

ノンパラメトリック検定

母集団の分布が**特定できない**とき

サンプル数が**少ない**とき

計数値、度数の場合

(計測値は計量値と呼びます)

パラメトリック検定	Z検定	平均 比率		
	t検定			
ノンパラメトリック検定	χ ² 検定	カテゴリーデータ 度数分布	単純集計表	適合度の検定
			クロス集計表	独立性の検定
		母分散の検定		
	U検定	順位データ	対応なし	
	サインランク検定		対応有り	

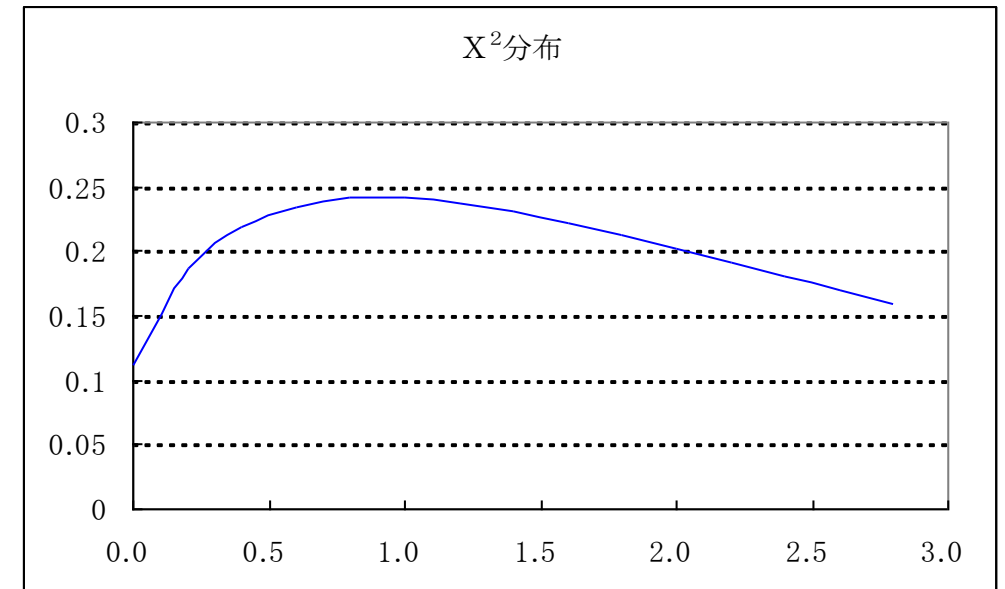
カイ χ^2 分布

統計量Tは χ^2 分布となる

$$\textcircled{1} T = \sum \frac{(\text{実測度数} - \text{期待度数})^2}{\text{期待度数}}$$

EXCEL: CHITEST(実測値範囲, 期待値範囲)

$$\textcircled{2} T = \frac{\text{標本サイズ} \times \text{標本分散}}{\text{母分散}}$$



←実験結果がこうなるはずだ
期待値からのズレを検定する

- ・適合度
- ・独立性

←母分散を推定する

適合度の検定 ホンパ・ラトリック検定 → χ^2 検定

ある大学で学生50人をランダムに選び、支持政党を調査したところ、下表の結果を得ました。この大学の学生は**特定の政党を支持する傾向を持つ**といえるでしょうか？有意水準0.05で検定しなさい。

政党	A	B	C	D	計
支持数	20	15	10	5	50

帰無仮説: $p_1 = p_2 = p_3 = p_4 = 1 \div 4 = 0.25$ 各政党の支持率は等しい

政党	A	B	C	D	計
n_i	20	15	10	5	$n = 50$
期待度数	$n \times p_1 = 12.5$	$n \times p_2 = 12.5$	$n \times p_3 = 12.5$	$n \times p_4 = 12.5$	

$$T = \sum \left(\frac{n_i - np_i}{np_i} \right)^2 = \frac{(20-12.5)^2}{12.5} + \frac{(15-12.5)^2}{12.5} + \frac{(10-12.5)^2}{12.5} + \frac{(5-12.5)^2}{12.5}$$

$$T = \sum \frac{(\text{実測度数} - \text{期待度数})^2}{\text{期待度数}} = \frac{7.5^2}{12.5} + \frac{2.5^2}{12.5} + \frac{(-2.5)^2}{12.5} + \frac{(-7.5)^2}{12.5} = 10$$

