

医学統計勉強会

第5回 比率と分割表

帝京大学臨床研究センター（TARC）・帝京大学大学院公衆衛生学研究科 共催

帝京大学大学院公衆衛生学研究科

宮田 敏

カテゴリデータの要約

度数分布：
水準 カテゴリの種類
度数 各水準のデータ数
割合（相対度数） 度数/サンプル数 (%)

カテゴリデータを提示する場合は、必ず**度数 (%)** の形で提示する。

変数によっては欠測値の数が異なることがあるため、度数あるいは相対度数だけでは、データ全体に占める各水準の相対的な大きさをうまく表現できない。

比率と分割表

疾患の発症率など、物事の頻度 (frequency) を議論する際、以下の三つの概念を使い分ける。

- **比 (ratio)** : A, B ($\neq 0$) が存在するとき, A/B を比という. A と B は互いを含まない.
 - 例 : 性比. $BMI = \text{体重} / \text{身長}^2$
- **割合 (proportion)** : A, B ($\neq 0$) が存在し, 分子 A が分母 B に含まれるとき, A/B を割合という. $0 \leq \text{割合} \leq 1$.
比率 \doteq 割合
- **率 (rate)** : 単位時間あたりのイベントの発生割合.
 - 率 = イベント発生件数 / 延べ観察時間

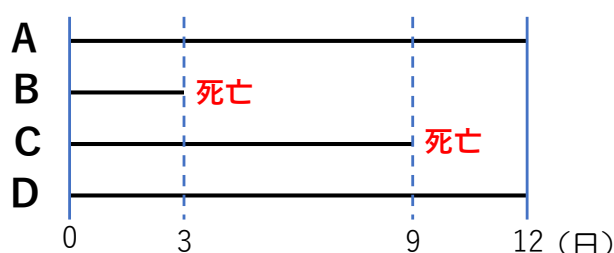
2023/10/25

医学統計勉強会 第5回 比率と分割表

3

割合と率 :

例 : 4匹のマウスの生存時間を観察した.



死亡割合 : イベント数 / 標本数 = $2/4 = 0.5$

死亡率 (人日法) : イベント数 / 延べ観察時間
 $= 2 / (12 + 3 + 9 + 12) = 2/36 = 0.0556$

「率」は、時間単位によって値が変わることに注意.

2023/10/25

医学統計勉強会 第5回 比率と分割表

4

母比率の推定と検定

X_1, \dots, X_n : イベントの有無

$$X_i = \begin{cases} 1: \text{イベント有り} \\ 0: \text{イベントなし} \end{cases}, P(X_i = 1) = p \quad \text{母比率}$$

推定量 : $\hat{p} = X/n$
期待値 : $E(\hat{p}) = p$
分散 : $V(\hat{p}) = \hat{p}(1 - \hat{p})/n$
標準誤差 : $s.e.(\hat{p}) = (\hat{p}(1 - \hat{p})/n)^{1/2}$

比率の推定量 \hat{p} が、近似的に正規分布に従うことを用いて、信頼区間や仮説検定を構成する。

母比率の**信頼区間** (信頼水準 $100(1 - \alpha)\%$)

$$\left(\hat{p} - z_{\alpha/2} \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}}, \hat{p} + z_{\alpha/2} \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} \right)$$

ただし, $z_{\alpha/2}$: 標準正規分布の上側 $100(\alpha/2)\%$ 点

母比率の**仮説検定** (有意水準 α)

$$H_0 : p = p_0$$

$$H_1 : p \neq p_0$$

棄却域 :

$$Z = \frac{|\hat{p} - p_0| - 1/2n}{\sqrt{\frac{p_0(1 - p_0)}{n}}} > z_{\alpha/2}$$

Clopper-Pearson正確信頼区間

前ページの信頼区間と仮説検定はよく用いられるが、標本数が十分大きいときに適用される近似を用いている。（通常 $np > 5$ 程度）また、 $\hat{p} = 0$ あるいは 1 となったとき、信頼区間が構成出来ない。

