

2つの平均値の差の検定

60161053 前田欣耶

資料

- 医薬研究者の視点からみた道具としての統計学
- 9章「2つの平均値の差の検定」



目次

1、t検定とは？

2、対応のないt検定

3、対応のあるt検定

4、まとめ

1、t検定とは？

ある母集団からの標本を2つ抽出したとき、二つの標本の平均の差が偶然誤差の範囲内にあるかどうかを調べるための検定

2、対応のないt検定

- 二つの標本に対応のないt検定には2種類ある
 - 等分散を仮定した、対応のないt検定
 - 等分散を仮定しない、対応のないt検定

2、対応のないt検定

- 条件

- 1、標本は母集団から無作為抽出される
- 2、母集団は正規分布に従う
- 3、標本が互いに独立している
- 4、二つの母分散は等しい

2、対応のないt検定

a郡	b郡
3.6	4.8
4.3	4.4
3.8	3.7
3.5	4.5
4.7	4.7
3.8	5.0
3.5	4.5
4.8	4.8
4.4	4.1
3.6	
n1 = 10	n2 = 9
$\bar{X}1 = 4.0$	$\bar{X}2 = 4.5$
SD1 = 0.503	SD2 = 0.4

- 等分散の検定(F検定)

- 1、標本分散の比(F値)

$$\begin{aligned}
 F &= SD1^2 / SD2^2 \\
 &= 0.50^2 / 0.40^2 \\
 &= 1.583
 \end{aligned}$$

- 2、自由度を求める

$$\text{自由度} = n-1$$

$$\text{第一自由度} = 10-1 = 9$$

$$\text{第二自由度} = 9-1 = 8$$

- 3、F値に対応する確率(P)が有意水準より小さいか確かめる

$$\begin{aligned}
 P &= \text{FDIST}(1.583, 9, 8) = \\
 &0.264 > 0.05
 \end{aligned}$$

n:標本数 $\bar{X}$:平均 SD:標本分散 有意水準:5%
両側検定

- 等分散が仮定できたので、二つの平均値に差があるかどうかを調べる

a郡	b郡
3.6	4.8
4.3	4.4
3.8	3.7
3.5	4.5
4.7	4.7
3.8	5.0
3.5	4.5
4.8	4.8
4.4	4.1
3.6	
n1 = 10	n2 = 9
$\bar{X}_1 = 4.0$	$\bar{X}_2 = 4.5$
SD1 = 0.503	SD2 = 0.4

- 標準誤差(SEMd)を求める

$$\text{SEMd} = \sqrt{[(n_1-1) \cdot \text{SD}_1^2 + (n_2-1) \cdot \text{SD}_2^2] / (n_1+n_2-2)} \cdot (1/n_1 + 1/n_2)$$

$$= 0.210$$

- t値を求める

$$t = (\bar{X}_1 - \bar{X}_2) / \text{SEMd}$$

$$= -2.378$$

- 自由度(df)を求める

$$\text{df} = n_1 + n_2 - 2$$

$$= 17$$

- t値の絶対値に対応する確率(P)が有意水準より小さいか確かめる

$$P = \text{TDIST}(2.378, 17, 2) = 0.029 < 0.05$$

差があるといえる

n:標本数 $\bar{X}$:平均 SD:標本分散 有意水準:5%
両側検定

2、対応のないt検定

a郡	b郡
3.6	4.8
4.3	4.4
3.8	3.7
3.5	4.5
4.7	4.7
3.8	5.0
3.5	4.5
4.8	4.8
4.4	4.1
3.6	
n1 = 10	n2 = 9
$\bar{X}_1 = 4.0$	$\bar{X}_2 = 4.5$
SD1 = 0.503	SD2 = 0.4

- 二つの平均値の差の信頼区間
二つの平均値の検定に際しては、二つの平均値の差の信頼区間も同時に求めることができる

$$|\bar{X}_1 - \bar{X}_2| - t(\alpha/2, df) * SEM_d < \mu_1 - \mu_2 < |\bar{X}_1 - \bar{X}_2| + t(\alpha/2, df) * SEM_d$$



TINV(α, 自由度)

$$0.056 < \mu_1 - \mu_2 < 0.944$$

二つの平均値の差の95%信頼区間は
(0.056mEq/l, 0.094mEq/l,)

n:標本数 X:平均 SD:標本分散 有意水準:5%
両側検定

3、対応のあるt検定

- 条件

- 1、標本何らかの形で対になっている
- 2、対になった標本は母集団から無作為に抽出されている
- 3、対になった標本の差が近似的に正規分布に従う

3、対応のあるt検定

被験者	投与前	投与後	差(di)
1	3.6	4.8	1.2
2	4.3	4.4	0.1
3	3.8	3.7	-0.1
4	4.5	5.0	0.5
5	4.2	4.9	0.7
6	4.2	5.1	0.9
7	3.5	4.5	1.0
8	4.8	4.9	0.1
9	4.4	4.2	-0.2
10	3.9	4.7	0.8
			$\bar{d}=0.5$
			SDd=0.494

- 1、t値を求める(SDd:差の標準偏差
 δ :差の母集団の平均値)

$$t = (\bar{d} - \delta / \text{SDd}) / \sqrt{n} \\ = 3.198$$

- 2、自由度(df)を求める

$$\text{df} = n - 1 \\ = 10 - 1 = 9$$

- 3、t値に対応する確率(P)が有意水準
より小さいか確かめる

$$p = \text{TDIST}(3.198, 9, 2) \\ = 0.011 < 0.05$$

差があるといえる

3、対応のあるt検定

被験者	投与前	投与後	差(di)
1	3.6	4.8	1.2
2	4.3	4.4	0.1
3	3.8	3.7	-0.1
4	4.5	5.0	0.5
5	4.2	4.9	0.7
6	4.2	5.1	0.9
7	3.5	4.5	1.0
8	4.8	4.9	0.1
9	4.4	4.2	-0.2
10	3.9	4.7	0.8
			$\bar{d}=0.5$
			$SDd=0.494$

- 二つの平均値の差の信頼区間

$$\bar{d}-t(\alpha/2,df)*SDd/\sqrt{n}<\delta<\bar{d}+t(\alpha/2,df)*SDd/\sqrt{n}$$



TINV(α,自由度)

$$0.147<\delta<0.853$$

二つの平均値の差の95%信頼区間は
(0.147mEq/l, 0.853mEq/l,)

4、まとめ

- 二つの標本に対応があるかないか(同じ母集団から抽出しているかどうか)で計算方法が変わってくる
- t検定を行う際は、同時に信頼区間も求めることができる

登場した関数について

- FDIST(F値, 自由度 1 , 自由度 2)
F分布の確立関数の値が返ってくる
- TDIST(t値の絶対値, 自由度, x)
x:両側検定→ 2、片側検定→ 1
t分布の確立関数の値が返ってくる
- TINV(α , 自由度)
t分布の逆関数の値が返ってくる