有意水準0.05 自由度 $f = C - 1 = 4 - 1 = 3$ χ^2 分布より

$T = 10 > 7.815 = \chi^2(f, \alpha) = \chi^2(3, 0.05)$ より帰無仮説は棄却

よって、この大学では、**各政党の支持率は等しくない(特定の政党を支持する傾向)**

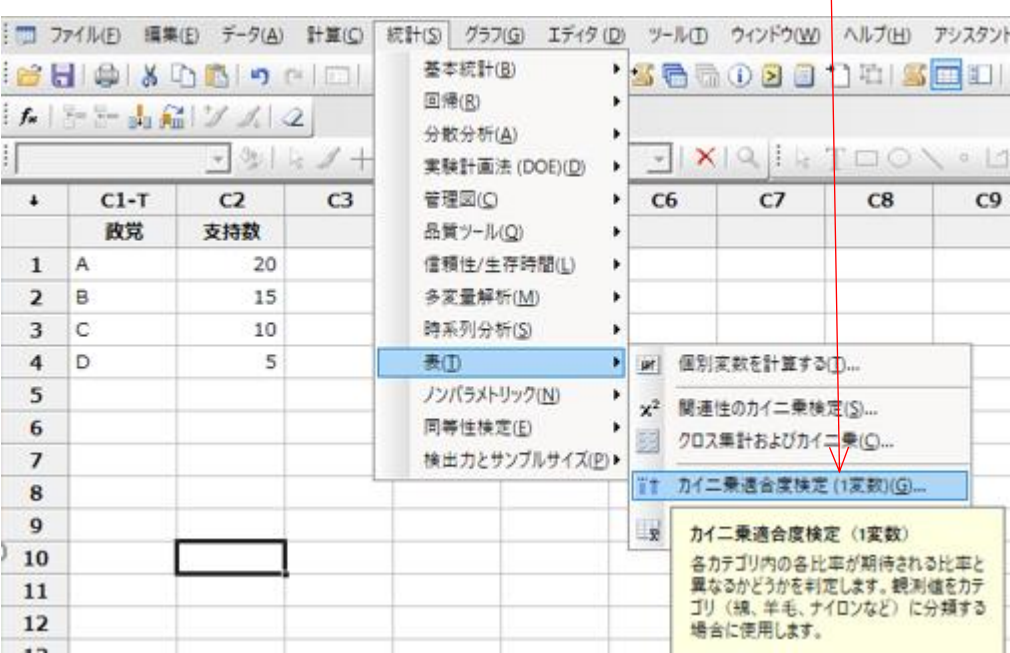
χ^2 分布表

Excel関数は CHIINV(確率,自由度)

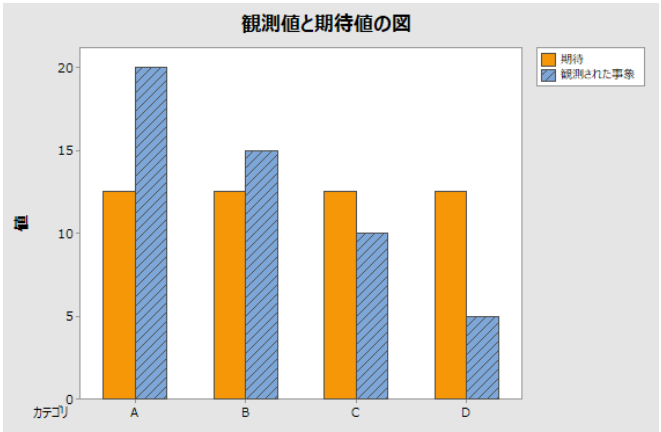
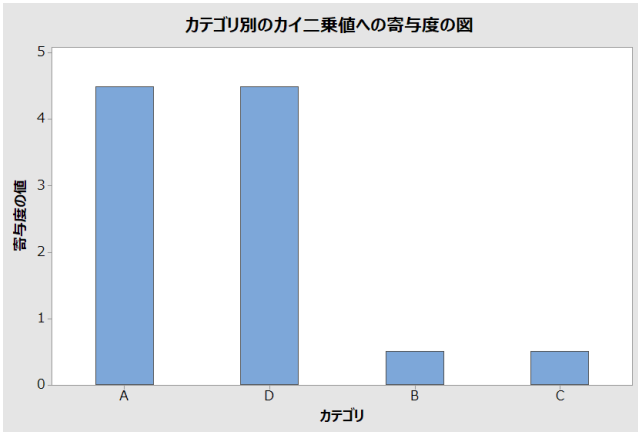
χ^2 分布表

