

カイ二乗検定とオッズ比 : Odds ratio (OR)

カイ二乗検定

基本は t 検定と似ています。t 検定では 2 つの集団が同じであると仮定するところからはじまっています。カイ二乗検定はしばしば Yes/no type の表データを検討する際に便利であり、まず観察された値(O: observed)が期待された値(E: expected)と同じであると仮説をたてます。そして個々の観察された値が期待された値からどれくらい隔たっているかを検討し、大きく隔たっている場合には最初の仮説を棄却して有意差ありという結論に至ります。

$[O - E] / \text{variance}$ O: observed, E: expected

$$\text{chi-squared} = (O_i - E_i)^2 / E_i$$

臨床研究においてはしばしば以下の表が用いられます。

	暴露あり	暴露なし	合計
疾患発生あり	a	b	a + b
疾患発生なし	c	d	c + d
合計	a + c	b + d	n

a, b c d は実際観察された値です。もし暴露と疾患発生の上に相関関係が存在しなければ各マスの期待値は下記のような表になります。

	暴露あり	暴露なし	合計
疾患発生あり	$(a+b)(a+c)/n$	$(a+b)(b+d)/n$	a + b
疾患発生なし	$(c+d)(a+c)/n$	$(c+d)(b+d)/n$	c + d
合計	a + c	b + d	n

4 つのマスのそれぞれが観察値と期待値の間でどれくらい異なるかを上記公式にあてはめ chi-square (χ^2) を算出します。t 検定では Z の値から p を割り出したように、 χ^2 の値から p を割り出します。自由度は (行 - 1) × (列 - 1) であり、上のような 2 × 2 の表においては自由度は 1 となります。2 × 2 の表では $\chi^2 > 3.8$ の場合有意であると覚えておくと便利です。

「コーヒーを頻繁に飲む人には不整脈を自覚することが多いか？」を調べたいと思います。下の表の結果を得ました。まず H_0 と H_A を定義してください。

	コーヒー - 非常飲者	コーヒー - 常飲者	合計
不整脈自覚あり	218	17	235
不整脈自覚なし	428	130	558
合計	646	147	793

H_0 : コーヒーを頻繁に飲む人も飲まない人も不整脈自覚に差はない。

H_A : コーヒーを頻繁に飲む人も飲まない人も不整脈自覚に差がある。

それでは、もしも差がなかったとするとコーヒー常飲者、非常飲者に何人ずつ不整脈自覚症状のある人あるいは無い人の数が期待されますか？下の表の空欄を埋めてください。

	コーヒー - 非常飲者	コーヒー - 常飲者	合計
不整脈自覚あり			235
不整脈自覚なし			558
合計	646	147	793

例えば不整脈自覚症状ありの人は $(147 \times 235)/793$ です。パターンを覚えると公式を覚える必要はまったくありません。

	コーヒー - 非常飲者	コーヒー - 常飲者	合計
不整脈自覚あり	191.4	43.6	235
不整脈自覚なし	454.6	103.4	558
合計	646	147	793

それぞれのマスで期待される数値と観察された数値の差をとり 2 乗して更に期待された数で割ります。そして全てを合計すると chi-square となります。T-test で行なったように chi-square distribution に相当する表と比較します。2 x 2 table の場合自由度は 1 であり、有意性は 3.8 より大きくて証明されます。t 検定のように one-sided とか two-sided はありません。多くの表が 2 x 2 table なので覚えてしまいましょう。

この例では

$$\text{chi}^2 = (17 - 43.6)^2/43.6 + (130 - 103.4)^2/103.4 + (218 - 191.4)^2/191.4 + (428 - 454.6)^2/454.6 = 16.23 + 6.84 + 3.70 + 1.56 = 28.33 \gg 3.8$$

で有意であり、 H_0 を棄却し「コーヒー常飲者とそうでないものを比較したとき不整脈自覚において差がある」という結論となります。しかし我々の知りたいのはどちらが多く、どれくらい違うのかです。これを知るにはオッズ比を用います。その前に chi-square 検定は 2 x 2 table より大きいものでも使えることを示したいと思います。

. cci 218 17 428 130

Proportion

	Exposed	Unexposed	Total	Exposed
Cases	218	17	235	0.9277
Controls	428	130	558	0.7670
Total	646	147	793	0.8146
	Point estimate		[95% Conf. Interval]	
Odds ratio	3.894997		2.30169	6.588306 (Cornfield)
Attr. frac. ex.	.7432604		.5655366	.8482159 (Cornfield)
Attr. frac. pop	.6894926			
chi2(1) = 28.26 Pr>chi2 = 0.0000				

オッズ比は後から説明しますが、上の表からは手で計算したのと同様に有意差ありの結論を導くことができます。

臨床診断と病理診断の一致率が A 病院と B 病院で違うかどうか検討しました。その一致度を完全一致、部分一致、不一致と 3 つのカテゴリーに分けました。病理診断を絶対的 golden standard と考えた場合、その一致率をみるとその病院の臨床診断の実力を知る材料になるかもしれません。A 病院と B 病院を比較してみてください。

