

カイ二乗検定（頻度、回数、人数などの相違）

使用目的：頻度、回数、人数などの質的データ（名義尺度）の相違を調べる。

・例えば、「携帯電話を持っていますか？」という質問を 100 名に Yes か No で答えてもらおうと、以下のような集計結果が得られたとします。

	人数
持っている	93
持っていない	7
合計	100

このような場合、カイ二乗 (χ^2) 検定を使います。英語では Chi-square test といいます。

- ・その他にも、あるコーパスにおける labor という語の頻度が、期待されるよりも多いか少ないかなどということを統計的に明らかにすることができます。
- ・どういうことを具体的にやっているかという、実際の回答の数（実測値）が、割合として期待される回答の数（期待値）とどれほど違っているのかを調べていると考えるとわかりやすいでしょう。例えば、上の携帯電話の例では、持っているか、持っていないかという 2 つのカテゴリーに対して、100 人の回答者がいたので、期待値は $100 \div 2 = 50$ となります。この期待値と実際の回答である実測値とどれだけずれているかをカイ二乗 (χ^2) 検定を使い確認します。

1. サンプルが 1 つの場合（適合度検定）

	人数
好き	25
ふつう	38
嫌い	18
計	81

【例】
 中学 2 年生の 2 クラスで「英語が好きですか？」という質問に対して、「好き」「ふつう」「きらい」という 3 つのどれかに回答するという調査をしました。回答を集計したとこと左の表のような結果になりました。
 この回答は統計的にどのような意味をもっているのでしょうか？

1.1. Excel を使って分析する

	O3		$=\$C\$6/3$
	A	B	C
1			
2		人数	期待値
3	好き	25	27
4	ふつう	38	27
5	嫌い	18	27
6	計	81	81
7			

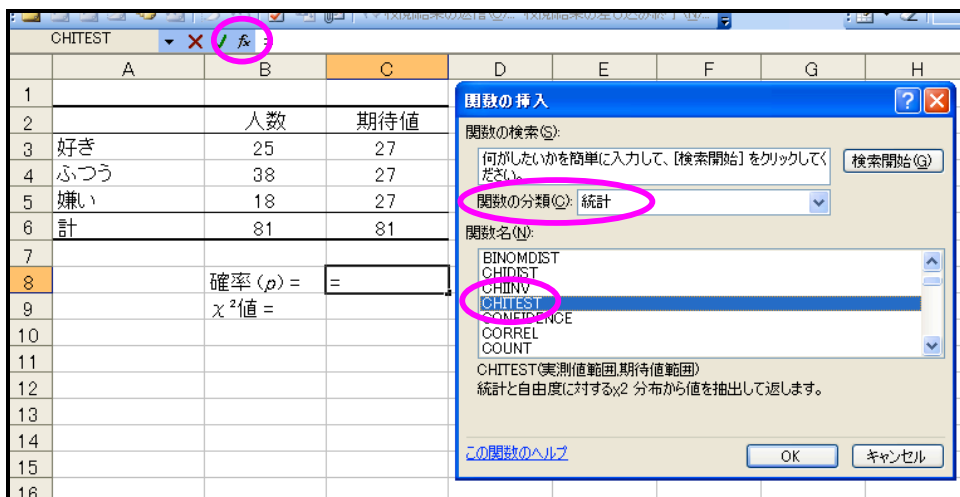
まずは左のように、実際の回答をまとめた隣に「期待値」を計算して入力します。

このデータの場合の期待値は、81（合計人数） \div 3（カテゴリー）で 27 となり、1 つのセルにつき、27 という数字が入ります。

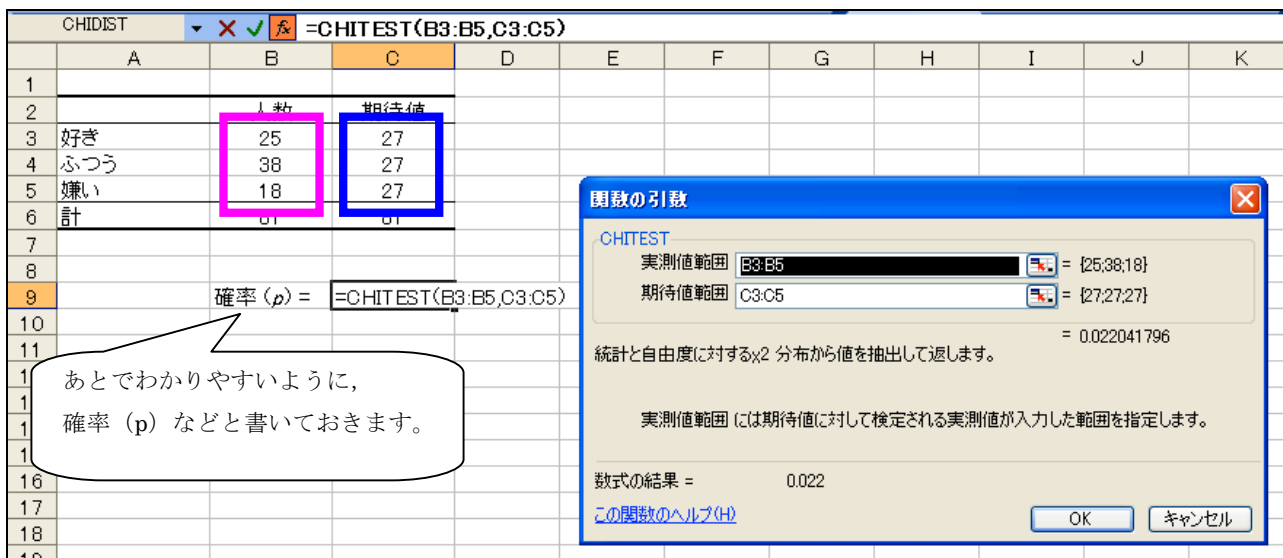
説明使用データ <http://www.mizumot.com/stats/chi-square.xls>