Clopper-Pearson正確信頼区間

$$\left(\frac{x}{(n-x+1)F_1 + x}, \frac{(x+1)F_2}{(x+1)F_2 + n-x} \right)$$

ただし、

$F_1 = F(\alpha/2, 2(n-x+1), 2x)$: $df1=2(n-x+1)$, $df2=2x$ の F 分布の上側 $100(\alpha/2)\%$ 点。
 $F_2 = F(\alpha/2, 2(x+1), 2(n-x))$: $df1=2(x+1)$, $df2=2(n-x)$ の F 分布の上側 $100(\alpha/2)\%$ 点。

小標本の下での比率の検定

標本数が少ない、あるいは比率の推定値が 0, 1 に極端に近いときは、信頼区間の場合と同様に仮説検定でも正規近似に基づく方法は使えない。二項分布に基づく正確検定を行う必要がある。

例： $x = 0, n = 10, \hat{p} = 0/10 = 0$

```
> binom.test(0, 10)
```

```
Exact binomial test
```

```
data: 0 and 10
```

```
number of successes = 0, number of trials = 10, p-value = 0.001953
```

```
alternative hypothesis: true probability of success is not equal to 0.5
```

```
95 percent confidence interval:  
0.0000000 0.3084971
```

```
sample estimates:
```

```
probability of success  
0
```

イベント率（人年法）の信頼区間と検定

イベント率 = イベント発生件数 / 延べ観察時間

```
> ##Examples from Rothman 1998, p. 238
> bc <- c(Unexposed = 15, Exposed = 41)
> pyears <- c(Unexposed = 19017, Exposed = 28010)
> dd <- matrix(c(41,15,28010,19017),2,2)
> dimnames(dd) <- list(Exposure=c("Yes","No"), Outcome=c("BC","PYears"))
>
> rateratio.wald(dd)
$data
  Outcome
Exposure BC PYears
Yes      41 28010
No       15 19017
Total   56 47027

$measure
  rate ratio with 95% C.I.
Exposure estimate lower upper
Yes 1.0000000 NA NA
No 0.5388632 0.2982792 0.9734956

$p.value
  two-sided
Exposure midp.exact wald
Yes NA NA
No 0.03545742 0.03736289
```

イベント率の比と95%信頼区間

イベント率の比較の検定の p-value
Midp-exact法とWald法の二つがある

2023/10/25

医学統計勉強会 第5回 比率と分割表

9

カテゴリデータの要約

分割表 (Contingency table), クロス集計表 (cross tabulation table) : 2種類のカテゴリデータの水準の組み合わせごとに、度数を求めた表。

(例)