	0.99	0.98	0.95	0.90	0.80	0.70	0.60	0.50	0.40	0.30	0.20	0.10	0.05	0.02	0.01	0.001
1	0.00016	0.00063	0.00393	0.01579	0.06418	0.14847	0.27500	0.45494	0.70833	1.07420	1.64238	2.70554	3.84146	5.41189	6.63490	10.82757
2	0.02010	0.04041	0.10259	0.21072	0.44629	0.71335	1.02165	1.38629	1.83258	2.40795	3.21888	4.60517	5.99146	7.82405	9.21034	13.81551
3	0.11483	0.18483	0.35185	0.58437	1.00517	1.42365	1.86917	2.36597	2.94617	3.66487	4.64163	6.25139	7.81473	9.83741	11.34487	16.26624
4	0.29711	0.42940	0.71072	1.06362	1.64878	2.19470	2.75284	3.35669	4.04463	4.87843	5.98862	7.77944	9.48773	11.66784	13.27670	18.46683
5	0.55430	0.75189	1.14548	1.61031	2.34253	2.99991	3.65550	4.35146	5.13187	6.06443	7.28928	9.23636	11.07050	13.38822	15.08627	20.51501
6	0.87209	1.13442	1.63538	2.20413	3.07009	3.82755	4.57015	5.34812	6.21076	7.23114	8.55806	10.64464	12.59159	15.03321	16.81189	22.45774
7	1.23904	1.56429	2.16735	2.83311	3.82232	4.67133	5.49323	6.34581	7.28321	8.38343	9.80325	12.01704	14.06714	16.62242	18.47531	24.32189
8	1.64650	2.03248	2.73264	3.48954	4.59357	5.52742	6.42265	7.34412	8.35053	9.52446	11.03009	13.36157	15.50731	18.16823	20.09024	26.12448
9	2.08790	2.53238	3.32511	4.16816	5.38005	6.39331	7.35703	8.34283	9.41364	10.65637	12.24215	14.68366	16.91898	19.67902	21.66599	27.87716
10	2.55821	3.05905	3.94030	4.86518	6.17908	7.26722	8.29547	9.34182	10.47324	11.78072	13.44196	15.98718	18.30704	21.16077	23.20925	29.58830

Minitab ①統計→表→カイニ乗適合度検定



実行結果



政党でカテゴリ名を使用

②選択

カイニ乗適合度検定

C1 政党
C2 支持数

観測度数(B): 支持数
カテゴリ名 (オプション)(C): 政党

検定
等しい比率(E)
特定の比率(S)
履歴度数で指定された比率(P): 入力列

選択
ヘルプ
グラフ(G)...
結果(R)...
OK(O)
キャンセル

カテゴリ	観測された事象	検定比率	期待	カイニ乗への寄与度の値
A	20	0.25	12.5	4.5
B	15	0.25	12.5	0.5
C	10	0.25	12.5	0.5
D	5	0.25	12.5	4.5

非欠損値	自由度	カイニ乗	p値
50	3	10	0.019

<0.05なので
各政党の支持率は等しくない
特定政党支持の傾向あり

独立性の検定 パラメトリック検定 → χ^2 検定

ある都市の有権者200名について、支持政党と性別を調査したところ、下表の結果を得ました。性別と支持政党は関連があるといえるでしょうか？有意水準0.05で検定しなさい。

		支持政党		
		A	B	C
性別	男	50	20	10
	女	30	40	50

帰無仮説： 2つの分類項目は独立で関連がない(期待度数との差がない)