はじめにカイ二乗 (χ^2) の有意水準を計算します。

「関数の挿入」⇒「関数」⇒関数の分類 (C) : 中の「統計」⇒「CHITEST」を選ぶ。



以下のような画面になるので、「実測値範囲」には「人数」の列（以下では B3 から B5）を選択し、「期待値範囲」は「期待値」と書いている列（以下では C3 から C5）を選択する。



次に、 χ^2 (カイ二乗) 値を求めます。確率 (p) と書いてある下に χ^2 とでも書いておきましょう。再び「関数の挿入」から、「関数」⇒関数の分類 (C) : 中の「統計」⇒「CHINV」を選ぶ。



説明使用データ <http://www.mizumot.com/stats/chi-square.xls>

「確率」は先ほど「CHITEST」関数で求めたセルを選びます（以下の例では C9）。
そして「自由度」は、「3（カテゴリー数） - 1」なので「2」と入力します。

	A	B	C	D	E	F	G	H	I	J
1										
2		人数	期待値							
3	好き	25	27							
4	ふつう	38	27							
5	嫌い	18	27							
6	計	81	81							
7										
8		確率 (p) =	0.022							
9		χ^2 値 =	INV(C9,2)							
10										
11										
12										
13										
14										
15										
16										
17										
18										

以下のように χ^2 値が出力されれば完成です。

	A	B	C
1			
2		人数	期待値
3	好き	25	27
4	ふつう	38	27
5	嫌い	18	27
6	計	81	81
7			
8			
9		確率 (p) =	0.022
10		χ^2 値 =	7.630
11			

【 結果の報告 】

「ある中学校の 2 年生の 2 クラスで英語に対する意識を、「好き」「ふつう」「嫌い」という 3 つのどれかに回答するという調査を行った。その結果、 $\chi^2(2) = 7.63, p < .05$ で回答には有意差が認められた。」

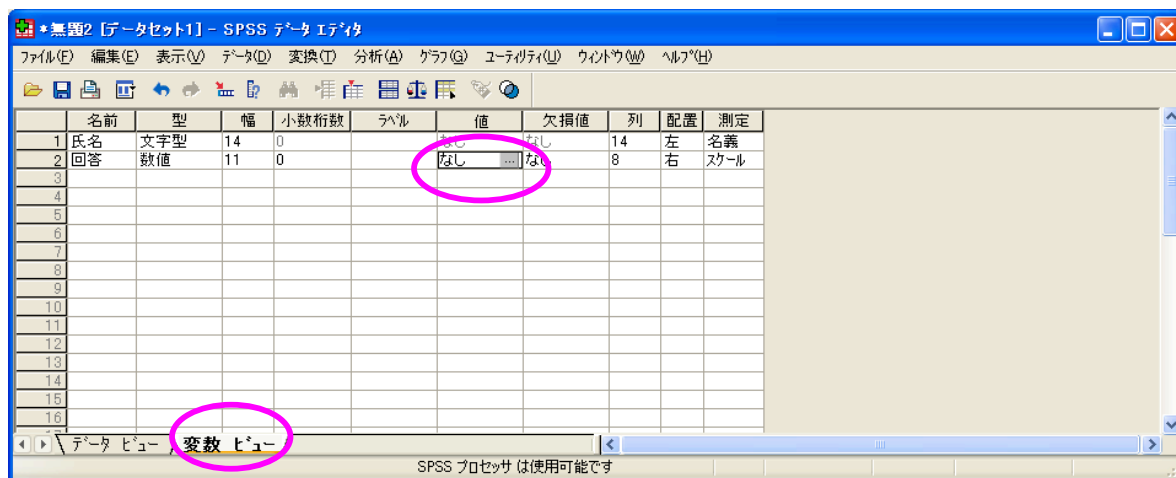
説明使用データ <http://www.mizumot.com/stats/chi-square.xls>

