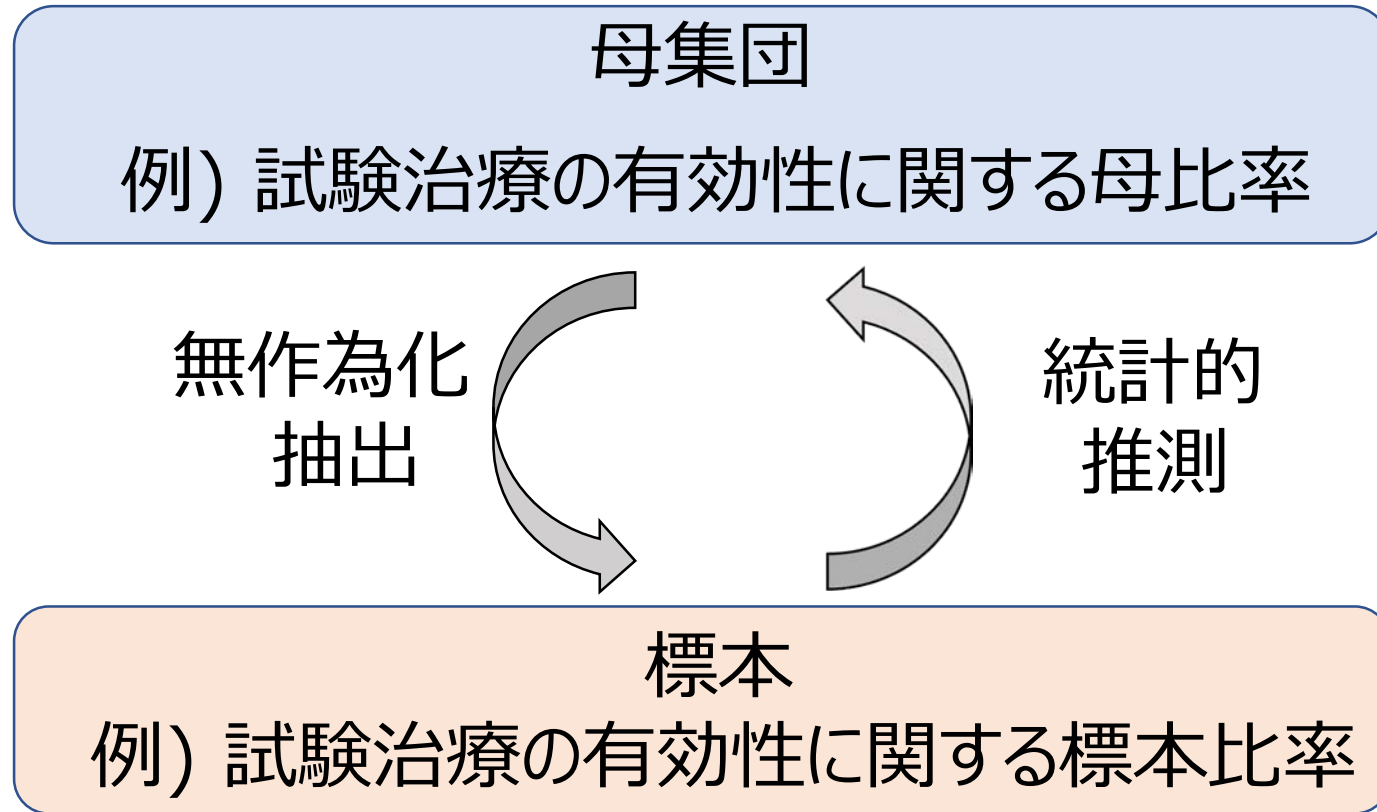
ケースコントロール研究において、「ケース：コントロール」の比を研究者が選択できるから

「1:2」や「1:4」等

ひとえに割合といっても気を付けないといけない点が多々存在する。

2. 一標本の母比率の検定

□ 母集団と標本について



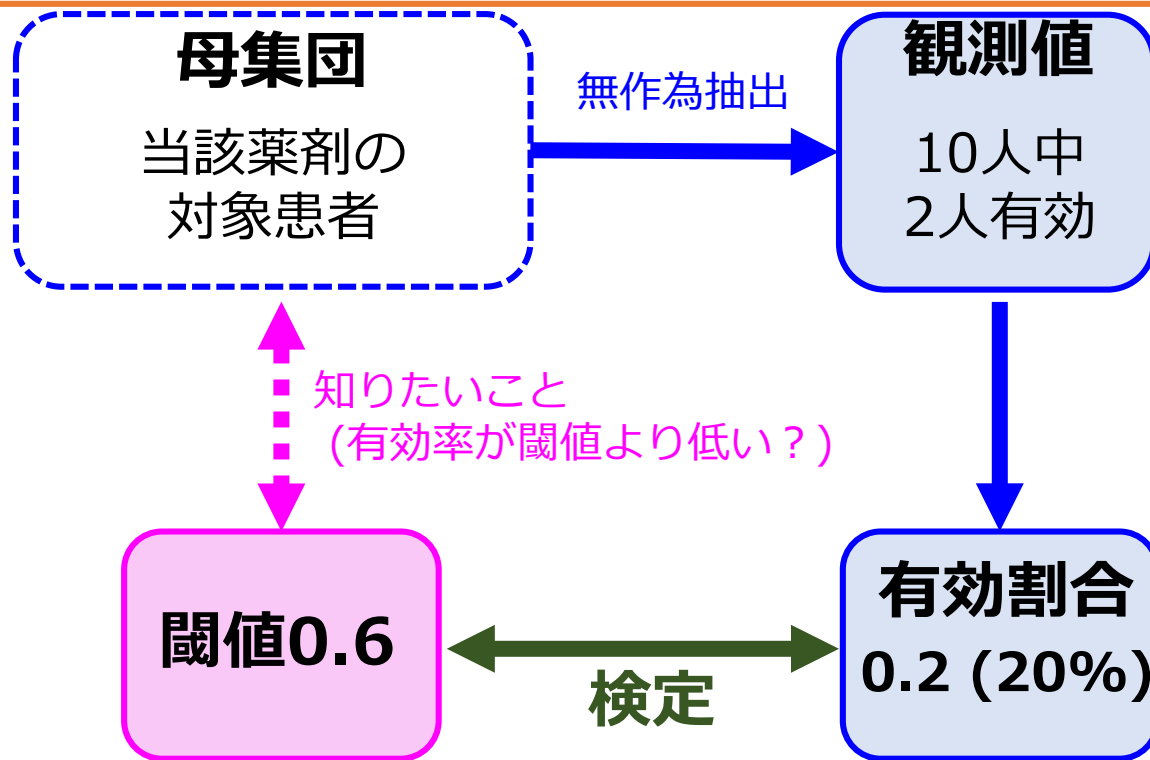
標本比率から母比率を推定する(区間推定・点推定)
母比率がある比率より小さいか (大きい) 等を検定する

