自由度１　の　χ^2 モデル　計算方法

山本　洋一

A

0104277441
0158930933
0178334532
0113590342
0145775079

　データ　A と B は、
双方とも　50 個の数字です
A は、
数字　01××××××××
先頭数字 0 と 1　他 8 桁乱数　総数 50
B は、
数字　××××××××××
…　10 桁乱数　　総数 50

B

4452823928
4991821667
1057364247
2606422093
4003174372

統計を取りますと下記の表のようになります。
χ^2公式を用いでどういう結果なるか調査してみます。

マトリックス　**2×2**　分割表の場合　　　χ^2**=3.508**
その他の方法　適合性の検定の場合　　　χ^2**=9.52**

(1) マトリックス　**2×2**　分割表

	0 と 1	2～9	計
A	16	34	50
B	8	42	50
計	24	76	100

上記数値を下記公式にて計算
公式　$\chi^2=\{n(eh-fg)^2\}/abcd$ にて計算します。

A	e	f	a
B	g	h	b
計	c	d	n

$\chi^2=3.508$　<自由度１ $\chi^2=3.841$　　5% 有意ならず。

(2)　適合性の検定　　0～1 …　8　2～9 …　42　を理論値とみなし
算出。

		実測	理論値
0～1	A_1	16	8
2～9	A_2	34	42

総数　50

公式　$\chi^2=(A_1-理論値_1)^2/理論値_1+(A_1-理論値_2)^2/理論値_2$
で計算します。

$\chi^2=(16-8)^2/8+(34-42)^2/42$　9.52

$\chi^2=9.52$　>自由度１ $\chi^2=3.841$　　5% 有意

(1) の結果　5%　有意ならず。　　$\chi^2=3.508$
(2) の結果　〃　　有意　　　$\chi^2=9.52$　　**(1)≦ (2)**
(2) が必ず大きくなります。
上記の差異について、どうしてなのか、吟味してみます。

違いの証明

マトリックス　2×2 部活表には、理論値が隠されている。

マトリックス 2×2

			計
A	e	f	a
B	g	h	b
計	c	d	n

χ自乗自由度1=0　マトリックス

			計
A	ac/n	ad/n	a
B	bc/n	bd/n	b
計	c	d	n

公式　$\chi^2=\{(eh-fg)^2 n\} / abcd$　おける　実測と理論。

マトリックス 2×2 実測

実測

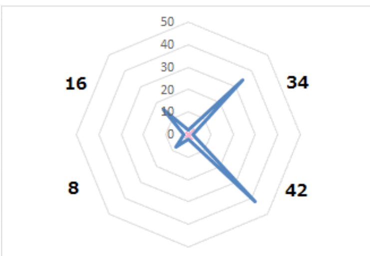
			計
A	16	34	50
B	8	42	50
計	24	76	100

χ自乗自由度1=0　マトリックス

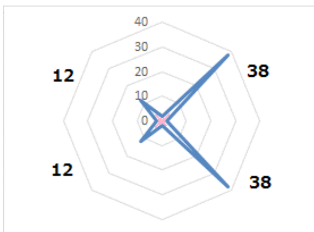
隠されている。

			計
A	12	38	50
B	12	38	50
計	24	76	100

実測値



理論値



実測値―理論値

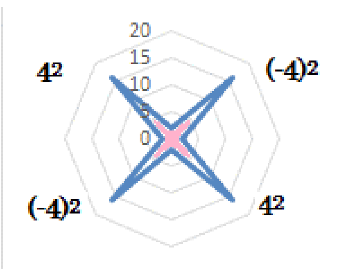
			計
A	16-12	34-38	0
B	8-12	42-38	0
計	0	76	0