1.2. SPSS を使って分析する

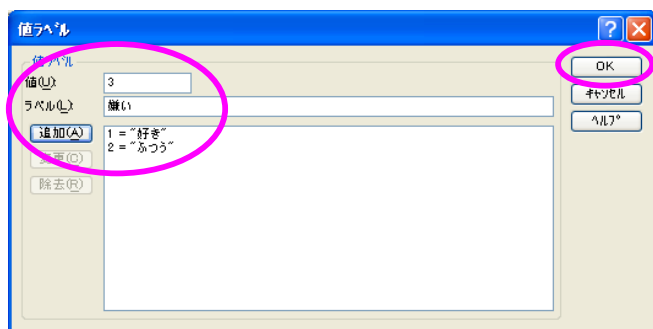
SPSS を使った分析では以下のように一人ずつの回答を「1=好き」, 「2=ふつう」, 「3=嫌い」という形で打ち込んだものがが必要です。Excel で入力したデータをドラッグ&ドロップで SPSS に移せば、かなり楽です。

	A	B	C	D
1	氏名	回答		
2	島田紳助	1		
3	安倍なつみ	2		1=好き
4	中島美嘉	3		2=ふつう
5	ベヨンジュン	2		3=嫌い
6	山下智久	2		
7	小堺一機	2		
8	にしきのあきら	2		
9	平井堅	2		
10	ビートたけし	1		
11	浅野忠信	2		
12	坂本冬美	3		
13	森口博子	1		
14	郷ひろみ	1		
15	長澤まさみ	1		

SPSS の「変数ビュー」をクリックして、値ラベルを選択して、「1=好き」, 「2=ふつう」, 「3=嫌い」というラベルをつけます。



以下の画面で「値」に「1」, 「ラベル」を「好き」として「追加」, 同様に「2」も「ふつう」, 「3」は「嫌い」と入力して追加したら、「OK」をクリック。



説明使用データ <http://www.mizumot.com/stats/chi-square.xls>

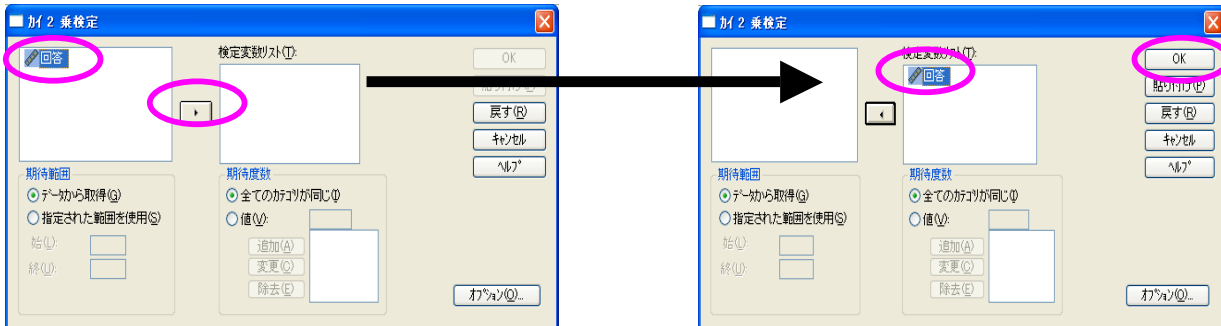
以下の画面に戻ったら、測定の尺度を「名義」に変更しておきましょう。



「分析」⇒「ノンパラメトリック検定」⇒「カイ2乗」を選ぶ。



以下のような画面が出てくるので、「回答」をクリックして、「検定変数リスト」へ移動させて「OK」。



回答			
	観測度数 N	期待度数 N	残差
好き	25	27.0	-2.0
ふつう	38	27.0	11.0
嫌い	18	27.0	-9.0
合計	81		

検定統計量	
	回答
カイ2乗 ^a	7.630
自由度	2
漸近有意確率	.022

a. 0.1セル(2%)の期待度数は 5 以下です。必要なセルの度数の最小値は 27.0 です。

【 結果の報告 】

「ある中学校の 2 年生の 2 クラスで英語に対する意識を、「好き」「ふつう」「嫌い」という 3 つのどれかに回答するという調査を行った。その結果、 $\chi^2(2) = 7.63$, $p < .05$ で回答には有意差が認められた。」

2. サンプルが2つ以上の場合（独立性の検定）

実測値	uptake (正しい文を繰り返した回数)	topic cont(話が続いた回数)	合計
No pf(ジェスチャーなし)	17	29	46
With pf(ジェスチャーあり)	35	7	42
	52	36	88

Source: Davies, M. (2007). Paralinguistic focus on form. *TESOL Quarterly*, 40, 841-855.