2×2分割表

	疾患有り	疾患なし
要因陽性	a	b
要因陰性	c	d

2023/10/25

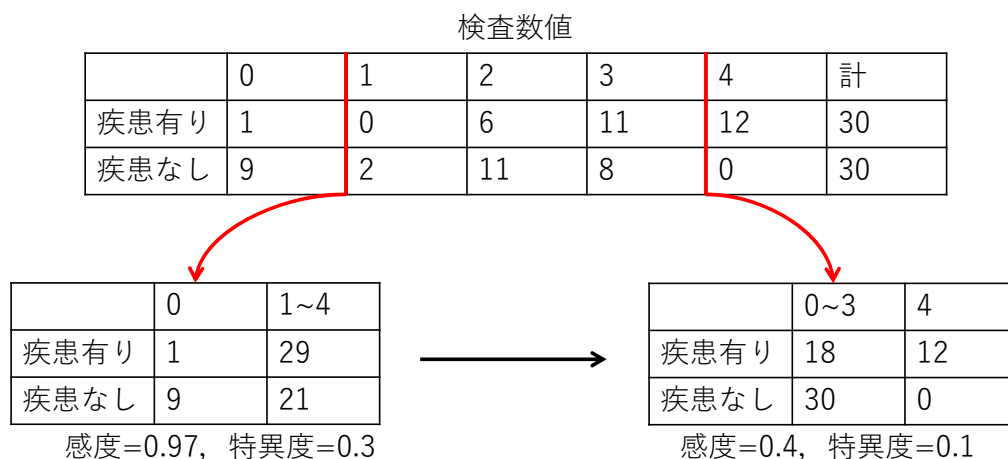
医学統計勉強会 第5回 比率と分割表

10

要因の有無により疾患を予測するとき、その精度を測るために以下の概念を定義する。

- **正答率 (accuracy rate)** : $(a+d)/(a+b+c+d)$
- **感度 (sensitivity)** : 疾患有りのうち、陽性と判断される割合. $a/(a+c)$
疾患有りなのに陰性と判断される割合を**偽陰性率 (false negative rate)**と呼ぶ。
- **特異度 (specificity)** : 疾患なしのうち、陰性と判断される割合. $d/(b+d)$
疾患なしなのに陽性と判断される割合を**偽陽性率 (false positive rate)**と呼ぶ。
- **陽性的中率 (positive predictive rate)** : 陽性のうち疾患有りの割合.
 $a/(a+b)$
- **陰性的中率 (negative predictive rate)** : 陰性のうち疾患なしの割合.
 $d/(c+d)$

検査数値が、三段階以上もしくは実数値をとる場合、カットポイントのとり場所により、複数の 2×2 分割表にデータをまとめることが出来る。検査数値の順序に従って分割表を作り、感度、特異度が求める。

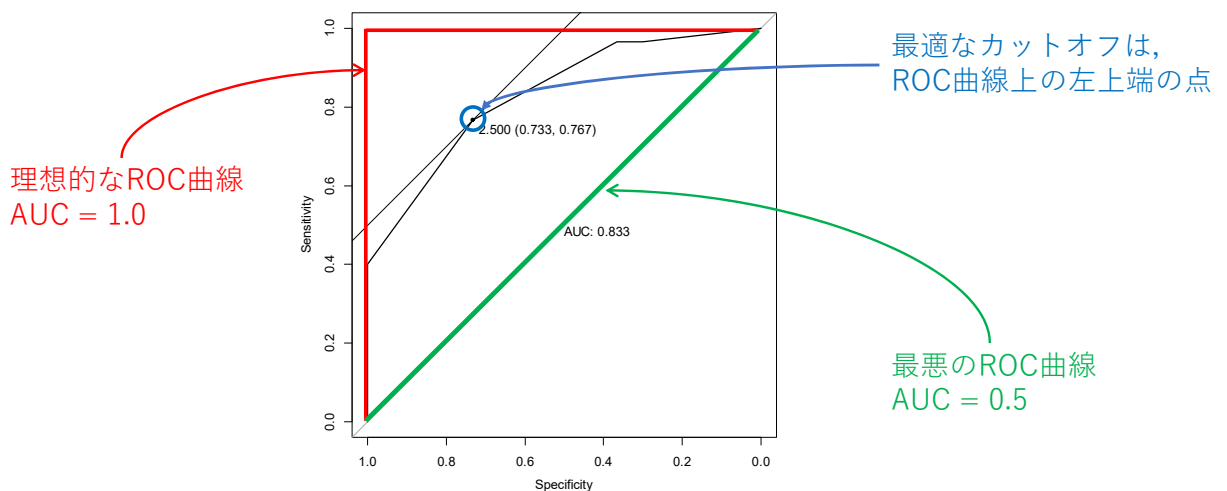


柳川 堯, 木 由布子 (著) 「バイオ統計の基礎—医薬統計入門」 近代科学社 (2010/02)

前項のように、検査数値の可能なカットポイントの順に、一連の分割表を作り感度、特異度が計算出来る。縦軸に感度、横軸に(1-特異度)をとったグラフを**ROC曲線** (Receiver Operating Characteristic curve) と呼び、ROC曲線の下面積を **AUC** (Area Under Curve) あるいは**C統計量** (C statistic) と呼ぶ。

感度は1に近く、特異度は1に近い=(1-特異度)は0に近いことが望ましいわけだから、AUCは大きいほうが望ましい。AUCは $0.5 \leq AUC \leq 1.0$ となることが示される。