		支持政党			合計
		A	B	C	
性別	男	0.16	0.12	0.12	0.4
	女	0.24	0.18	0.18	0.6
合計		0.4	0.3	0.3	1.0

		期待度数 支持政党			合計
		A	B	C	
性別	男	32	24	24	80
	女	48	36	36	120
合計		80	60	60	200

$$T = \sum \frac{(\text{実測度数} - \text{期待度数})^2}{\text{期待度数}}$$

$$T = \sum_{j=1}^A \sum_{i=1}^B \frac{(n_{ij} - \frac{n_j \times n_i}{n})^2}{\frac{n_j \times n_i}{n}}$$

$$T = \frac{(50 - 80 \times 80 \div 200)^2}{80 \times 80 \div 200} + \frac{(20 - 60 \times 80 \div 200)^2}{60 \times 80 \div 200} + \frac{(10 - 60 \times 80 \div 200)^2}{60 \times 80 \div 200} + \frac{(30 - 60 \times 120 \div 200)^2}{60 \times 120 \div 200} + \frac{(40 - 60 \times 120 \div 200)^2}{60 \times 120 \div 200} + \frac{(50 - 60 \times 120 \div 200)^2}{60 \times 120 \div 200} = 31.6$$

自由度 $f = (A-1)(B-1) = (2-1)(3-1) = 2$ 有意水準0.05のとき

$$T = 31.6 > 5.991 = \chi^2(f, \alpha) = \chi^2(2, 0.05)$$

帰無仮説を棄却する。性別と支持政党は独立でない。性別により支持政党が異なる傾向がある。

χ^2 分布表

Excel関数は CHIINV(確率,自由度)

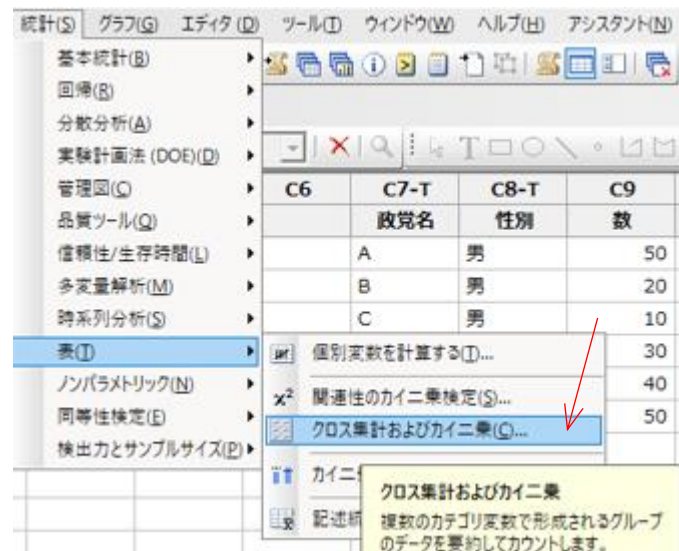
χ^2 分布表

	0.99	0.98	0.95	0.90	0.80	0.70	0.60	0.50	0.40	0.30	0.20	0.10	0.05	0.02	0.01	0.001
1	0.00016	0.00063	0.00393	0.01579	0.06418	0.14847	0.27500	0.45494	0.70833	1.07420	1.64238	2.70554	3.84146	5.41189	6.63490	10.82757
2	0.02010	0.04041	0.10259	0.21072	0.44629	0.71335	1.02165	1.38629	1.83258	2.40795	3.21888	4.60517	5.99146	7.82405	9.21034	13.81551
3	0.11483	0.18483	0.35185	0.58437	1.00517	1.42365	1.86917	2.36597	2.94617	3.66487	4.64163	6.25139	7.81473	9.83741	11.34487	16.26624
4	0.29711	0.42940	0.71072	1.06362	1.64878	2.19470	2.75284	3.35669	4.04463	4.87843	5.98862	7.77944	9.48773	11.66784	13.27670	18.46683
5	0.55430	0.75189	1.14548	1.61031	2.34253	2.99991	3.65550	4.35146	5.13187	6.06443	7.28928	9.23636	11.07050	13.38822	15.08627	20.51501
6	0.87209	1.13442	1.63538	2.20413	3.07009	3.82755	4.57015	5.34812	6.21076	7.23114	8.55806	10.64464	12.59159	15.03321	16.81189	22.45774
7	1.23904	1.56429	2.16735	2.83311	3.82232	4.67133	5.49323	6.34581	7.28321	8.38343	9.80325	12.01704	14.06714	16.62242	18.47531	24.32189
8	1.64650	2.03248	2.73264	3.48954	4.59357	5.52742	6.42265	7.34412	8.35053	9.52446	11.03009	13.36157	15.50731	18.16823	20.09024	26.12448
9	2.08790	2.53238	3.32511	4.16816	5.38005	6.39331	7.35703	8.34283	9.41364	10.65637	12.24215	14.68366	16.91898	19.67902	21.66599	27.87716
10	2.55821	3.05905	3.94030	4.86518	6.17908	7.26722	8.29547	9.34182	10.47324	11.78072	13.44196	15.98718	18.30704	21.16077	23.20925	29.58830