2. 一標本の母比率の検定

あるクリニックで有効率が0.6といわれている薬剤を10人の患者に投薬したところ、2人しか有効でなかった。

⇒ 薬剤の有効率が疑問

有効率が0.6より低いかどうかを確認せよ (柳川・荒木, 2010).



2. 一標本の母比率の検定

帰無仮説 : 薬剤の母集団での有効率 = 0.6

対立仮説 : 薬剤の母集団での有効率 < 0.6

検定の概要

Step 1: **帰無仮説**と**対立仮説**を設定する

Step 2: **帰無仮説**が成り立つと仮定する

Step 3: **帰無仮説**のもとで**標本（データ）比率**が得られる可能性を検討する

Step 4: 帰無仮説のもとで標本比率が得られる可能性が低い
⇒そもそも帰無仮説のもとで得られた標本比率でないと考える
帰無仮説を棄却（**対立仮説**を主張する）

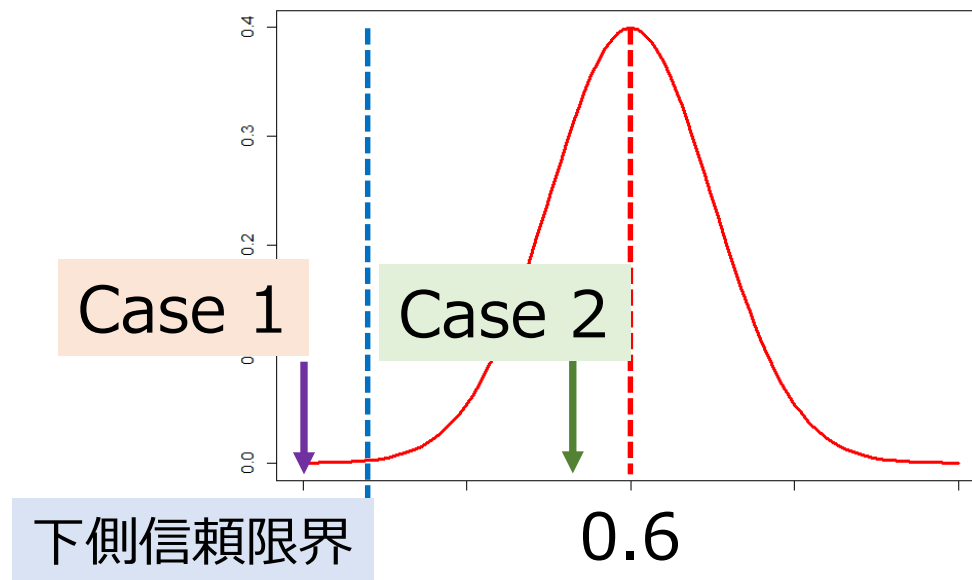
帰無仮説のもとで標本比率が得られる可能性を否定できない
帰無仮説を否定する根拠がない（**帰無仮説**を受容する）

2. 一標本の母比率の検定

帰無仮説：薬剤の母集団での有効率 = 0.6

対立仮説：薬剤の母集団での有効率 < 0.6

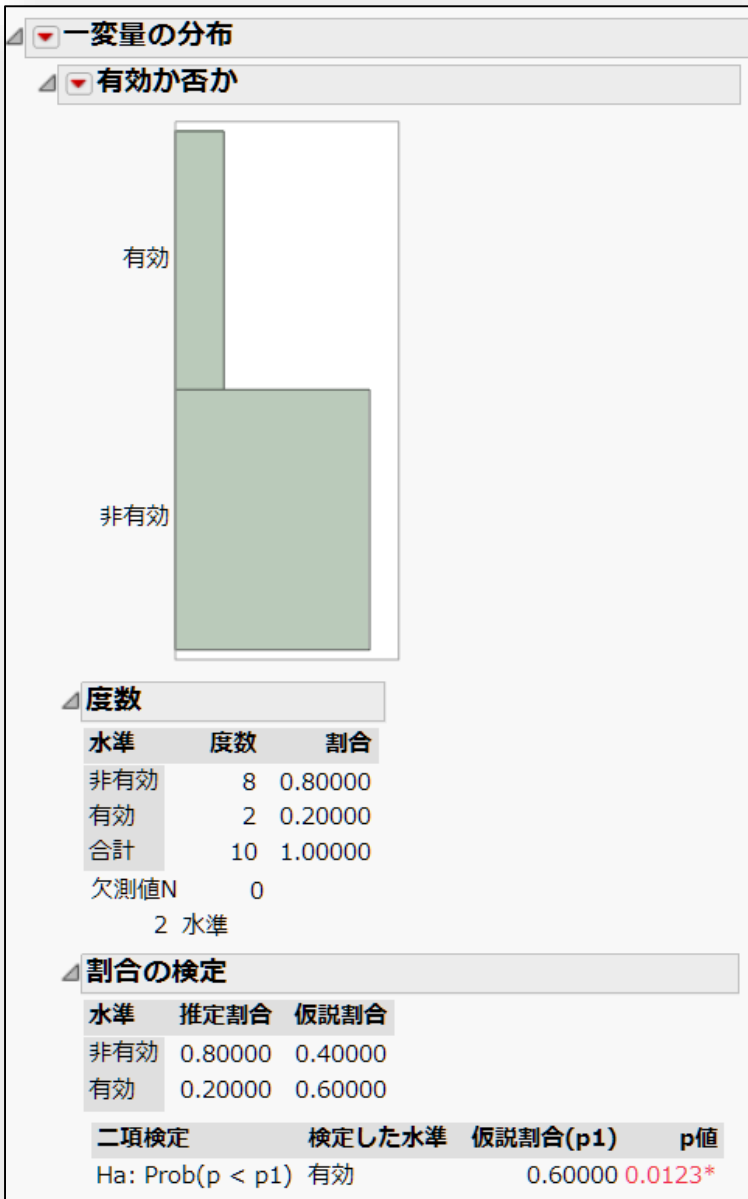
帰無仮説：薬剤の母集団での有効率 = 0.6



Case 1: **帰無仮説**のもとで当該母比率が得られる可能性低い
⇒ **帰無仮説**を棄却 (**対立仮説**を主張する)

