

複合集計 2変数: 質 × 質

- 1つの変数の分類ごとにもう一方の度数を数える
= 分割表, クロス表

ID	[1]				[2]
	(1)	(2)	(3)	(4)	好き嫌い
	性別	年齢	出身	居住	
1	1	4	1	2	1
2	1	4	8	2	1
3	1	4	2	2	2
4	1	4	1	2	3
5	1	4	1	2	4
6	2	3	1	2	4
7	2	3	1	2	1
8	2	3	1	2	1
9	1	3	2	2	2
10	2	3	1	2	2
11	1	3	1	2	1
12	1	3	1	2	3
13	1	3	1	2	4
14	1	3	1	2	1
15	1	3	1	2	2
16	1	3	2	2	1
17	1	3	7	2	4
18	1	3	1	2	3
19	1	3	1	2	3
20	1	3	1	2	4
21	1	3	7	2	1
22	1	3	1	2	4



意見	大変好き	少し好き	少し嫌い	大変嫌い	合計
男	82	44	14	12	152
女	30	16	14	74	134
合計	112	60	28	86	286

クロス表の分析: 問題設定

- 2つ製品PとQのどちらを好むかを標本調査により調べた。150人から回答を得、男女別に集計すると次のようになった。

性別	製品		計
	P	Q	
男性	36	34	70
女性	24	56	80
計	60	90	150

関心事 = 製品P, Qへの好みは、性別によって異なるといえるか。

クロス表の分析: クロス表 (分割表) (1)

性別	製品		計
	P	Q	
男性	36	34	70
女性	24	56	80
計	60	90	150

表側

表頭

セル

周辺度数
(周辺分布)

因果が想定されるとき

原因と考えられる方 (独立変数 / 目的変数) … 表側
結果と考えられる方 (従属変数 / 説明変数) … 表頭

クロス表の分析: クロス表 (分割表) (2)

■ 比率 (%) に注目した分析

- 行パーセント・・・セルの度数 ÷ 行の周辺度数
- 列パーセント・・・セルの度数 ÷ 列の周辺度数

行パーセント

性別	製 品		計
	P	Q	
男性	36 51%	34 49%	70 100%
女性	24 30%	56 70%	80 100%
計	60 40%	90 60%	150 100%

男性の好みは, P:Q = 51% : 49% と, ほとんど差はないが, 女性の場合は, P:Q = 30% : 70% と製品Qの方を好んでいる。

列パーセント

性別	製 品		計
	P	Q	
男性	36 60%	34 38%	70 47%
女性	24 40%	56 62%	80 53%
計	60 100%	90 100%	150 100%

製品Pを好む人は, 男性:女性 = 60% : 40% と男性の方が多いが, 製品Qは, 男性:女性 = 38% : 62% と女性の方が多。

クロス表の分析: クロス表 (分割表) (3)

女性の回答数が今のちょうど10倍であった場合

行パーセントの考察

結果に変わりはない

列パーセントの考察

製品Pは, 男性:女性 = 36人:240人 = 13%:87%

製品Qは, 男性:女性 = 34人:560人 = 6%:94%

となり, 製品Pについては, 結論が逆になってしまう。

行パーセント

性別	製 品		計
	P	Q	
男性	36 51%	34 49%	70 100%
女性	240 30%	560 70%	800 100%
計	276 32%	594 68%	870 100%

列パーセント

性別	製 品		計
	P	Q	
男性	36 13%	34 6%	70 8%
女性	240 87%	560 94%	800 92%
計	276 100%	594 100%	870 100%

行パーセントによる考察の方が適当
(行列を入れ替えている場合は, 列パーセントによる考察となる)。

クロス表の分析: 統計的仮説検定 (1)

- いったいどれぐらい性別が製品P, Qの好みに関係しているのか?

