カイ二乗検定とオッズ比：Odds ratio (OR)

カイ二乗検定

- 基本はt検定と似ている。t検定では２つの集団が同じであると仮定すところからはじまっている。カイ二乗検定はしばしばYes/no type の表データを検討す際に便利であり、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」と仮説をたてる。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る。
- [O – E]/variance O: observed, E: expected
- chi-squared = $\Sigma(O_i-E_i)2/E_i$
- 臨床研究においてはしばしば表１が用いられる。
- a, b c d は実際観察された値である。もし暴露と疾患発生の間に相関関係が存在しなければ各期待値は下記のような表２になる。
- ４つのマス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 (chi2)を算出す。t検定ではtの値からpを割り出したように、chi2の値からｐ（面積）を割り出す。自由度は（行—1）χ（列—1）であり、上のような２χ２の表において、自由度は１となる。2χ2の表ではchi2＞3.8の場合有意であると覚えておくと便利である。

表1　暴露ありとなし、疾患発生ありとなしを２ｘ２の表として示す。

	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

表2　もしも暴露因子の有無に関わらず疾患が発生するとしたら、表のようなデータであることが期待される（合計値から比例計算で算出することができる）。

	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

表3　コーヒー常飲者と不整脈自覚の関係

	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

- 「コーヒーを頻繁に飲む人には不整脈を自覚すことが多いか？」を調べたところ表3の結果を得た。まず帰無仮説：H0と対立仮説：HAを定義すると以下のようになる。
- H0: 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。
- HA: 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るか？例えば不整脈自覚症状ありの人は(147 x 235)/793 である。パターンを覚えると公式を覚える必要はまったくない

表4　コーヒー常飲が不整脈自覚に影響しないと仮定すると、表内の値であることが期待される。

	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

- この例では
- chi^square = (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 >> 3.8
- で有意である。すなわち、H0 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」という結論となる。
- 　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病院の臨床診断の実力を知る材料になるかもしれない。A病院とB病院を比較してみる（表5）。
- H0: 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。
- HA: 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。

表5　臨床診断と病理診断の一致率をA病院とB病院で比較する。

	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

- 我々はH0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る（表6）。
- 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
- 自由度は(2 – 1)(3 – 1) = 2
- これを表などに照らし合わせると$p < 0.001$となり有意である。つまりH0を棄却してHAを採用し「2つの病院における臨床診断と病理診断の一致率は異なる」と結論できる。これは、例えばランダム化していない比較試験において、標準治療Ａのがんステージと新規治療薬Ｂのがんステージが同じか異なるかを検定する際にもしばしば用いられる。しかし、私たちは決まって「どっちの病院がどれくらいいいんだ？」とたずねるであろう。

表6　臨床診断と病理診断の一致率をA病院とB病院で同じであると想定したとき、表内の値であることが期待される。

	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

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ためのものであったが、相関の程度や方向を知ることができない。これはt検定と同じである。この点を解決するためのものがオッズ比である。ある事象が確率pで発生すとすると、この事象が発生しない確率は1－pである。オッズとはp /(1-p) であらわされるものである。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) だった。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) である。よって暴露された集団のオッズは[a/(a+c)] / [c/(a+c)] = a/c となる。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 となる。暴露されなかった集団のオッズに対す暴露された集団のオッズの比をとるとオッズ比は ad/bc である。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定量的に表現することができるため、カイ二乗検定より情報量が多くなる。

- 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)と呼ぶ。Standard error: se[ln(OR)] = square root of [1/a + 1/b + 1/c + 1/d]
- 95% CI = eln(OR) - 1.96se[ln(OR)], e ln(OR) + 1.96se[ln(OR)]
- 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のリスクが同一であることを意味し有意差なしと解釈できる。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、1を含まない場合は有意差ありと判断できる。信頼区間の幅が狭いほど真の値に近いことを指す。つまり、p値は両者が異なるか否かの判定情報しか与えてくれず、95%信頼区間は真の結果にどれくらい近いのかという情報まで提供してくれる。そのため、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。
- 具体的には以下の通りである。

- 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ある（表7）。
- OR = [382 x 2,745] / [229 x 2,492] = 1.72,   ln(OR) = 0.542
- Se[ln(OR)] = square root of [1/358+1/229+1/2492+1/2745] = 0.089
- 95% CI = e0.542 – 1.96 x 0.089, e0.542 + 1.96 x 0.089 = e0.368, e0.716 = (1.44 〜 2.05)
- よってORは1.7、95%CIは1.4から2.1である。95%CIは1を含まないので統計学的には有意である。「運動量が少ないと冠動脈疾患の発生リスクが1.7倍になる」と結論できる。あるいは、「バイアス、交絡因子が存在しないと仮定して、運動量が少ないと冠動脈疾患の発生リスクは95%の確率で1.4倍から2.1倍になる」とも表現できる。
- variance, 95%信頼区間の算出方法は、クローズド・コホート研究、オープン・コホート研究、ケース・コントロール研究によって異なるため、数式資料を参考にしていただきたい。variance の算出に関しては、研究者によって多少異なっているようである。先に示したカイ二乗検定の際のvariance と数式資料に示したvariance も異なっている。前者は生物統計学教授の主張する方法であり、後者は疫学教授の方法であったが、個々に確認したところ、自分が正しいという主張を曲げなかったので、敢えて２つの方法を示すことにした。

表7　運動量の多い少ないと、冠動脈疾患発生との関係

	運動量少ない	運動量多い	合計
冠動脈疾患あり	358	229	587
冠動脈疾患なし	2,492	2,745	5,237
合計	2,850	2,974	5,824