Case 2: **帰無仮説**のもとで当該母比率が得られる可能性を否定できない
⇒ **帰無仮説**を受容

2. 一標本の母比率の検定



- 1 : 「分析」→「1変量の分布」で変数を選択
- 2 : 変数(ここでは「有効か否か」)の▼から「割合の検定」を選択
- 3 : 「仮説割合」の右側の白い箱の中に0.60を入力し, 「割合が仮説値より小さい」にチェックを入れる.

P値が0.0123

⇒帰無仮説が棄却され対立仮説が支持される

⇒当該クリニックでの当該薬剤の有効率が0.6よりも低いことが示された.

2. 一標本の母比率の検定

母比率の点推定値

母比率の点推定値は通常の比率を計算可能

例) 薬剤のさきほどの例

$$2 / 10 = 0.20$$

母比率の区間推定値

母比率の95%信頼区間の意味は下記の通り

例) 95%の確率で当該区間に母比率が入る

例) 薬剤のさきほどの例

$$[0.035, 0.558]$$

⇒母比率の95%信頼区間を狭く推定しようとする
量的なものと比較して（分散に勿論依存するが）
広くなる傾向にある

2. 一標本の母比率の検定

Kurokawa et al., *BJC*, (2014)

対象：HER2陽性の進行・再発胃癌患者

ヒストリカルコントロール：フッ化ピリミジン系抗がん剤(S1)+CDDP

奏効率 = 35%

試験治療：S1+CDDP+Tmabの併用療法

奏効率 = 68% (36例 / 53例)

母集団

HER2陽性進行・
再発胃癌患者

無作為抽出

観測値

53人中36人
奏効

p 値 $<0.001 \Rightarrow$ 有意
95%CI $\Rightarrow [0.535, 0.797]$

知りたいこと
(奏効率が閾値より高い?)

閾値0.35

Historical
Control

奏効割合

0.68 (68%)

検定

本治療レジメンは、
HER2陽性の進行・再発
胃癌患者に対して、
 $\Rightarrow$ 非常に**有望**であると結論

3. カイ自乗検定

□ 母集団と標本について

2標本の比率の差の検定について扱う

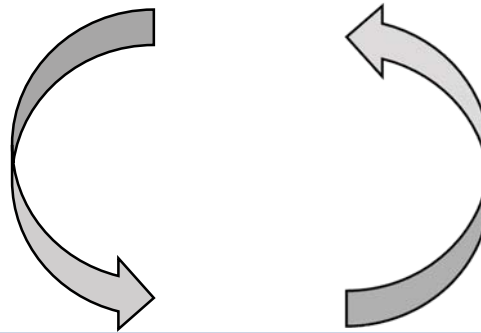
母集団

例)

試験治療の有効性に関する母比率 — 既存治療の有効性に関する母比率

無作為化
抽出

統計的
推測



標本

例)

試験治療の有効性に関する標本比率 — 既存治療の有効性に関する標本比率

3.カイ自乗検定

□カイ自乗検定を適用する具体例

(Tsujinaka et al., Lancet, 2013).

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: 創合併症の発現

	創合併症あり	創合併症なし	計
真皮縫合術	47 (0.084)	511 (0.916)	558
スキンステープラー	59 (0.115)	455 (0.855)	514
計	106 (0.099)	966 (0.901)	1072

縫合術の違いによって、創合併症に違いがあるだろうか？

3. カイ自乗検定

□ カイ自乗検定を適用する具体例

帰無仮説：「外科治療に対する縫合術」と「創合併症」は独立である

対立仮説：「外科治療に対する縫合術」と「創合併症」は独立でない

「外科治療に対する縫合術」と「創合併症」は独立とはどんな状況？

元の表

	創合併症あり	創合併症なし	計
真皮縫合術	47	511	558
スキンステープラー	59	455	514
計	106	966	1072

帰無仮説のもとでの表

	創合併症あり	創合併症なし	計
真皮縫合術			558
スキンステープラー			514
計	106	966	1072

3.カイ自乗検定

□カイ自乗検定を適用する具体例

帰無仮説：「外科治療に対する縫合術」と「創合併症」は独立である

対立仮説：「外科治療に対する縫合術」と「創合併症」は独立でない

「外科治療に対する縫合術」と「創合併症」は独立とはどんな状況？

帰無仮説のもとでの表

	創合併症あり	創合併症なし	計
真皮縫合術	$\frac{106 \times 558}{1072} = 55.2$	$\frac{966 \times 558}{1072} = 502.8$	558
スキンステープラー	$\frac{106 \times 514}{1072} = 50.8$	$\frac{966 \times 514}{1072} = 463.2$	514
計	106	966	1072

例：真皮縫合術のもとで創合併症が生じる確率が独立ということは

「創合併症が生じる」の確率 と 「真皮縫合術を実施する」確率の
はそれぞれ影響を与えない

⇒それが各セル（上記だと4つのセル）で起こっている

3. カイ自乗検定

□ カイ自乗検定を適用する具体例

帰無仮説：「外科治療に対する縫合術」と「創合併症」は独立である

対立仮説：「外科治療に対する縫合術」と「創合併症」は独立でない

「外科治療に対する縫合術」と「創合併症」は**独立でない状況**とは？

帰無仮説のもとでの表

	創合併症あり	創合併症なし	計
真皮縫合術	$\frac{106 \times 558}{1072} = 55.2$	$\frac{966 \times 558}{1072} = 502.8$	558
スキンステープラー	$\frac{106 \times 514}{1072} = 50.8$	$\frac{966 \times 514}{1072} = 463.2$	514
計	106	966	1072