	完全一致	不完全一致	不一致	合計
A 病院	157	18	54	229
B 病院	268	44	34	346
合計	425	62	88	575

H_0 : A病院とB病院において臨床診断と病理診断の一致率が同じである。

H_A : A病院とB病院において臨床診断と病理診断の一致率が同じでない。

我々は H_0 が正しいとして、先ほどと同じように期待される数を書きます。

	完全一致	不完全一致	不一致	合計
A 病院	169.3	24.7	35.0	229
B 病院	255.7	37.3	53.0	346
合計	425	62	88	575

$$\text{Chi-squared} = (157 - 169.3)^2/169.3 + (18 - 24.7)^2/24.7 + (54 - 35.0)^2/35.0 + (268 - 255.7)^2/255.7 + (44 - 37.3)^2/37.3 + (34 - 53.0)^2/53.0 = 0.89 + 1.82 + 10.31 + 0.59 + 6.81 + 1.20 = 21.62$$

$$\text{自由度} = (2 - 1)(3 - 1) = 2$$

これを表などに照らし合わせると $p < 0.001$ となり有意です。つまり H_0 を棄却して H_A を採用し 2 つの病院は異なる結論します。

オッズ比 : Odds ratio (OR)

カイ二乗検定は2つの変数の間に相関があるかどうかを検定するものですが、相関の程度や方向を知ることができません。これはt検定と同じです。この点を解決するためのものがオッズ比です。ある事象が確率pで発生するとすると、この事象が発生しない確率は1 - pです。オッズとは $p / (1-p)$ であらわされるものです。先の表で暴露された集団において病気の発生する確率は $a / (a + c)$ でした。一方暴露された集団において病気の発生しない確率は $c / (a + c)$ です。よって暴露された集団のオッズは $[a/(a+c)] / [c/(a+c)] = a/c$ となります。また暴露されなかった集団についてのオッズは同様に b/c となります。暴露されなかった集団のオッズに対する暴露された集団のオッズの比をとるとオッズ比は ad/bc です。これを計算により因子に暴露されたことによって疾患発生率が何倍に増えたと言うことができ、カイ二乗検定より情報量が多くなります。

オッズ比はケースコントロール研究で使用される場合にはバイアスを生じません。一方、コホート研究他においては相対比が好んで用いられます。相対比は発生率が低い場合においてオッズ比と近似しますが、数が少ない場合には相対比の方が正確だからです。

臨床研究において一定の症例数の結果を用いて全体を推論するわけですが、その予測される値が95%の確率で収まる範囲を95%信頼区間(CI)と呼びます。

オッズ比が1であるということは暴露あるいは非暴露における疾患発生が同一であることを意味し有意差なしと解釈できます。すなわち95%信頼区間が1.0を含んでいれば有意差なしといえ、また信頼区間の幅が広いほど有意差の程度が低いこととなります。臨床研究においてはp値より95%信頼区間が好んで用いられる所以です。

以下は毎日一定以上の運動をする人とそうでない人で冠動脈疾患の発生頻度を比較してみたものです。

	運動量多い	運動量少ない	合計
冠動脈疾患あり	158	351	509
冠動脈疾患なし	4592	6106	10698
合計	4750	6457	11207

OR とその95%CIを計算してみてください。結果に対して解釈を加えてください。

. cci 158 351 4529 6106

		Proportion	
Exposed	Unexposed	Total	Exposed
-----+-----+-----			

Cases	158	351	509	0.3104
Controls	4529	6106	10635	0.4259
-----+-----+-----				
Total	4687	6457	11144	0.4206
	Point estimate		[95% Conf. Interval]	
-----+-----+-----				
Odds ratio	.6068823	.5011657	.7349015	(Cornfield)
Prev. frac. ex.	.3931177	.2650985	.4988343	(Cornfield)
Prev. frac. pop	.1674123			
-----+-----+-----				
	chi2(1) = 26.57 Pr>chi2 = 0.0000			

よって OR は 0.6、95%CI は 0.5 から 0.7 です。95%CI は 1 を含まないので統計学的には有意です。「一定以上の運動は冠動脈疾患の発生を 60%にまで予防する」と結論できます。

2 x 2 table は臨床研究をする際の基本です。我々の身の回りには臨床研究をする題材がいくらかでも転がっています。しかし以上は基本的な話で、例えばコーヒー飲料といっ
ても性別、年齢、喫煙や体重、職業などとの関係を調べなくてはなりません。しかし、本書を読み終わることにはこの点についても習熟していただけたと思います。

Doll&Hill は男性医師を対象に何らかの疾患が発生しないかどうか4年5ヶ月経過観察しました。そして喫煙者、非喫煙者で疾患の発生に違いがあるかどうか比べています。

	喫煙者	非喫煙者	合計
死亡	1582	166	1748
健康	27116	5630	32746
合計	28698	5796	34494

喫煙をしていると死亡するリスクはどれくらいあがるのでしょうか？

上のような研究スタイルは cohort study と呼ばれます。Exposure base で検討するものを cohort, outcome base で考えるものを case-control study と呼びます。上では喫煙という exposure を軸にものを考えているので cohort study です。Cohort study の場合リスクの差や比で比べるのが一般的ですが、分母が大きい場合には OR でも問題ありません。

```
. input Y X count
```

```

      Y      X      count
1.  1  1  1582
2.  1  0   166
3.  0  1 27116
4.  0  0  5630
5. end

```

```
. cc Y X [freq=count], woolf
```

```

      | X              |          Proportion
      | Exposed  Unexposed | Total   Exposed
-----+-----+-----+-----+-----+
Cases |      1582      166 |    1748   0.9050
Controls |    27116     5630 |   32746   0.8281
-----+-----+-----+-----+
Total |    28698     5796 |   34494   0.8320
      |
      | Point estimate | [95% Conf. Interval]
      |-----+-----+-----+-----+
Odds ratio |      1.978705 | 1.681997  2.327754 (Woolf)

```

```
Attr. frac. ex. |      .4946191      | .4054687  .5704013 (Woolf)
Attr. frac. pop |      .4476472      |
```

```
+-----
```

```
chi2(1) = 70.31 Pr>chi2 = 0.0000
```

男性医師において、喫煙者の死亡率は非喫煙者の 2 倍近いことになります。