性別	製 品		計
	P	Q	
男性	36	34	70
女性	24	56	80
計	60	90	150

独立であることから
どれだけ差があるかを見極めればよい
= 統計的仮説検定

まったく関連していない (行と列は独立) ----- 最大に関連している

クロス表の分析: 統計的仮説検定 (2)

- ずれを1つの数値として表したものは χ^2 値

この値が十分大きい

今回観測された度数は独立である度数分布と「ずれている」

||

「差が有意である」

||

「関連がある」

性別	製品		計
	P	Q	
男性	36	34	70
女性	24	56	80
計	60	90	150

まったく関連していない (行と列は独立)  最大に関連している

クロス表の分析:

2独立性格検定 (1)

独立であることの
理論度数

	品		計
	P	Q	
男性	28 40%	42 60%	70 100%
女性	32 40%	48 60%	80 100%
計	60 40%	90 60%	150 100%

今回の調査による
観測度数

	品		計
	P	Q	
男性	36 51%	34 49%	70 100%
女性	24 30%	56 70%	80 100%
計	60 40%	90 60%	150 100%

製品の好みを男女で分けても
違いの考察ができない

||

製品の好みの違いに性別は
影響がない(無関係)

||

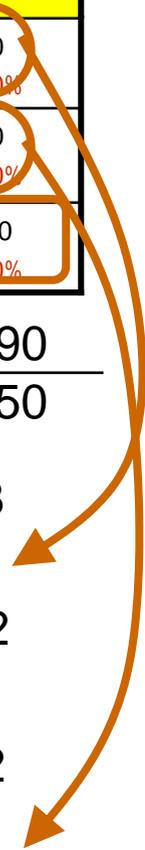
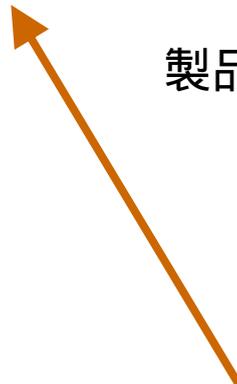
製品の好みと性別は独立

完全に独立

||

男性と女性の好み方が
全体(周辺度数)と同じ

製品P	$\frac{60}{150}$	製品Q	$\frac{90}{150}$
男 P	$70 \times \frac{60}{150} = 28$		
男 Q	$70 \times \frac{90}{150} = 42$		
女 P	$80 \times \frac{60}{150} = 32$		
女 Q	$80 \times \frac{90}{150} = 48$		



クロス表の分析:

2²独立性検定 (2)

性別	製品		計
	P	Q	
男性	28 40%	42 60%	70 100%
女性	32 40%	48 60%	80 100%
計	60 40%	90 60%	150 100%

性別	製品		計
	P	Q	
男性	36 51%	34 49%	70 100%
女性	24 30%	56 70%	80 100%
計	60 40%	90 60%	150 100%



まったく関連していない
(行と列は独立)

最大に関連している

1. 仮説の設定

H_0 : 男女間で製品への好みに差がない(行・列の度数配置は独立)

H_1 : 男女間で製品への好みに差がある(行・列の度数配置に偏り)

クロス表の分析: χ^2 独立性検定 (3)

性別	製品		計
	P	Q	
男性	28 40%	42 60%	70 100%
女性	32 40%	48 60%	80 100%
計	60 40%	90 60%	150 100%

性別	製品		計
	P	Q	
男性	36 51%	34 49%	70 100%
女性	24 30%	56 70%	80 100%
計	60 40%	90 60%	150 100%

理論度数と観測度数の差

性別	製品		計
	P	Q	
男性	8	-8	0
女性	-8	8	0
計	0	0	0

差の2乗 ÷ 理論度数

性別	製品		計
	P	Q	
男性	$\frac{8^2}{28} = 2.286$	$\frac{(-8)^2}{42} = 1.524$	3.810
女性	$\frac{(-8)^2}{32} = 2.000$	$\frac{8^2}{48} = 1.333$	3.333
計	4.286	2.857	7.143

2. 統計量の算出

χ^2 値 = 7.143

総和

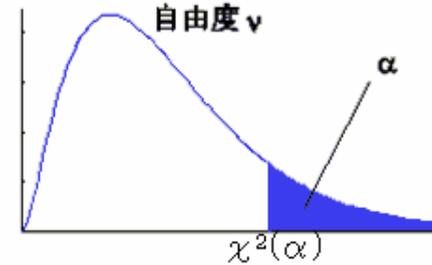
クロス表の分析:

χ^2 独立性検定 (4)