「外科治療に対する縫合術」と「創合併症」は**独立である状況**の表から乖離している表を

「外科治療に対する縫合術」と「創合併症」は**独立でない状況**の表とする

3. カイ自乗検定

□ カイ自乗検定を適用する具体例

カイ2乗検定とは、縦の項目(変数)と横の項目(変数)に関連性がない状況(縫合術によって創合併症の違いがない)でのクロス集計表からの乖離を評価する。

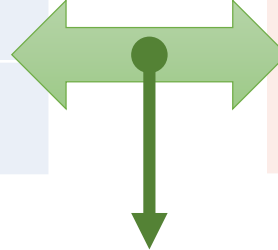
観測データ

	創合併症あり	創合併症なし
真皮縫合術	47	511
スキンステープラー	59	455

期待度数(関連性がない状況)

創合併症あり	創合併症なし
$\frac{106 \times 558}{1072} = 55.2$	$\frac{966 \times 558}{1072} = 502.8$
$\frac{106 \times 514}{1072} = 50.8$	$\frac{966 \times 514}{1072} = 463.2$

乖離



	創合併症あり	創合併症なし
真皮縫合術	$47 - 55.2 = -8.2$ $(-8.2)^2 / 55.2 = 1.218$	$511 - 502.8 = 8.2$ $(8.2)^2 / 502.8 = 0.134$
スキンステープラー	$59 - 50.8 = 8.2$ $(8.2)^2 / 50.8 = 1.324$	$455 - 463.2 = -8.2$ $(-8.2)^2 / 463.2 = 0.145$

合計はカイ2乗値と呼ばれ、カイ2乗値を用いることで p 値を計算することになる。

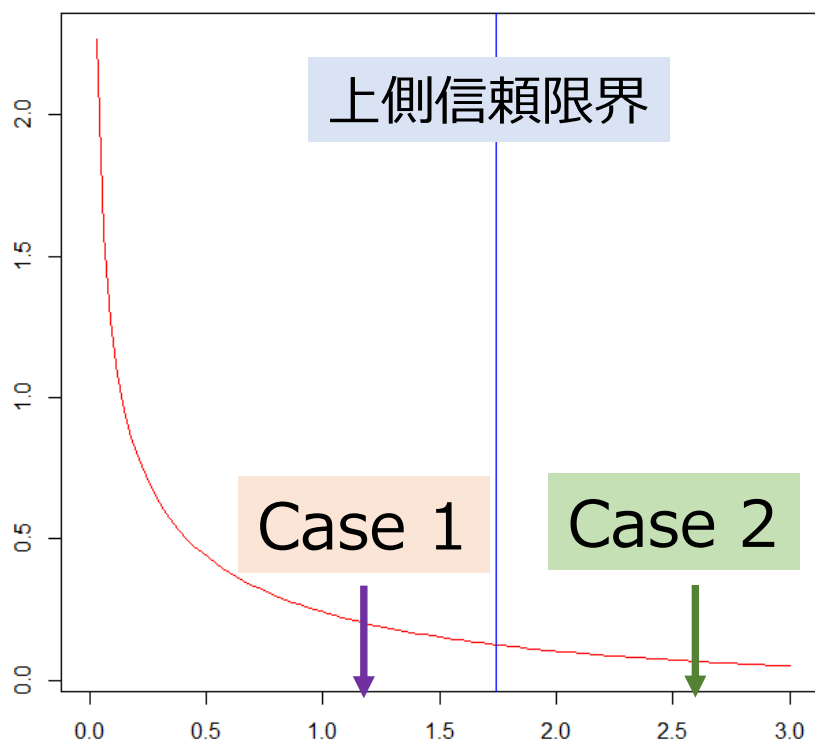
3. カイ自乗検定

□ カイ自乗検定を適用する具体例

帰無仮説：「外科治療に対する縫合術」と「創合併症」は独立である

対立仮説：「外科治療に対する縫合術」と「創合併症」は独立でない

カイ自乗分布（自由度1）



Case1

手元の表が独立な表から得られた可能性を否定できない
⇒ 帰無仮説を受容

Case2

手元の表が独立な表から得られた可能性が低い
⇒ そもそも帰無仮説を仮定していたことが間違っていたのでは（棄却）

横軸：独立な表との乖離度（意識）

3. カイ自乗検定

□ JMPの結果

度数: 度数

分割表

	創合併症		
	あり	なし	合計
度数			
全体%			
列%			
行%			
スキンステーパー	59	455	514
	5.50	42.44	47.95
	55.66	47.10	
	11.48	88.52	
真皮縫合術	47	511	558
	4.38	47.67	52.05
	44.34	52.90	
	8.42	91.58	
合計	106	966	1072
	9.89	90.11	

検定

N	自由度	(-1)*対数尤度	R2乗(U)
1072	1	1.4015636	0.0041

検定	カイ2乗	p値(Prob>ChiSq)
尤度比	2.803	0.0941
Pearson	2.804	0.0941

データセットからクロス集計表作成・検定の場合

1: 「分析」→「2変量の関係」で変数を選択

- ・目的変数: 創合併症
- ・説明変数: 縫合術

クロス集計表から検定の場合

1: 「分析」→「2変量の関係」で変数を選択

- ・目的変数: 創合併症
- ・説明変数: 縫合術
- ・度数: 度数

3/0列		縫合術	創合併症	度数
1	真皮縫合術	あり		47
2	真皮縫合術	なし		511
3	スキンステップ...	あり		59
4	スキンステップ...	なし		455

p値=0.0941であることから、有意水準 $\alpha=0.05$ のもとで有意でなかった。したがって、縫合術によって創合併症に違いがあるとは言えなかった。

3. カイ自乗検定

□ 何故検定で評価するのか？ 比に関する留意点

2重盲検試験で、200例の患者に新薬が、別の200例の患者に偽薬が投与された。前者では10例が感染症に罹患し、後者では20例が罹患した。この新薬は感染症の予防に有効か。
Reigier (1982) [土屋・肥田・後藤 (1988)]

■ 誤った解析

$$\text{新薬での罹患率 } \hat{p}_1 = \frac{5}{200} = 0.05$$

$$\text{疑薬での罹患率 } \hat{p}_2 = \frac{10}{200} = 0.10$$

偽薬を服用した患者の疾病発生は、新薬を服用した患者における疾病発生のおける2倍であった。したがって、この薬は罹患率を半分にする。

■ 正しい解析(カイ2乗検定)