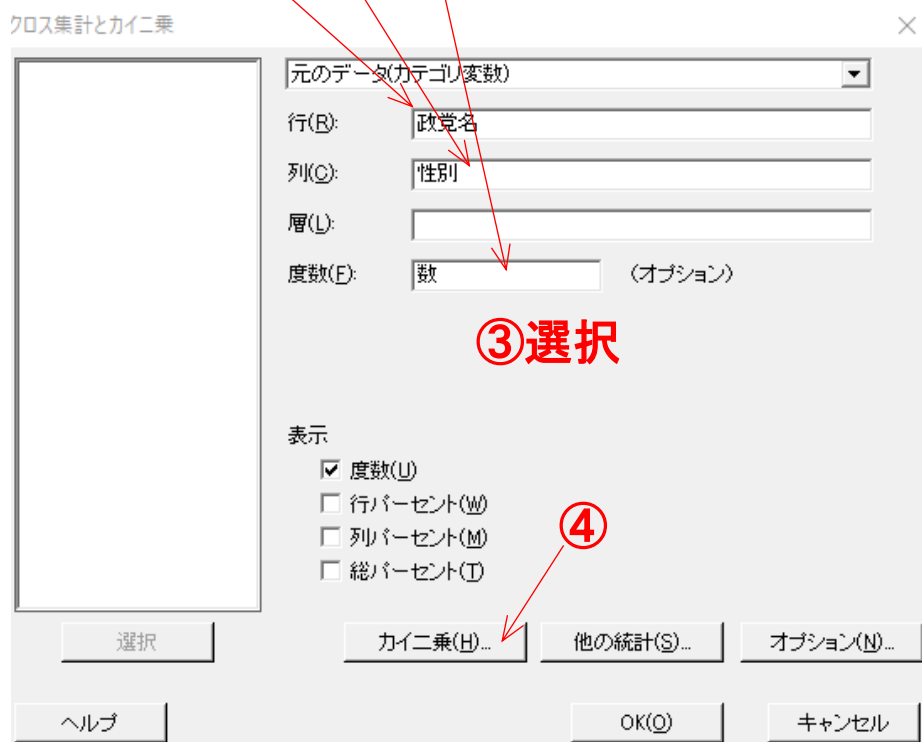
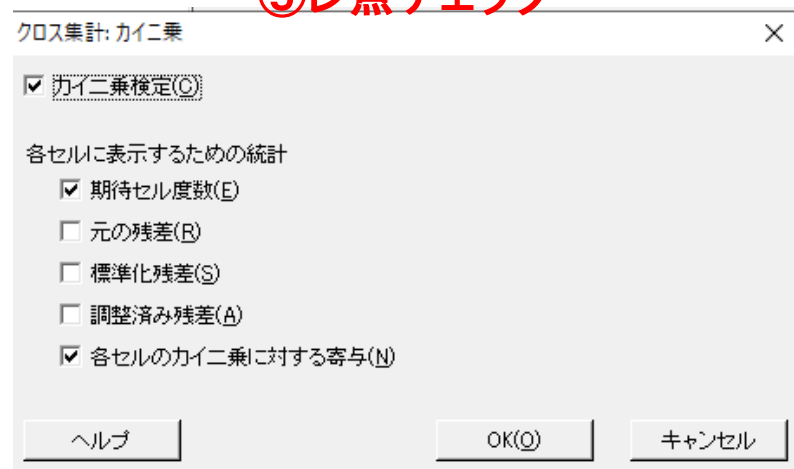
①データ入力