- 感度、特異度共に大きい(1.0に近い)方が望ましい。
- ROC曲線の左上端の点に対応する検査数値が、最適な分岐点。
- ロジスティック回帰の適合度の評価にも用いる。



分割表の検定（独立性の検定）

	6カ月以内死亡	6カ月以上生存
牛乳抗体陽性	29	80
牛乳抗体陰性	10	94

$$\hat{p}_1 = 29/(29 + 80) = 0.266$$

$$\hat{p}_2 = 10/(10 + 94) = 0.0962$$

帰無仮説 $H_0: p_1 = p_2$ 母比率が一定.

対立仮説 $H_1: p_1 \neq p_2$ 母比率が異なる.

もし、二つの変数に**関連がなければ**、グループによらず（=v陽性でも陰性でも）、母比率（=6カ月以内に死亡する確率）は**一定のはず** ⇒
帰無仮説 H_0

関連があれば、母比率が**異なる** ⇒ **対立仮説 H_1**

2×2分割表に限り、**片側仮説**が可能. $H_1: p_1 > (<) p_2$

分割表の検定（独立性の検定） 続き

χ^2 検定 (chi-squared test) : サンプル数が多いとき、検定統計量の分布が χ^2 分布で近似されることを利用した検定. 分割表の度数は、最低 **5** は必要とされる. 必ず **Yatesの連続補正 (Yates's continuity correction)** を行う.

Fisherの直接法 (Fisher's exact test) : サンプル数によらず、正確な p -value を計算できる検定.

Fisherの直接法が第一選択. 分割表が大きすぎてFisher検定が実行できない場合に限り、Yatesの連続補正をした χ^2 検定を選択する.

例：高血圧を合併した安定期慢性心不全患者に対するオルメサルタンの有効性に関する薬物介入臨床試験（SUPPORT試験）

	Olmesartan (+)	Olmesartan (-)
β-Blocker (+)	405	416
β-Blocker (-)	173	152

Fisherの直接法： $p = 0.2388$

χ^2 検定 (Yatesの連続補正あり)： $p = 0.2606$

χ^2 検定 (Yatesの連続補正なし)： $p = 0.2339$

分割表の検定の結果の提示

例：

各群の総数を明示する

Name	Statin=0 (n=2933)	Statin=1 (n=1803)	p-value
HxSmoke01	1243 (42.4%)	864 (47.9%)	0
HxHFad	1616 (55.1%)	853 (47.3%)	0
HxHT	2200 (75%)	1472 (81.6%)	0
indx.DM	866 (29.5%)	786 (43.6%)	0
Cancer	380 (13%)	174 (9.7%)	0.001

カウントとパーセントを両方提示する。

検定のp値（Fisher's exact testが基本）を、必ず明示する。

対応のあるデータに対する比率の検定

同じサンプルから、対照群と処理群のデータを対応のあるデータ（1対1マッチングデータ）として取られる場合の有効率の比較を考える。

	control	treatment
Item 1	+	+
Item 2	+	-
Item 3	-	-
Item n	+	-

→

	Treatment (+)	Treatment (-)
Control (+)	a	b
Control (-)	c	d

(i, j) セルの確率を P_{ij} とすると、分割表全体の確率は以下のようなになる。

	Treatment (+)	Treatment (-)
Control (+)	P_{11}	P_{12}
Control (-)	P_{21}	P_{22}

対照群の有効率は $P_{\text{cnt}} = P_{11} + P_{12}$ 処置群の有効率は $P_{\text{trt}} = P_{11} + P_{21}$ であるから、検定する仮説は以下の通り。

$$H_0 : P_{\text{cnt}} = P_{\text{trt}} \text{ vs. } H_1 : P_{\text{cnt}} \neq P_{\text{trt}} \Leftrightarrow H_0 : P_{12} = P_{21} \text{ vs. } H_1 : P_{12} \neq P_{21}$$