	感染	非感染	合計
新薬	10(A)	190(B)	200(n_1)
偽薬	20(B)	180(D)	200(n_2)
合計	30(m_1)	370(m_2)	400(n)

$$\begin{aligned}\chi_c^2 &= \frac{n \cdot (|A \cdot D - B \cdot C| - n/2)^2}{n_1 \cdot n_2 \cdot m_1 \cdot m_2} \\ &= \frac{400(|10 \cdot 180 - 20 \cdot 190| - 400/2)^2}{200 \cdot 200 \cdot 30 \cdot 370} \\ &= 2.92 \quad \text{p 値 } 0.087\end{aligned}$$

したがって、有意差は認められない。

4.Fisherの正確検定

ある新薬の初期の臨床試験で，新薬19例，対照薬9例に割り当てた，その結果，副作用の有無が次の通りだった．副作用の発生率に差があるだろうか．

	副作用		計
	あり	なし	
新薬	1	18	19
対照薬	4	5	9
計	5	23	28

Fisherの正確検定では，周辺度数(青色の部分)を固定したときに，現在のクロス集計表の状況が帰無仮説のもとでどれぐらいの確率で起こり得るかを評価する．

	副作用		計
	あり	なし	
新薬	O_{11}	$O_{12} = n_1 - O_{11}$	n_1
対照薬	$O_{21} = m_1 - O_{11}$	$O_{22} = N - O_{11} - O_{21} - O_{12}$	n_2
計	m_1	m_2	N

周辺度数が固定されると，一つの度数(O_{11})が決めれば，残りの度数が定まる．

4. Fisherの正確検定

Fisherの正確検定とは

$$O_{11} = 0$$

	副作用		計
	あり	なし	
新薬	0	19	19
対照薬	5	4	9
計	5	23	28

こうなる確率 : $p = 0.0013$

$$O_{11} = 1$$

	副作用		計
	あり	なし	
新薬	1	18	19
対照薬	4	5	9
計	5	23	28

こうなる確率 : $p = 0.0244$

$$O_{11} = 2$$

	副作用		計
	あり	なし	
新薬	2	17	19
対照薬	3	6	9
計	5	23	28

こうなる確率 : $p = 0.01462$

$$O_{11} = 3$$

	副作用		計
	あり	なし	
新薬	3	16	19
対照薬	2	7	9
計	5	23	28

こうなる確率 : $p = 0.3549$

$$O_{11} = 4$$

	副作用		計
	あり	なし	
新薬	4	15	19
対照薬	1	8	9
計	5	23	28

こうなる確率 : $p = 0.3549$

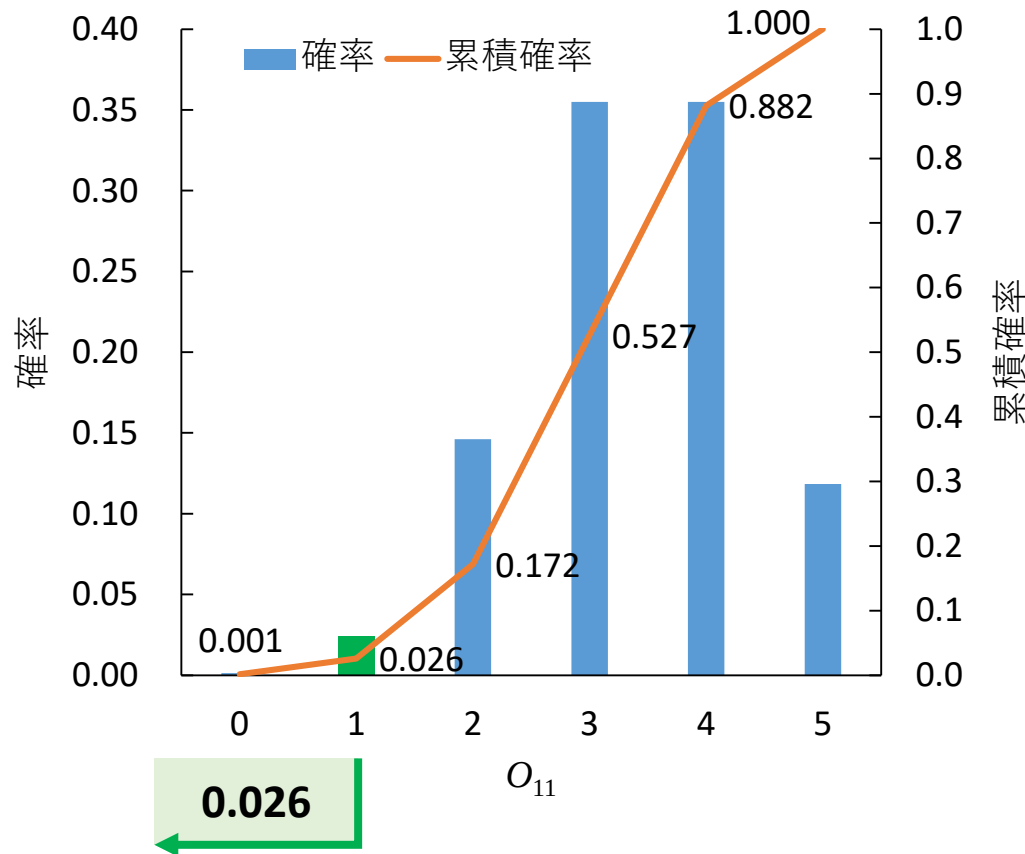
$$O_{11} = 5$$

	副作用		計
	あり	なし	
新薬	5	14	19
対照薬	0	9	9
計	5	23	28

こうなる確率 : $p = 0.1183$

因みに、これらの確率分布は超幾何分布という確率分布を用いて計算できる。

4. Fisherの正確検定



観測度数および観測度数よりも偏った確率の和(累積確率あるいは1-累積確率で0.5をまたがないいずれかの値)を求める。

両側対立仮説の場合は、0.026を2倍した0.052がp値である。

事例のp値は片側検定の場合に0.0256である:H1
両側検定の場合に0.0504である:H0

すなわち、両側検定のもとでは帰無仮説を支持できない(有意でない)

4.Fisherの正確検定

癌腫瘍をもつ20匹のマウスが得られているとする。
新しい試験薬の腫瘍に及ぼす効果を検討したい。

対照群のマウスにはプラセボだけを投与し，他方の群のマウスには新薬を投与する．腫瘍の大きさに変化があったかどうかを調べたところ，次のような結果になった．

Yang and Robinson (1986) [永井・後藤他 (1994)]

	プラセボ	新薬	合計
腫瘍の縮小	1	7	8
腫瘍の不変・拡大	9	3	12
合計	10	10	20

この新薬は腫瘍の大きさを縮小するのに役立つといえるか．

4.Fisherの正確検定

誤った解析

■カイ2乗検定

p 値は0.023である.