政党名	性別	数
A	男	50
B	男	20
C	男	10
A	女	30
B	女	40
C	女	50

②統計→表→クロス集計およびカイニ乗



⑤レ点チェック



③選択

④

行: 政党名 列: 性別

	女	男	すべて
A	30 48 6.750	50 32 10.125	80
B	40 36 0.444	20 24 0.667	60
C	50 36 5.444	10 24 8.167	60
すべて	120	80	200

セルの内容: 計数
期待度数
カイニ乗への寄与度

実行結果

<0.05なので
性別と支持政党は独立でない

ピアソンカイニ乗 = 31.597, 自由度 = 2, p値 = 0.000
尤度比カイニ乗 = 32.906, 自由度 = 2, p値 = 0.000

独立性の検定 パラメトリック検定→ χ^2 検定

開発中のある医薬品を、200症例についてテストしました。下表がその結果です。**この医薬品に効果があったといえるでしょうか？**有意水準0.05で検定しなさい。

	効果		計
	あり	なし	
投与	70	20	90
非投与	50	60	110
計	120	80	200

帰無仮説:「投与・非投与」と「効果あり・なし」は互いに独立である

$$T = \sum \frac{(\text{実測度数} - \text{期待度数})^2}{\text{期待度数}}$$

$$T = \sum_{j=1}^A \sum_{i=1}^B \frac{\left(n_{ij} - \frac{n_j \times n_i}{n}\right)^2}{\frac{n_j \times n_i}{n}} \quad T = \frac{(70-54)^2}{54} + \frac{(20-36)^2}{36} + \frac{(50-66)^2}{66} + \frac{(60-44)^2}{44} = 21.55$$

自由度 $f = (A-1)(B-1) = (2-1)(2-1) = 1$ 有意水準0.05のとき

$$T = 21.55 > 3.841 = \chi^2(f, \alpha) = \chi^2(1, 0.05)$$

帰無仮説を棄却する。「投与・非投与」と「効果あり・なし」は互いに独立でない。
この医薬品は効果があったといえる。

χ^2 分布表

Excel関数は CHIINV(確率,自由度)

χ^2 分布表

	0.99	0.98	0.95	0.90	0.80	0.70	0.60	0.50	0.40	0.30	0.20	0.10	0.05	0.02	0.01	0.001
1	0.00016	0.00063	0.00393	0.01579	0.06418	0.14847	0.27500	0.45494	0.70833	1.07420	1.64238	2.70554	3.84146	5.41189	6.63490	10.82757
2	0.02010	0.04041	0.10259	0.21072	0.44629	0.71335	1.02165	1.38629	1.83258	2.40795	3.21888	4.60517	5.99146	7.82405	9.21034	13.81551
3	0.11483	0.18483	0.35185	0.58437	1.00517	1.42365	1.86917	2.36597	2.94617	3.66487	4.64163	6.25139	7.81473	9.83741	11.34487	16.26624
4	0.29711	0.42940	0.71072	1.06362	1.64878	2.19470	2.75284	3.35669	4.04463	4.87843	5.98862	7.77944	9.48773	11.66784	13.27670	18.46683
5	0.55430	0.75189	1.14548	1.61031	2.34253	2.99991	3.65550	4.35146	5.13187	6.06443	7.28928	9.23636	11.07050	13.38822	15.08627	20.51501
6	0.87209	1.13442	1.63538	2.20413	3.07009	3.82755	4.57015	5.34812	6.21076	7.23114	8.55806	10.64464	12.59159	15.03321	16.81189	22.45774
7	1.23904	1.56429	2.16735	2.83311	3.82232	4.67133	5.49323	6.34581	7.28321	8.38343	9.80325	12.01704	14.06714	16.62242	18.47531	24.32189
8	1.64650	2.03248	2.73264	3.48954	4.59357	5.52742	6.42265	7.34412	8.35053	9.52446	11.03009	13.36157	15.50731	18.16823	20.09024	26.12448
9	2.08790	2.53238	3.32511	4.16816	5.38005	6.39331	7.35703	8.34283	9.41364	10.65637	12.24215	14.68366	16.91898	19.67902	21.66599	27.87716
10	2.55821	3.05905	3.94030	4.86518	6.17908	7.26722	8.29547	9.34182	10.47324	11.78072	13.44196	15.98718	18.30704	21.16077	23.20925	29.58830