【例】

コミュニケーション的な指導法の中で、**Focus on Form**（言語形式の焦点化）と呼ばれるフィードバックの方法があります。**Focus on Form**では、会話の流れを止めることなく、学習者に間違いを気づかせて（もしくは気づかせずに）訂正させる指導を行います。

この**Focus on Form**がジェスチャーを交えることによって、いかに効果的になるかという研究を行いました。4クラスで5週間に渡り、ジェスチャーを交えた指導を行った中で、学習者がフィードバックで訂正された正しい文章を**uptake**（繰り返した）回数と、**uptake**することなく、話が続いた回数をビデオ撮影で振り返ってカウントしました。それをまとめたものが上の表です。表を見てみると、ジェスチャーを交えたほうが、**uptake**が起こりやすかったようです（35回）。

2.1. Excel を使って分析する

以下のような分割表のときには、式1の形になるように計算すれば χ^2 値を出すことができます。

	B ₁	B ₂	合計
A ₁	a	b	n ₁
A ₂	c	d	n ₂
合計	m ₁	m ₂	N

$$\chi^2 = \frac{(ad - bc)^2 \times N}{n_1 \times n_2 \times m_1 \times m_2} \quad (\text{式1})$$

詳しい計算はファイルの方を見てチェックしましょう。

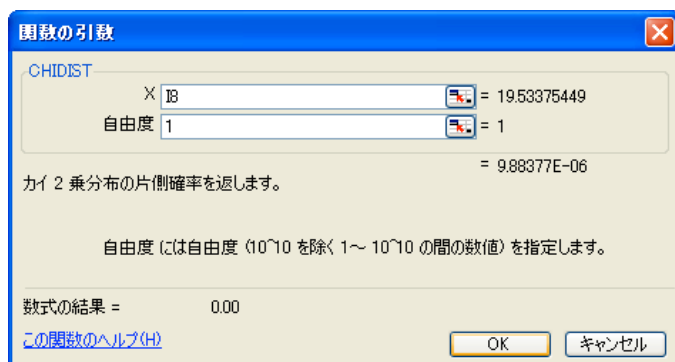
	H	I	J	K	L
1					
2	実測値	uptake (正しい文を繰り返した回数)	topic cont(話が続いた回数)	合計	
3	No pf(ジェスチャーなし)	17	29	46	
4	With pf(ジェスチャーあり)	35	7	42	
5		52	36	88	
6					
7					
8		=(((I3*I4)-(J3*I4))^2)*K5/(K3*K4*I5*J5)			
9					
10					

説明使用データ <http://www.mizumot.com/stats/chi-square.xls>

χ^2 値（以下の例では 19.53）が計算されたセルの隣に、有意確率を計算します。

「関数の挿入」⇒「統計」の中の「CHIDIST」を選択します。

	H	I	J	K
1				
2	実測値	uptake(正しい文を繰り返した回数)	topic cont(話が続いた回数)	合計
3	No pf(ジェスチャーなし)	17	29	46
4	With pf(ジェスチャーあり)	35	7	42
5		52	36	88
6				
7				
8			19.53375449	=CHIDIST(I8,1)
9				
10				



「X」には先ほど計算した χ^2 値を入れます（上の例では I8 のセル）。

「自由度」は (2-1) で「1」にします。

完成。

実測値	uptake(正しい文を繰り返した回数)	topic cont(話が続いた回数)	合計
No pf(ジェスチャーなし)	17	29	46
With pf(ジェスチャーあり)	35	7	42
	52	36	88
	χ^2	確率 (p)	
χ^2 検定	19.53	0.00	

★ イエーツの補正

今回の例のような 2×2 型のカイ二乗検定では、(1) データ数が少ないときと、(2) セルの中に「5」より小さい実測値がある場合には「イエーツの補正」(Yates' correction) を用いた方が良くとされます。今回は With pf (ジェスチャーあり) の topic cont (話が続いた回数) のセルに「7」という頻度がありますので、出典の論文ではイエーツの補正を行ったものを結果として挙げています。

実測値	uptake(正しい文を繰り返した回数)	topic cont(話が続いた回数)	合計
No pf(ジェスチャーなし)	17	29	46
With pf(ジェスチャーあり)	35	7	42
	52	36	88

イエーツの補正は式の中に $(-N/2)$ が入っているだけですが、以下の式 2 に書いているものです。

$$\chi^2 = \frac{\left(|ad - bc| - \frac{N}{2} \right)^2}{n_1 \times n_2 \times m_1 \times m_2} \quad (\text{式 2})$$