有意水準0.05のもとで、新薬は腫瘍の大きさを縮小する効果がある

正しい解析

■Fisherの正確検定

p 値は0.0099である.

有意水準0.01のもとで、新薬は腫瘍の大きさを縮小する効果がある

- 例数が少ないときにカイ2乗検定を用いるのは誤り.
- 一部のセルに属する例数が少ない場合は要注意.
[帰無分布(p値を計算するための確率分布)での近似性能が悪くなる]

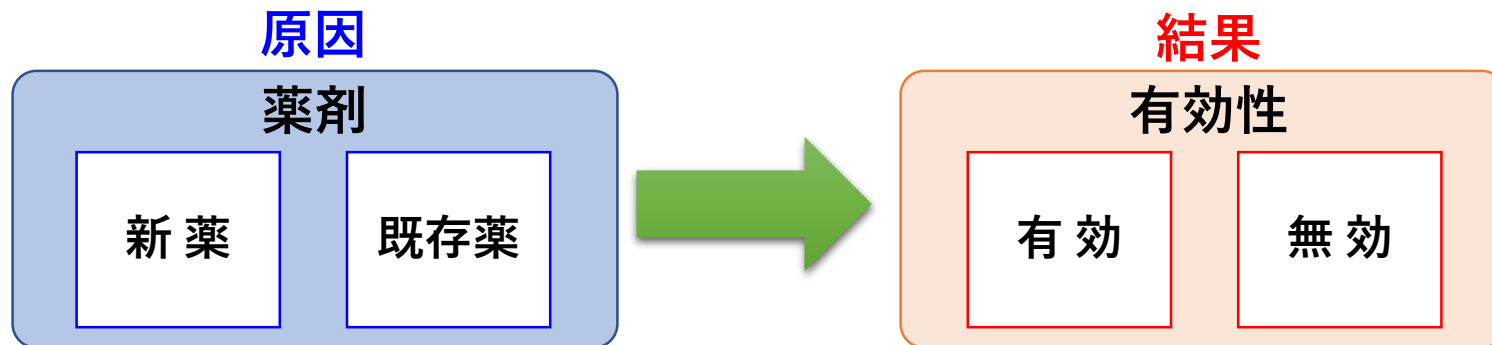


Fisherの正確検定を用いる

5.因果関係と連関関係

因果関係

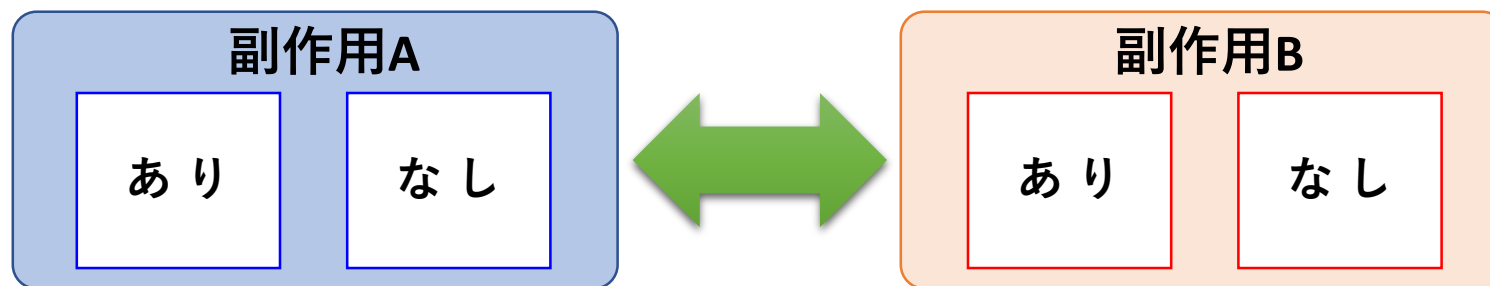
因果関係とは、クロス集計表において2変数の関係が原因(説明変数)と結果(応答変数)で表すことができる状況である。



因果関係では、原因のほうの結果に比べて時間的に先でなければいけない。

連関関係

連関関係とは、クロス集計表において2変数の関係のみを表す関係である。



連関関係には、時間的な概念は存在しない。

6. オッズ比およびその信頼区間

消化器癌(胃癌, 大腸癌)に対する外科治療に対する縫合術(真皮縫合術, スキンステーパー)の比較のデータを用いる.

R	真皮縫合術	創感染症あり		創感染症あり	創感染症なし	陽性率
		創感染症なし				
	スキンステーパー	創感染症あり	真皮縫合術	47 (a)	511 (b)	$p_1=0.084$
		創感染症なし	スキンステーパー	59 (c)	455 (d)	$p_2=0.115$

オッズとは, ある結果が生じる比率とある結果が生じない比率との比である.

$$\text{真皮縫合術} = \frac{0.084}{1-0.084} = 0.092, \quad \text{ステーパー} = \frac{0.115}{1-0.115} = 0.130$$

オッズ比は, 結果(所見の有無)に対してどれぐらい要因(検査法)が寄与しているかを表す.

$$\text{オッズ比} = \frac{p_1 / (1 - p_1)}{p_2 / (1 - p_2)} = \frac{a \times d}{b \times c} = \frac{47 \times 455}{59 \times 511} = 0.709$$

- 真皮縫合術はスキンステーパーに比べて**0.709倍の創感染症の上昇**が認められる.
- スキンステーパーは真皮縫合術に比べて **$1/0.709=1.410$ 倍の創感染症の上昇**が認められる.