Minitab

①データ入力

	C1-T	C2-T	C3
	投与	効果	結果
1	投与	あり	70
2	投与	なし	20
3	非投与	あり	50
4	非投与	なし	60

クロス集計とカイニ乗

元のデータ(カテゴリ変数)

行(R): 投与

列(C): 効果

層(L):

度数(E): 結果 (オプション)

表示

☒ 度数(U)

☐ 行パーセント(W)

☐ 列パーセント(M)

☐ 総パーセント(T)

選択

カイニ乗(H)...

他の統計(S)...

オプション(O)...

ヘルプ

OK(O)

キャンセル

③選択

クロス集計: カイニ乗

☒ カイニ乗検定(C)

各セルに表示するための統計

☒ 期待セル度数(E)

☐ 元の残差(R)

☐ 標準化残差(S)

☐ 調整済み残差(A)

☒ 各セルのカイニ乗に対する寄与(N)

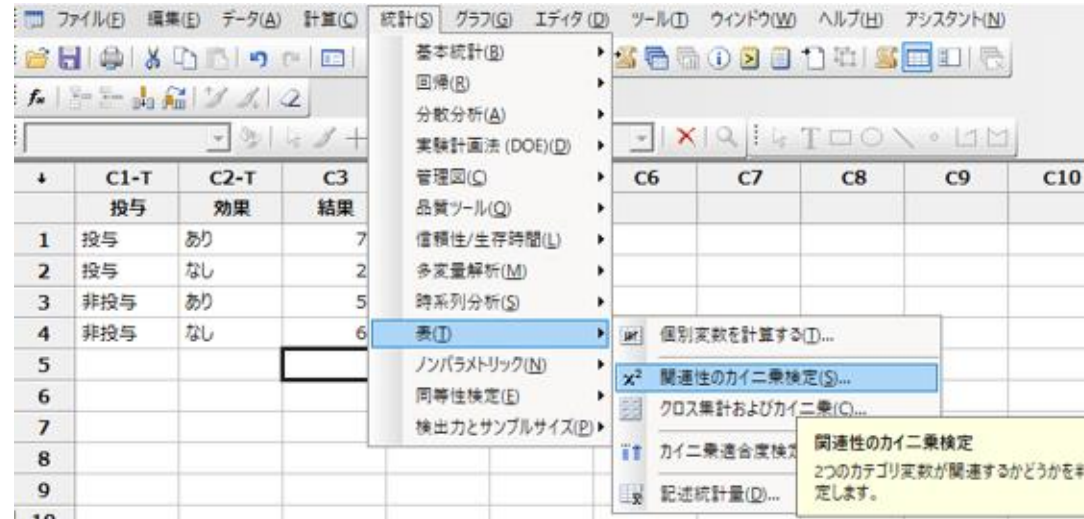
ヘルプ

OK(O)

キャンセル

④

②統計→表→クロス集計およびカイニ乗



集計統計量: 投与, 効果

結果の度数を使用します

	あり	なし	すべて
投与	70 54 4.741	20 36 7.111	90
非投与	50 66 3.879	60 44 5.818	110
すべて	120	80	200

セルの内容: 計数
期待度数
カイニ乗への寄与度

<0.05なので
投与・非投与と効果あり・なしは独立でない

ピアソンカイニ乗 = 21.549, 自由度 = 1, p値 = 0.000
尤度比カイニ乗 = 22.276, 自由度 = 1, p値 = 0.000