イエーツの補正を行った χ^2 値と有意確率を以下では求めています。どのような式になっているかは Excel ファイルを確認してみましょう。

実測値	uptake (正しい文を繰り返した回数)	topic cont (話が続いた回数)	合計
No pf (ジェスチャーなし)	17	29	46
With pf (ジェスチャーあり)	35	7	42
	52	36	88
	χ^2	確率 (p)	
χ^2 検定	19.53	0.00	
	=ABS(((\$C\$3*\$D\$4)-(\$D\$3*\$C\$4)-(\$E\$5/2))^2*\$E\$5/(\$E\$3*\$E\$4*\$C\$5*\$D\$5))		

前ページと同じように、「関数の挿入」⇒「統計」の中の「CHIDIST」を選択します。

実測値	uptake (正しい文を繰り返した回数)	topic cont (話が続いた回数)	合計
No pf (ジェスチャーなし)	17	29	46
With pf (ジェスチャーあり)	35	7	42
	52	36	88
	χ^2	確率 (p)	
χ^2 検定	19.53	0.00	
イエーツの補正	17.66	=CHIDIST(C9,1)	
		CHIDIST(χ^2 , 自由度)	

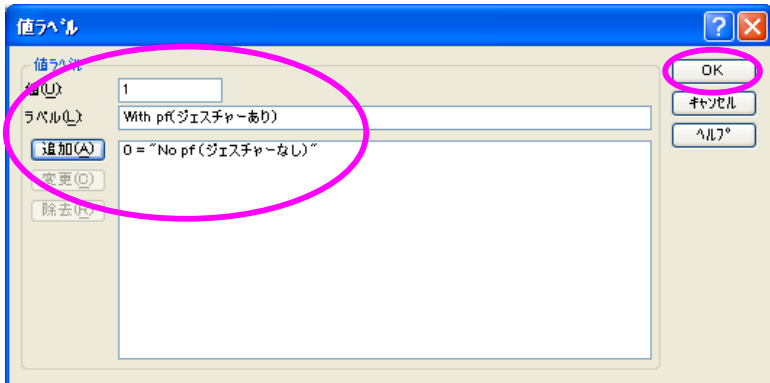
完成。

実測値	uptake (正しい文を繰り返した回数)	topic cont (話が続いた回数)	合計
No pf (ジェスチャーなし)	17	29	46
With pf (ジェスチャーあり)	35	7	42
	52	36	88
	χ^2	確率 (p)	
χ^2 検定	19.53	0.00	
イエーツの補正	17.66	0.00	

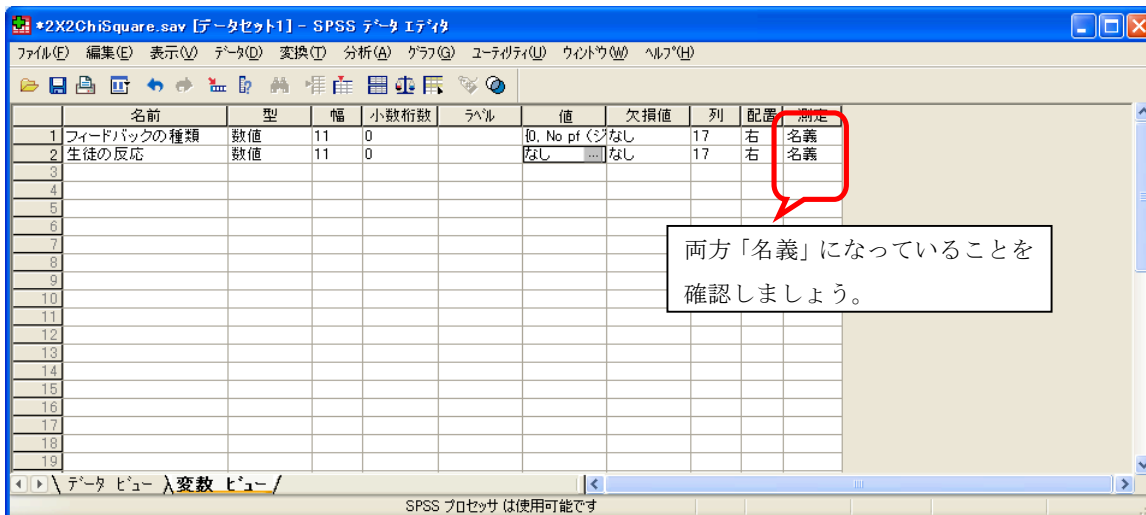
【 結果の報告 】

「ジェスチャーを交えたフィードバックにおいて、学習者がフィードバックで訂正された正しい文章を uptake (繰り返した) 回数と、uptake することなく、話が続いた回数をまとめたものが上の表である。これらの差をイエーツの補正を用いたカイ二乗検定で分析したところ、 $\chi^2(1) = 17.66, p < .01$ という結果となり、ジェスチャーがあるときとないときでは正しい文章を uptake する回数が異なるかもしれないということが明らかになった。