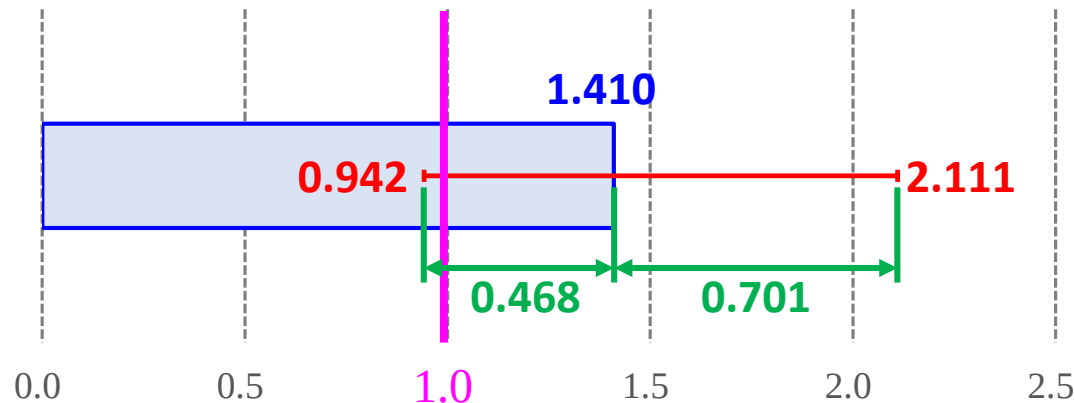
6. オッズ比およびその信頼区間

オッズ比		
オッズ比	下側95%	上側95%
1.40982	0.94174	2.110554

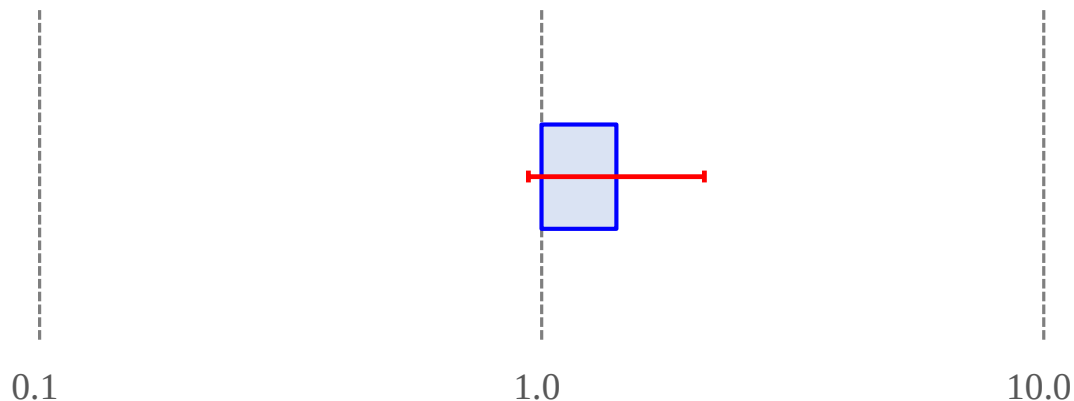
JMPの計算(カイ2乗検定の手順後)

▼からオッズ比を選択する.

$1/1.40982=0.7093$ なので, 逆数をとると先ほどと同じになる.



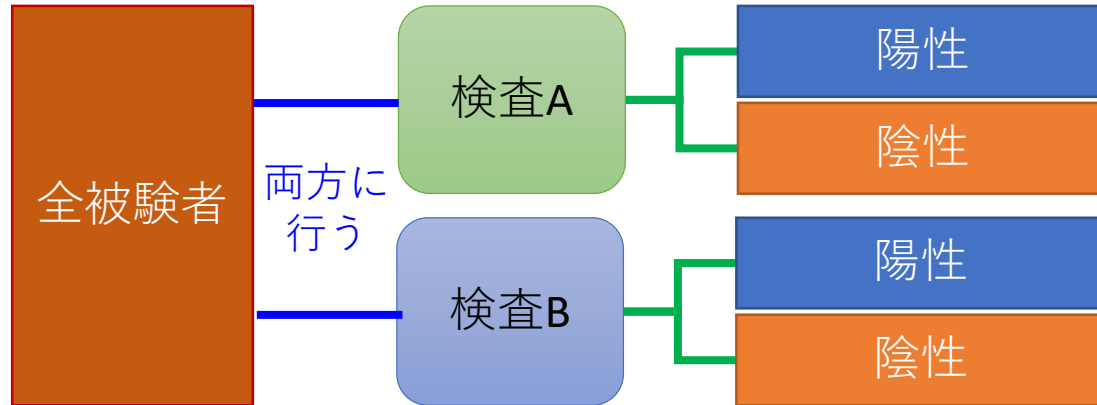
オッズ比の95%信頼区間が1.0をまたいでいることから, 帰無仮説「オッズ比が1である」が有意水準0.05のもとで受容される.



オッズ比をグラフ表現する場合に対数目盛を用いることで, 左右対称な信頼区間で表現できる.

7. 対応のあるデータにおけるクロス集計表

対応がある場合 (同一被験者に2種類の検査をした場合)



クロス集計表

		検査B	
		陽性	陰性
検査A	陽性	a	b
	陰性	c	d

対応がある場合 (同一被験者に2種類の治療をした場合)：クロスオーバー試験



クロス集計表

		治療B	
		あり	なし
治療A	あり	a	b
	なし	c	d

対応がない場合では、2群に分けられた後でそれぞれの群に対して異なる治療が行われる。これに対して、対応がある場合には、治療前後の比較、あるいは同一の個体に異なる治療・検査法が施される。

7. 対応のあるデータにおけるクロス集計表

治療B		治療間 被験者が独立 でない
治療A		
有効	治療A:有効 治療B:有効	治療A:有効 治療B:無効
無効	治療A:無効 治療B:有効	治療A:無効 治療B:無効

いま, ある2種類の治療法(治療A, 治療B)が同一の被験者に対して実施された(クロスオーバー試験). このとき, 治療効果は後治療に影響がないほどウォッシュアウトされているとする. 治療Aと治療Bのどちらがよいか.

