

二つの差の検定（平均値と比率）

道具としての統計学

著者 奥田千恵子

第九章

平均値の差の検定

比較の種類

- 対応のない比較（分類要因）
 - － 例 男性と女性の肺活量の比較
 - － 性別などから分類できる二つの母集団から無作為抽出したものの比較
- 対応のある比較（処置要因）
 - － 例 Aさんの薬を投与前後の脈拍の比較
 - － 抽出したものの処置で比較する

二種類のt検定

- 対応のないt検定
 - Unpaired t test
- 対応のあるt検定
 - Paired t test

Unpaired t testの条件

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

(等分散の検定)計算手順

1. 標本分散の比を求め、Fを出す

$$F = \frac{sd_1^2}{sd_2^2} \quad (\text{sd1, sd2は二つ標本の分散})$$

2. それぞれの標本数 n_1, n_2 から自由度を求める
第一自由度 $= n_1 - 1$ 第二自由度 $= n_2 - 1$

3. F分布と自由度を用いてPを求めて有意水準 α より大きければ等分散が等しいとする

$$P = FDIST(F, \text{第一自由度}, \text{第二自由度})$$

Unpaired t test計算手順1

1. 二つの標本平均の差から SEM_d を求める

$$SEM_d = \sqrt{\left(\frac{(n_1-1) \times sd_1^2 + (n_2-1) \times sd_2^2}{n_1 + n_2 - 2} \right)}$$

2. 標本平均の差 d ,母平均値の差 δ から
標準値 t を求める

$$t = \frac{d - \delta}{SEM_d} \quad (d = \bar{X}_1 - \bar{X}_2, \delta = u_1 - u_2)$$

Unpaired t test計算手順2

1. 自由度,有意水準,片側か両側検定か決める
自由度 $df = n_1 + n_2 - 2$
2. t分布と自由度を用いてPを求める
 $P = \text{TDIST}(t, \text{自由度}, \text{両側/片側の区別})$
3. $P < \alpha$ (有意水準) ならば母平均に有意差があるとする

Paired t testの条件

1. 標本が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っている
 - 大きく違わなければ問題ない

第十章

2つの比率の差の検定

比較の種類

- 対応のない比較（分類要因）
 - － 例 男性と女性の肺活量の比較
 - － 性別などから分類できる二つの母集団から無作為抽出したものの比較
- 対応のある比較（処置要因）
 - － 例 Aさんの薬を投与前後の脈拍の比較
 - － 抽出したものの処置で比較する

四種類の比率検定

- 対応のないある比率検定
 - 正規分布に近似させる検定法
 - χ^2 検定
 - フィッシャーの直接確率法
- 対応のある比率検定
 - マニクマーの検定

正規分布の近似の条件

1. 標本は2値変数である
2. 標本は母集団から無作為抽出している
3. 標本が互いに独立している
4. 標本数や各カテゴリの生起確率が大きい

計算手順1

1. 母比率 p_1, p_2 と標本数 n_1, n_2 生起数 m_1, m_2 から
生起確率 $\dot{p}_1, \dot{p}_2$ を求める

$$\dot{p}_1 = \frac{m_1}{n_1} \quad \dot{p}_2 = \frac{m_2}{n_2}$$

2. 母比率に差がないという帰無仮説の場合
二つの併合した確率 $\dot{p}$ は以下の様に表すことができる

$$\dot{p} = \frac{m_1 + m_2}{n_1 + n_2}$$

計算手順2

1. 2つの標本比率の差から SE_{pd} を求める

$$SE_{pd} = \sqrt{\dot{p}(1 - \dot{p}) \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}$$

2. SE_{pd} を用いて標準値 z を求める

$$z = \frac{\dot{p}_1 - \dot{p}_2}{SE_{pd}}$$

3. 有意水準 α と片側検定か両側検定か決め
 z 値に対応する確率 P を求め

計算手順3

1. 有意水準 α と片側検定か両側検定か決める
2. 標準正規分布を用いて z 値に対応する確率 P を求める
 $P = \text{NORMSDIST}(z)$
3. P が α よりも大きければ十分に差があるといえる

χ^2 検定の条件

1. 標本は母集団から無作為抽出している
2. 標本が互いに独立している
3. 標本数や各カテゴリの生起確率が大きい

フィッシャーの直接確率法の条件

1. 標本は2値変数である
2. 標本は母集団から無作為抽出している
3. 標本が互いに独立している
4. 標本数や、各カテゴリの生起確率が小さい

マクニマーの検定の条件

1. 標本は2値変数である
2. 標本が何らかの形で対になっている
3. 対になった標本は母集団から無作為抽出している
4. 標本数や各カテゴリの生起確率が大きい

まとめ

- t検定には様々な検定方法がある
- 検定方法には対応のあるものやないものなどでも分けられる