すなわち、(1, 1), (2, 2) セルは無視してよい。 $\theta = P_{12} / (P_{12} + P_{21})$ とすれば、

$$H_0 : \theta = 1/2 \text{ vs. } H_1 : \theta \neq 1/2$$

対応のあるデータに対する比率の検定

χ^2 分布を用いてこの検定を行う方法を、**McNemar検定**と呼ぶ。
McNemar検定は近似検定であり、連続補正が行われる。**二項分布**を用いて、正確な p 値を求める検定も可能である（こちらが望ましい）。

例：2004年と2008年のアメリカ大統領選挙における一般住民調査（男性）における投票行動

	2008年 民主党	2008年 共和党
2004年 民主党	175	16
2004年 共和党	54	204

Agresti, A. 2013. *Categorical Data Analysis*, 3rd ed. Hoboken, NJ: Wiley. P. 414, Table 11.1

```
> tbl2
      2008 election
2004 election Democrat Republican
      Democrat      175         16
      Republican    54         188
> binom.test(16, (16 + 54))# 二項分布による正確検定

Exact binomial test

data: 16 and (16 + 54)
number of successes = 16, number of trials = 70, p-value = 5.854e-06
alternative hypothesis: true probability of success is not equal to 0.5
95 percent confidence interval:
 0.136657 0.344475
sample estimates:
probability of success
0.2285714

> mcnemar.test(tbl2)# McNemar検定 (連続補正あり)

McNemar's Chi-squared test with continuity correction

data: tbl2
McNemar's chi-squared = 19.5571, df = 1, p-value = 9.764e-06

> mcnemar.test(tbl2)# McNemar検定 (連続補正なし)

McNemar's Chi-squared test with continuity correction

data: tbl2
McNemar's chi-squared = 19.5571, df = 1, p-value = 9.764e-06
```

シンプソンのパラドックス

分割表で考える2つの要因の双方に影響を与える因子を、**交絡因子 (confounding factor)** と呼ぶ。交絡因子を無視して検定を行うと、本来存在した関連が見えなくなったり、その逆が起こることがある。

いま、A, B2つの要因の間の分割表を考えているとする。第3の要因Zの値によって、標本が二つの層に分けられたとする。それぞれの分割表は以下の通りとする。

Z1			Z2		
	B	Not B		B	Not B
A	80	16	A	50	452
Not A	160	160	Not A	10	452

Odds ratio=(80/16)/(160/160)=**5** Odds ratio=(50/452)/(10/452)=**5**

2023/10/25

医学統計勉強会 第5回 比率と分割表

23

交絡因子Zを無視して、二つの分割表を統合してしまうと、

	B	Not B
A	130	468
Not A	170	612

$$\text{Odds ratio} = ((80+50)/(16+452))/((160+10)/(160+452)) = \mathbf{1}$$

この分割表のオッズ比を計算すると、**オッズ比=1**となる。交絡因子Zで層別しておけばオッズ比=5、すなわち要因A有りのほうが5倍オッズが高かったことが分かるが、交絡因子を無視したためにリスクが全く見えなくなっている。

(参照：柳川堯「離散多変量データの解析」(1986) 共立出版)