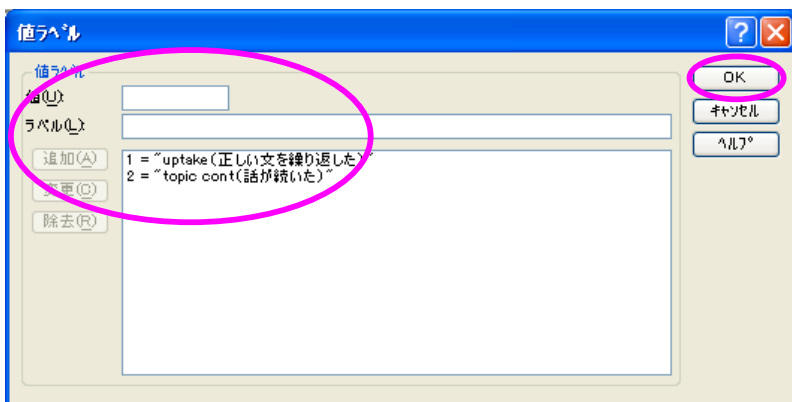
以下の画面で「値」に「0」,「ラベル」を「No pf (ジェスチャーなし)」として「追加」, 同様に「1」も「With pf (ジェスチャーあり)」と入力して追加したら, 「OK」をクリック。



もう一度以下の画面に戻ったら, 生徒の反応の値ラベルを選択して, 「1=uptake (正しい文を繰り返した)」, 「2=topic cont (話が続いた)」というラベルをつけます。



「値」に「1」,「ラベル」を「uptake (正しい文を繰り返した)」として「追加」, 同様に「2」も「topic cont (話が続いた)」と入力して追加したら, 「OK」をクリック。



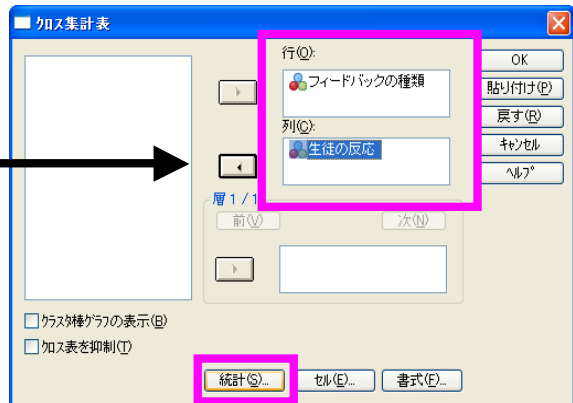
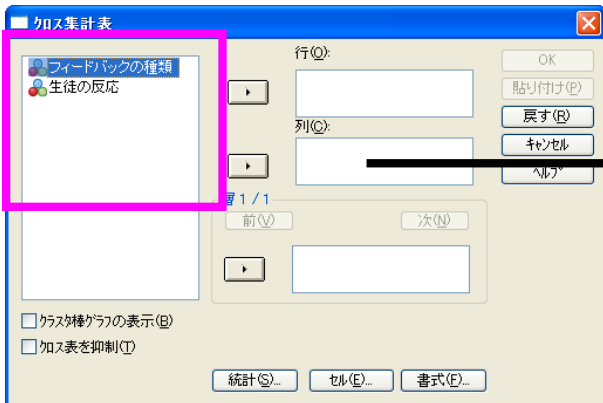
「データビュー」をクリックして、以下のような並びになっていることを確認してみよう。