χ^2 分布表

$$\alpha = P(\chi^2_{\alpha}(v) \leq \chi^2)$$

自由度 $v=5$ で $\alpha=0.05$ を満たす $\chi^2_{\alpha}(v)$ は, 11.070



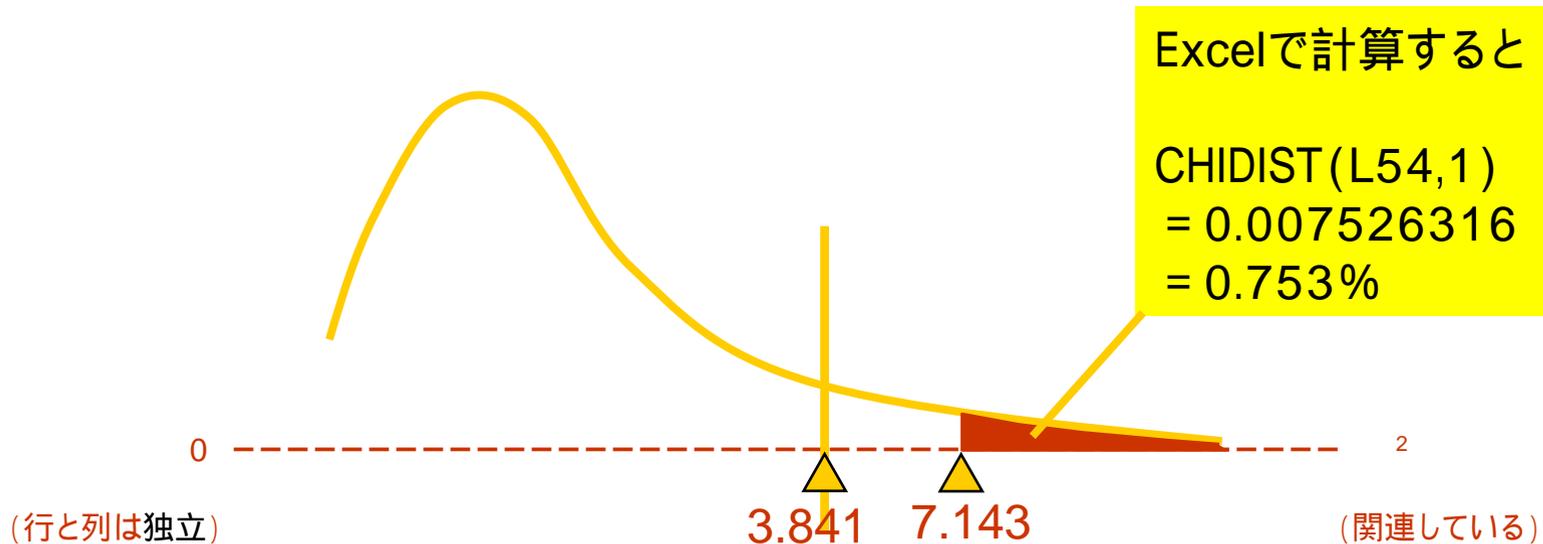
$\nu \setminus \alpha$	0.100	0.050	0.025	0.020	0.010	0.005	0.001
1	2.706	3.841	5.024	5.412	6.635	7.879	10.828
2	4.605	5.991	7.378	7.824	9.210	10.597	13.816
3	6.251	7.815	9.348	9.837	11.345	12.838	16.266
4	7.779	9.488	11.143	11.668	13.277	14.860	18.467
5	9.236	11.070	12.833	13.388	15.086	16.750	20.515
6	10.645	12.592	14.449	15.033	16.812	18.548	22.458
7	12.017	14.067	16.013	16.622	18.475	20.278	24.322
8	13.362	15.507	17.535	18.168	20.090	21.955	26.124
9	14.684	16.919	19.023	19.679	21.666	23.589	27.877
10	15.987	18.307	20.483	21.161	23.209	25.188	29.588

3. 確率の算出

- 自由度1の χ^2 分布表から5% (=0.050) 点の値をとる
 $\chi^2 = 7.143 > 3.841 = \chi^2_{0.050}(1)$

クロス表の分析:

2²独立性検定 (4)



4. 判定

男女間の製品の好み方に差がないとしたときに、今回のような差が起こるのは5%より確率が小さい。したがって、仮説 H_0 を棄却して H_1 を採用し、**「有意水準5%で、男女間で製品PとQへの好みに違いがあるといえる」と結論する。**