⇒ **Mantel-Haenszel検定**

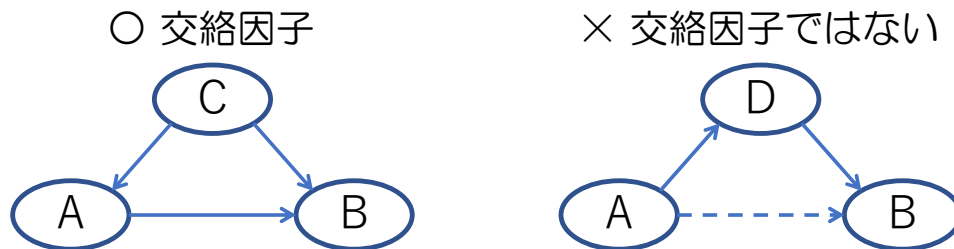
2023/10/25

医学統計勉強会 第5回 比率と分割表

24

分割表で考える A, B 二つの要因の, 双方に影響を与える第3の要因 C が交絡因子であった。

要因 A が原因, 要因 B が結果である時, 因果関係の連鎖の途中にある要因 D は, 交絡因子にはなりえないことに注意する。



何が交絡因子になるかは, 統計学だけでは答えられない。研究対象に関する科学的知識が必要。

交絡因子が判明した場合, 層ごとに分割表を作り, それらを統合して検定を行う。この検定を, **Mantel-Haenszel検定**という。

```
> tbl <- array(c(80,160,16,160, 50, 10, 452, 452), dim=c(2,2,2))
> tbl
, , 1
     [,1] [,2]
[1,]  80   16
[2,] 160  160
, , 2
     [,1] [,2]
[1,]  50  452
[2,]  10  452
> mantelhaen.test(tbl)

Mantel-Haenszel chi-squared test with continuity correction

data:  tbl
Mantel-Haenszel X-squared = 57.2109, df = 1, p-value = 3.915e-14
alternative hypothesis: true common odds ratio is not equal to 1
95 percent confidence interval:
 3.205444 7.799232
sample estimates:
common odds ratio
5
```

用量・反応関係の検定

片方の要因の水準が何らかの処置の用量に対応し、そこに増大若しくは減少の傾向があるとする。

	用量			
	d0	d1		dk
反応有り	x0	x1		xk
反応なし	n0 - x0	n1 - x1		nk - xk
	n0	n1		nk

	用量			
	d0	d1		dk
反応有り	p0	p1		pk
反応なし	1 - p0	1 - p1		1 - pk
	1	1		1

ここで用量は $d_0 < d_1 < \dots < d_k$ 若しくは不等号が逆であるとする。このとき、反応有りの確率にも $p_0 \leq p_1 \leq \dots \leq p_k$ あるいはその逆の傾向 (trend) があるかを検定したい。

用量・反応関係の検定

Cochran-Armitage Trend Test:

$$H_0 : p_0 = p_1 = \dots = p_k \text{ vs. } H_1 : p_0 \leq p_1 \leq \dots \leq p_k$$

本質的には、 $p_i = h(a + bd_i)$ の形の **ノンパラメトリックな回帰モデル** を検定することになる。

例：妊婦の飲酒量と奇形児の発生の関係

	アルコール消費				
	0	< 1	1-2	3-5	>= 6
奇形なし	17066	14464	788	126	37
奇形あり	48	38	5	1	1

Agresti, A. 2013. *Categorical Data Analysis*, 3rd ed. Hoboken, NJ: Wiley. P. 89, Table 3.8

```

> prop.trend.test(event, total, score=c(0, 0.5, 1.5, 4.0, 7.0))# Cochran-Armitage Trend Test
      Chi-squared Test for Trend in Proportions
data:  event out of total ,
      using scores: 0 0.5 1.5 4 7
X-squared = 6.5701, df = 1, p-value = 0.01037
> prop.test(event, total)# (傾向無し)の独立性の検定
      5-sample test for equality of proportions without continuity correction
data:  event out of total
X-squared = 12.0821, df = 4, p-value = 0.01675
alternative hypothesis: two.sided
sample estimates:
      prop 1      prop 2      prop 3      prop 4      prop 5
0.002804721 0.002620328 0.006305170 0.007874016 0.026315789

```

$p_i = h(a + bd_i)$ において $H_0: b = 0$ を検定する際、本例のように用量の水準が実数値の範囲で指定される場合は、範囲の中点を d とする。単に序数である時は、1, 2, 3, ... などとする。

Take Home Message

1. 比と割合と率
2. 母比率の推定と検定
 - Clopper-Pearsonの正確信頼区間
3. 分割表の推定
 - 感度, 特異度, その他
 - ROC曲線
4. 分割表の検定
 - 対応のあるデータの分割表に関する検定
5. シンプソンのパラドックス
 - 交絡因子
 - Mantel-Haenszel検定
6. 用量・反応関係の検定