	フィードバックの種類	生徒の反応
10	No pf (ジェスチャーなし)	uptake (正しい文を繰り返した)
11	No pf (ジェスチャーなし)	uptake (正しい文を繰り返した)
12	No pf (ジェスチャーなし)	uptake (正しい文を繰り返した)
13	No pf (ジェスチャーなし)	uptake (正しい文を繰り返した)
14	No pf (ジェスチャーなし)	uptake (正しい文を繰り返した)
15	No pf (ジェスチャーなし)	uptake (正しい文を繰り返した)
16	No pf (ジェスチャーなし)	uptake (正しい文を繰り返した)
17	No pf (ジェスチャーなし)	uptake (正しい文を繰り返した)
18	No pf (ジェスチャーなし)	topic cont (話が続いた)
19	No pf (ジェスチャーなし)	topic cont (話が続いた)
20	No pf (ジェスチャーなし)	topic cont (話が続いた)
21	No pf (ジェスチャーなし)	topic cont (話が続いた)
22	No pf (ジェスチャーなし)	topic cont (話が続いた)
23	No pf (ジェスチャーなし)	topic cont (話が続いた)
24	No pf (ジェスチャーなし)	topic cont (話が続いた)
25	No pf (ジェスチャーなし)	topic cont (話が続いた)
26	No pf (ジェスチャーなし)	topic cont (話が続いた)
27	No pf (ジェスチャーなし)	topic cont (話が続いた)
28	No pf (ジェスチャーなし)	topic cont (話が続いた)
29	No pf (ジェスチャーなし)	topic cont (話が続いた)
30	No pf (ジェスチャーなし)	topic cont (話が続いた)
31	No pf (ジェスチャーなし)	topic cont (話が続いた)
32	No pf (ジェスチャーなし)	topic cont (話が続いた)
33	No pf (ジェスチャーなし)	topic cont (話が続いた)
34	No pf (ジェスチャーなし)	topic cont (話が続いた)
35	No pf (ジェスチャーなし)	topic cont (話が続いた)
36	No pf (ジェスチャーなし)	topic cont (話が続いた)
37	No pf (ジェスチャーなし)	topic cont (話が続いた)
38	No pf (ジェスチャーなし)	topic cont (話が続いた)
39	No pf (ジェスチャーなし)	topic cont (話が続いた)
40	No pf (ジェスチャーなし)	topic cont (話が続いた)
41	No pf (ジェスチャーなし)	topic cont (話が続いた)
42	No pf (ジェスチャーなし)	topic cont (話が続いた)
43	No pf (ジェスチャーなし)	topic cont (話が続いた)
44	No pf (ジェスチャーなし)	topic cont (話が続いた)
45	No pf (ジェスチャーなし)	topic cont (話が続いた)
46	No pf (ジェスチャーなし)	topic cont (話が続いた)
47	With pf (ジェスチャーあり)	uptake (正しい文を繰り返した)
48	With pf (ジェスチャーあり)	uptake (正しい文を繰り返した)
49	With pf (ジェスチャーあり)	uptake (正しい文を繰り返した)
50	With pf (ジェスチャーあり)	uptake (正しい文を繰り返した)
51	With pf (ジェスチャーあり)	uptake (正しい文を繰り返した)
52	With pf (ジェスチャーあり)	uptake (正しい文を繰り返した)
53	With pf (ジェスチャーあり)	uptake (正しい文を繰り返した)
54	With pf (ジェスチャーあり)	uptake (正しい文を繰り返した)
55	With pf (ジェスチャーあり)	uptake (正しい文を繰り返した)
56	With pf (ジェスチャーあり)	uptake (正しい文を繰り返した)
57	With pf (ジェスチャーあり)	uptake (正しい文を繰り返した)
58	With pf (ジェスチャーあり)	uptake (正しい文を繰り返した)
59	With pf (ジェスチャーあり)	uptake (正しい文を繰り返した)
60	With pf (ジェスチャーあり)	uptake (正しい文を繰り返した)
61	With pf (ジェスチャーあり)	uptake (正しい文を繰り返した)
62	With pf (ジェスチャーあり)	uptake (正しい文を繰り返した)
63	With pf (ジェスチャーあり)	uptake (正しい文を繰り返した)
64	With pf (ジェスチャーあり)	uptake (正しい文を繰り返した)
65	With pf (ジェスチャーあり)	uptake (正しい文を繰り返した)

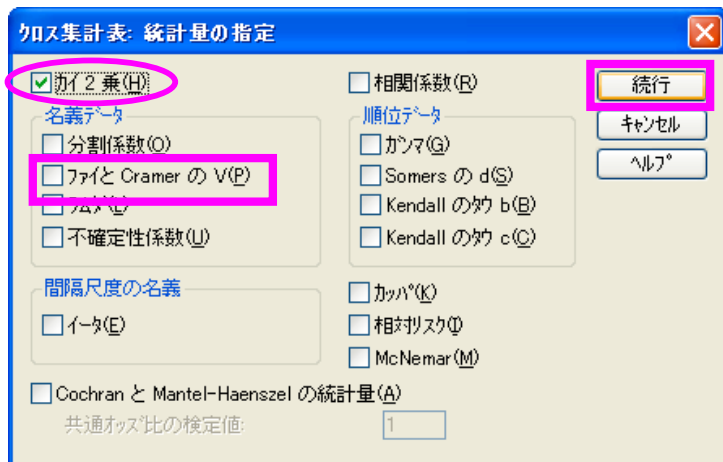
「分析」⇒「記述統計」⇒「クロス集計表」を選ぶ。

メニュー操作: 分析(A) > 記述統計(E) > クロス集計表(C)...