結果

			計
A	4	-4	0
B	-4	4	0
計	0	0	0

$\chi^2=\{(eh-fg)^2 n\} / abcd$

＝



$(16\times42-34\times8)^2\times100 / (50\times50\times24\times76)$

$=3.508$

適合性の検定別解

$\chi^2=\sum\{(\text{実測値-理論値})^2 / \text{理論値}\} / \text{自由度}$

$=\chi^2=4^2/12+(-4)^2/38+(-4)^2/12+4^2/38$

$\chi^2=3.508$　　自由度=1

σ= √(npq)　二項分布による証明　　Z 値　算出
σ= √(npq)　を用いた。　$z^2 \Rightarrow \chi^2$ 自由度１ 証明

もしも、A を実測値　B を理論値にすれば…

p＝0.16　q＝0.84　n=100
σ= √(npq)＝√(50×0.16×0.84)
z =(16-8) / √(50×0.16×0.84)=2.59　16 は、実測値　　8 は、理論値
-z =(34-42) /√(50×0.16×0.84)=-2,59　　34 は、実測値　　42 は、理論値
$z^2=(16-8)^2$ / (50×0.16×0.84) =9.52　　$z^2 \Rightarrow \chi^2$　　5%有意…(3)

適合性の検定
χ^2 = (A_1 理論値 $_1$) ^{2} / 理論値 $_1$＋ (A_1－理論値 $_2$) ^{2} / 理論値 $_2$
＝9.52　　5%有意　　…(4)
結果　(1)　≦　　結果(2) 、　(3)　、　(4)　一致
必ず　(2) (3) (4) はマトリックスより大きくなります。

マトリックス　n=100 検定

マトリックス 2 × 2 実測

			計
A	e	f	a
B	g	h	b
計	c	d	n

A	16	34	50
B	8	42	50
	24	76	100

理論値

			計
A	ac	ad	a
B	bc	bd	b
計	c	d	n

			計
A	12	38	50
B	12	38	50
	24	76	100

ここで　　eh≠fg　　ac×bd＝ad×bc

p = a / n　q = c / n　　　r = c / n　s = d / n
p= 0.5　q= 0.5　　　r= 0.24　s= 0.76　　　n= 100

マトリックスを　二項分布用の Z 値算出による検定
σ＝√(npqrs) 標準偏差より

			計
A	pr	sp	p
B	rq	sq	q
計	r	s	1

			計
A	0.12	0.38	0.5
B	0.12	0.38	0.5
計	0.24	0.76	1

ここで、　　pr×sq = sp×rq

σ＝√(npqrs)　　標準偏差
a=16　実測値　　b=34　実測値　　npr=12 理論値　　nsp=38 理論値
手順
z = (a-npr) / √(npqrs) または　-z = (b-nsp) / √(npqrs)

z^2 = (a-nrp)^2 / (npqrs)
z^2 ⇒　χ自乗　自由度１

下記実際計算例　…
　a=16　b=34　c=8　d=42　rp=0.12　sp=0.38
z^2 =(a-npr)^2 / (npqrs) =4^2 / {100×(0.5×0.5×0.24×0.76)}＝3.508
z^2 ⇒　χ自乗　自由度１

上記は、二項分布より算出したものであるが、マトリックスのものともぴったり合うか？
次にほかの資料で試してみたいと思います。

二項分布の公式　z^2 =(a-npr)^2 / (npqrs)
マトリックス公式　χ^2={n(eh－fg)^2} / abcd

新薬と旧薬と薬効比較　計算例　…　鈴木義一郎著　統計分析のいざない

	治癒	変化無	合計	
新薬	304	104	408	p=408 / 927
旧薬	353	166	519	q=519 / 927
合計	657	270	927	

r=652 / 927　s=270 / 927

z^2 =(a-npr)^2 / (npqrs)={304-927×(408 / 927)×(657 / 927)}^{2}
÷{927×(408 / 927)×(519 / 927)×(657 / 927)×(270 / 927)}
＝4.667　…(1)

χ^2={n(eh－fg)^2} / abcd ={927×(304×166-104×353)}^{2}
÷408×519×657×270
=4.667　…(2)

補足　旧薬は、比率が決まっているものとしての適合性検定
二項分布からの検定は経過が同じなので省略します。

	治癒	無変化	合計
新薬	304	104	408
旧薬	353	166	519

合計が　新薬と同じように配分を変えます。

治癒	無変化	合計
353÷519×408	166÷519×408	519÷519×408
≒277.5	≒130.5	≒408

適合性の検定
(304－277.5)^2 /277.5 + (104－130.5)^2 /130.5　7.91…(3)

結果は　(1)4.667 ＝　(2)4.667　≦　(3) 7.91
必ず大きくなります。

参照　鈴木義一郎著　現在統計学小事典　統計分析のいざない
石川幹人著、体感する統計解析