比較という観点からいくと

- ・ 治療Aで有効なのに治療Bで無効(Aが優れている)割合「 の領域」

$$\hat{p}_A = \frac{\text{「治療A無し」かつ「治療B有り」数}}{\text{標本サイズ}}$$

- ・ 治療Bで有効なのに治療Aで無効(Bが優れている)割合「 の領域」

$$\hat{p}_B = \frac{\text{「治療A有り」かつ「治療B無し」数}}{\text{標本サイズ}}$$

を比較することになる. このような検定が**McNemear検定**である.

7. 対応のあるデータにおけるクロス集計表

下表は、皮膚かぶれの患者51名に湿布薬Aを右腕、湿布薬Bを左腕に湿布して効果を見たデータであるA, Bに差があるかどうか有意水準0.05で検定しなさい(柳川・荒木, 2010).

		湿布薬 B		
		有効	無効	合計
湿布薬 A	有効	11	6	17
	無効	10	24	34
	合計	21	30	51

■ 湿布薬 A が**有効**，湿布薬 B が**有効** = 11名

■ 湿布薬 A が**無効**，湿布薬 B が**無効** = 24名

■ 湿布薬 A が**有効**，湿布薬 B が**無効** = 6名 (11.8%) : A 薬のほうが有効

■ 湿布薬 A が**無効**，湿布薬 B が**有効** = 10名 (19.6%) : B 薬のほうが有効

7. 対応のあるデータにおけるクロス集計表

データセットからの場合

- 1: 「分析」→「2変量の関係」で変数を選択
 - ・目的変数: 湿布薬 A (湿布薬 B)
 - ・説明変数: 湿布薬 B (湿布薬 A)
- 2: ▼のなかから「一致制の統計量」を選択

Bowkerの検定

対称性 **カイ2乗** **p値(Prob>ChiSq)**

1

0.3173

2x2表の場合、Bowker検定はMcNemar検定と等価です。

クロス集計表からの場合

- 1: 「分析」→「2変量の関係」で変数を選択
 - ・目的変数: 湿布薬 A (湿布薬 B)
 - ・説明変数: 湿布薬 B (湿布薬 A)
 - ・度数: 度数
- 2: ▼のなかから「一致制の統計量」を選択

湿布薬A	湿布薬B	度数
有効	有効	11
有効	無効	6
無効	有効	10
無効	無効	24

p値=0.3173であることから、有意水準 $\alpha=0.05$ のもとで有意でなかった。したがって、湿布薬 A と湿布薬 B のあいだの有効率に違いがあるとは言えなかった。

* Bowkerの検定(McNemar-Bowkerの対称検定)とは、オリジナルのMcNemar検定が2×2クロス集計表のみだったものをk×lクロス集計表に拡張したようなもの(体格を挟んで対称になっているか否かを評価する検定)であり、2×2クロス集計表ではMcNemar検定と同じになる。

8. $K \times I$ クロス集計表におけるカイ2乗値

観測度数

	1	2	...	l	計
1	O_{11}	O_{12}	...	O_{1l}	n_1
2	O_{21}	O_{22}	...	O_{2l}	n_2
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$
k	O_{k1}	O_{k2}	...	O_{kl}	n_k
計	m_1	m_2	...	m_l	N

期待度数

					計
1	E_{11}	E_{12}	...	E_{1l}	n_1
2	E_{21}	E_{22}	...	E_{2l}	n_2
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$
k	E_{k1}	E_{k2}	...	E_{kl}	n_k
計	m_1	m_2	...	m_l	N

$$\text{期待度数の計算：} E_{ij} = \frac{n_i m_j}{N}$$

観測度数と期待度数との乖離度を用いている点では、 2×2 分割表と全く同じ

帰無仮説 H_0 ：行方向の項目と列方向の項目に関連性が**ない**。

対立仮説 H_1 ：行方向の項目と列方向の項目に関連性が**ある**。

8. K×Iクロス集計表におけるカイ2乗値

下表は、ある疾患に対する薬剤(新薬／既存薬)の有効性を3段階で評価した結果である。薬剤によって有効性に違いがあるか否かを検討しなさい。

	改善	不変	悪化	計
新薬	158	42	33	233
既存薬	112	67	66	245
計	270	109	99	478

観測度数

	改善	不変	悪化
新薬	158	42	33
既存薬	112	67	66



期待度数(帰無仮説のもとでの度数)

	改善	不変	悪化
新薬	$\frac{270 \times 233}{478} = 131.61$	$\frac{109 \times 233}{478} = 53.13$	$\frac{99 \times 233}{478} = 48.26$
既存薬	$\frac{270 \times 245}{478} = 138.39$	$\frac{109 \times 245}{478} = 55.87$	$\frac{99 \times 245}{478} = 50.74$

観測度数と期待度数の乖離度 = カイ2乗値

8.K×Iクロス集計表におけるカイ2乗値

観測度数

	改善	不変	悪化
新薬	158	42	33
既存薬	112	67	66

期待度数(帰無仮説のもとでの度数)

	改善	不変	悪化
新薬	$\frac{270 \times 233}{478} = 131.61$	$\frac{109 \times 233}{478} = 53.13$	$\frac{99 \times 233}{478} = 48.26$
既存薬	$\frac{270 \times 245}{478} = 138.39$	$\frac{109 \times 245}{478} = 55.87$	$\frac{99 \times 245}{478} = 50.74$

乖離

観測度数で標準化

	改善	不変	悪化
新薬	$158 - 131.61 = 26.39$	$42 - 53.13 = -12.13$	$33 - 48.26 = -15.26$
既存薬	$112 - 138.39 = -26.39$	$67 - 55.87 = 11.13$	$66 - 50.74 = 15.26$

	改善	不変	悪化
新薬	$\frac{(26.39)^2}{131.61} = 5.29$	$\frac{(-12.13)^2}{53.13} = 2.77$	$\frac{(-15.26)^2}{48.26} = 4.83$
既存薬	$\frac{(-26.39)^2}{138.39} = 5.03$	$\frac{(11.13)^2}{55.87} = 2.22$	$\frac{(15.26)^2}{50.74} = 4.59$

合計したものがカイ2乗値

JMPの結果

検定	カイ2乗	p値(Prob>ChiSq)
尤度比	24.573	<.0001*
Pearson	24.285	<.0001*

p値<0.0001であることから、有意だった。したがって、薬剤によって有効性に違いがあることが認められた。

6.おわりに

ご清聴ありがとうございました