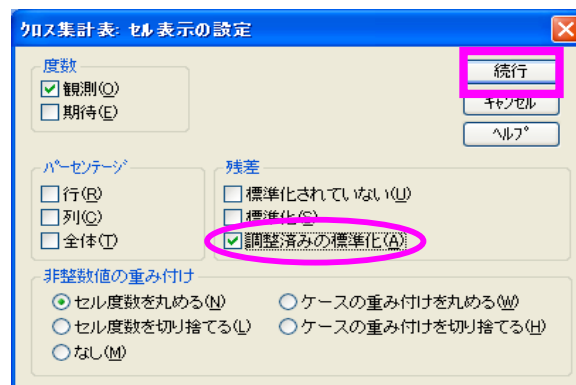
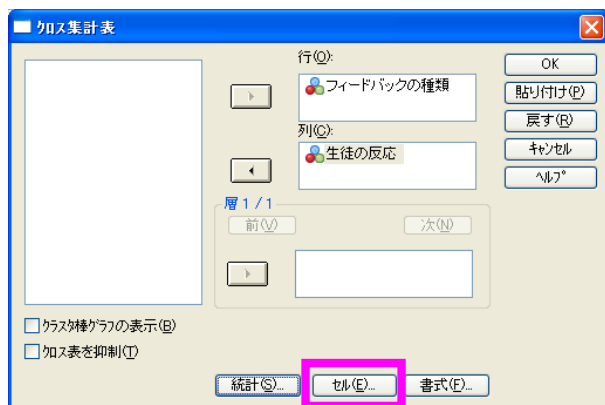
以下の画面になったら、「フィードバックの種類」を「行」のボックスへ移動し、「生徒の反応」を「列」に移動して「統計」をクリック。



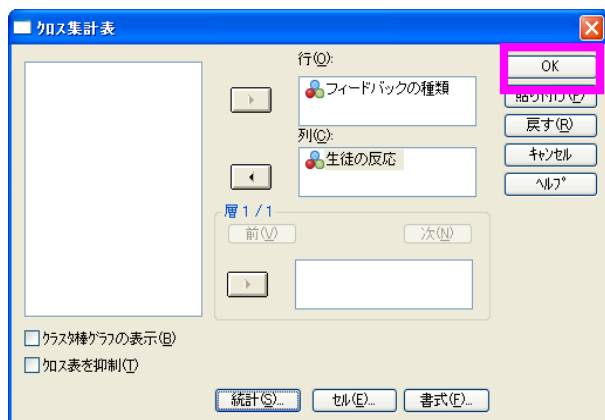
以下の画面が現れたら、「カイ 2 乗」にチェックを入れて、「続行」をクリック。
 ここで、ファイと Cramer の V にチェックを入れておけば、効果量が計算されます。



この画面に戻ったら、「セル」をクリックして、「調整済みの標準化」にチェックを入れて「続行」。



これで OK を押せば結果が出力される。



以下は結果が出力されたものです。

		生徒の反応		合計	
		uptake(正しい文章を繰り返した)	topic cont(話が続いた)		
フィードバックの種類	No pf (ジェスチャーなし)	度数	17	29	46
		期待度数	27.2	18.8	46.0
		調整済み残差	-4.4	4.4	
	With pf (ジェスチャーあり)	度数	35	7	42
		期待度数	24.8	17.2	42.0
		調整済み残差	4.4	-4.4	
合計		度数	52	36	88
		期待度数	52.0	36.0	88.0

	値	自由度	漸近有意確率 (両側)	正確有意確率 (両側)	正確有意確率 (片側)
Pearson のカイ2乗	19.534 ^b	1	.000		
連続修正 ^a	17.662	1	.000		
尤度比	20.619	1	.000		
Fisher の直接法				.000	.000
線型と線型による連関係のあるケースの数	19.312	1	.000		
	88				

a. 2x2 表に対してのみ計算
b. 0 セル (0%) は期待度数が 5 未満です。最小期待度数は 17.18 です。

「調整済み残差」はカイ二乗検定の結果、有意な関係が見られたときに、どのセルが関係をもたらしめているかを特定するための「残差分析」(residual analysis) の結果です。±1.96 以上であれば、そのセルが影響があるということになると考えられます。左の結果では、すべてのセルが大きな影響を持っていることがわかります。

カイ二乗の χ^2 値と有意確率は「Pearson のカイ 2 乗」の部分を確認する。

「連続修正」はイエーツの補正を行った χ^2 値を示している (Excel の結果と少し違うが、有意確率は同じになっている)。

ちなみに、「Fisher の直接法」は正式には「フィッシャーの正確確率検定」と呼ばれるものです。(1) セルの中に 0 に近い値があるときや、(2) 期待値が 5 より小さいとき、もしくは(3) 合計のセルである周辺度数が 10 以下程度の小さな値のときに用いるとされています。

【 結果の報告 】

実測値	uptake	topic cont	合計
No pf (ジェスチャーなし)	17	29	46
With pf (ジェスチャーあり)	35	7	42
合計	52	36	88

「ジェスチャーを交えたフィードバックにおいて、学習者がフィードバックで訂正された正しい文章を uptake (繰り返した) 回数と、uptake することなく、話が続いた回数をまとめたものが上の表である。これらの差をイエーツの補正を用いたカイ二乗検定で分析したところ、 $\chi^2(1) = 17.66, p < .01$ という結果となり、ジェスチャーがあるときとないときでは正しい文章を uptake する回数が異なるかもしれないということが明らかになった。」