

## ■ 『EXCEL 統計 Ver. 8.0』の独立性の検定の動作イメージ

『EXCEL 統計 Ver. 8.0』に搭載されている独立性の検定、通称「カイ2乗検定」と呼ばれる手法の動作イメージを紹介します。

	A	B	C	D	E	F	G	H	I	J	K	L
1		商品1	商品2	商品3	商品4							
2	20代以下	43	9	10	24	86	<div style="border: 1px solid gray; padding: 5px;"> <p>独立性の検定 <span style="float: right;">×</span></p> <p>データタイプ クロス集計表 <span style="margin-left: 20px;">OK</span> <span style="margin-left: 20px;">キャンセル</span></p> <hr/> <p>範囲指定</p> <p>データ範囲 <input type="text" value="\$A\$1:\$E\$4"/> <span style="float: right;">1行目と1列目</span></p> <p>結果出力先 <input type="text" value="\$B\$9"/> <span style="float: right;"><input checked="" type="radio"/> データラベル</span></p> <p style="text-align: right;"><input type="radio"/> データ</p> <p><input type="checkbox"/> 2×2より大きいクロス表に対して正確検定を実行しない</p> <hr/> <p>残差分析</p> <p><input type="radio"/> しない <input checked="" type="radio"/> 併記表で出力 <input type="radio"/> 分離表で出力</p> <p><input type="checkbox"/> 実測値 <input type="checkbox"/> 横% <input type="checkbox"/> 縦% <input type="checkbox"/> 全%</p> <p><input type="checkbox"/> 期待度数 <input type="checkbox"/> 単純残差 <input checked="" type="checkbox"/> 調整残差 <input checked="" type="checkbox"/> 判定</p> </div>					
3	30~40代	44	7	9	37	97						
4	50代以上	33	0	2	21	56						
5		120	16	21	82							
6												

※ 上記は表側が3カテゴリー、表頭が4カテゴリーのクロス集計です。

- (1) メニューから「独立性の検定」を選びます。
- (2) ダイアログボックスが開きますので、セルA1~E4を範囲指定します。5行目およびF列の合計は範囲指定しません。
- (3) 結果出力先はセルB9を指定します。
- (4) 「残差分析」で「併記表で出力」を選択して、「調整残差」と「判定」の2つをオンにしたら、「OK」ボタンを押します。
- (5) 次ページのように、検定結果と残差分析が出力されます。

## ◎ 独立性の検定

件数	分布	自由度	クラメールの連関係数
239	カイ二乗分	6	0.152

	統計量	p値	判定
適合度	11.08	0.086	n.s.
尤度比	15.01	0.020*	

期待度数が5以下のセルが存在します。12セルのうち、2セル(16.7%)あります。

## ■ Fisherの正確検定

	p値	判定
両側	0.056	n.s.

n.s.: p値 $\geq$ 0.05、\*: p値 $<$ 0.05、\*\*: p値 $<$ 0.01

## ■ 残差分析

		商品1	商品2	商品3	商品4
20代以下	調整残差	-0.05	1.75	1.16	-1.56
	判定	[ ]	[ ]	[ ]	[ ]
30~40代	調整残差	-1.24	0.27	0.22	1.03
	判定	[ ]	[ ]	[ ]	[ ]
50代以上	調整残差	1.49	-2.29	-1.58	0.57
	判定	[ ]	[/]	[ ]	[ ]

[/]: 調整残差 $<$ -2.58

[/]: -2.58 $<$ 調整残差 $<$ -1.96

[ ]: -1.96 $<$ 調整残差 $<$ 1.96

[\*]: 1.96 $<$ 調整残差 $<$ 2.58

[\*\*]: 調整残差 $>$ 2.58

『EXCEL 統計 Ver. 8.0』では **2×2 (2行×2列) 以外**のクロス表に対してもフィッシャーの正確検定(フィッシャーの正確確率検定、フィッシャーの直接確率検定、フィッシャーの直接法、フィッシャーの直接確率計算)ができるようになりました。

例えば、前ページのデータであれば、70秒ほど(CPU: Intel Core i5-8500T、メモリ: 8GB)で結果(左記)が出力されます。

このデータでは正確検定のp値は0.056でした。

16行目には期待度数が5以下のセルが何個あったのかも出力されます。

p値が全部で3つ出力されますが、独立性の検定(カイ2乗検定)としてよく使われるのは15行目にある「適合度」のp値です。

残差分析で「しない」以外を選択すると、残差分析が出力され、どのセルが有意に大きいまたは小さいかがわかります。

(27, 29, 31行目の判定で判断します。\*印がついていれば度数の多いセル、/印がついていれば度数の少ないセルを意味しますが、一般的な表記方法ではないので、発表する際は32~36行目の凡例も必要です。)

このデータでは「50代以上」の「商品2」が小さいという結果でした。

**残差分析は「適合度」検定の下位検定ではありません**ので、適合度検定で有意になっていないのに、残差分析で有意なセルがあることもあります。

説明が前後しますが、p値の大小だけでなく、12行目に出力されるクラメールの連関係数(クラメールのV、Cramer's V、クラマーのV)も確認してください。

統計学的に有意であっても、連関係数が小さければ実用上の意味は小さいからです。クラメールの連関係数は名義変数同士の相関係数のようなもので、0~1の数値になります。

フィッシャーの正確検定(Fisher's exact test, Fisher-Irwin test)について、**度数や組み合わせ数が多くなると計算できません**が、そうでなければ、適用できるクロス表の大きさは以下のとおりです。

・一方が2行または2列の場合:	2行3列(3行2列)、2行4列(4行2列)、2行5列(5行2列)、2行6列(6行2列)、2行7列(7行2列)、2行8列(8行2列)、2行9列(9行2列)、2行10列(10行2列)
・一方が3行または3列の場合:	3行4列(4行3列)、3行5列(5行3列)、3行6列(6行3列)、3行7列(7行3列) 3行4列(4行3列)、3行5列(5行3列)、3行6列(6行3列)、3行7列(7行3列)
・一方が4行または4列の場合:	4行5列(5行4列)、4行6列(6行4列)
・一方が5行または5列の場合:	5行5列