

復習（標準正規分布～統計的仮説検定）

道具としての統計学

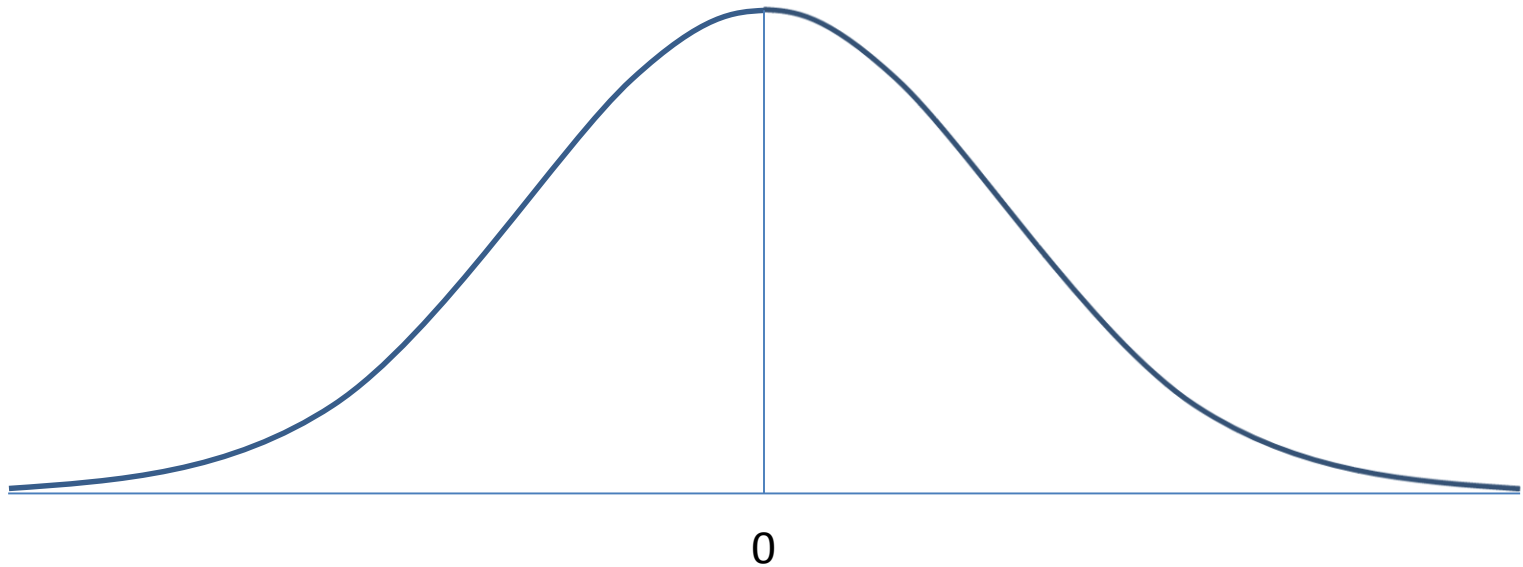
著者：奥田 千恵子

標準化(正規化)

- 与えられたデータを平均が0で分散が1のデータになるように変換する操作のこと
- $z = \frac{x - \mu}{\sigma}$ or $z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$
- zはz値、 μ は母平均、 σ は母標準偏差
- 利点: 異なった基準で測られた値が比較できる

標準正規分布

- 平均値が0、標準偏差が1の正規分布



t分布

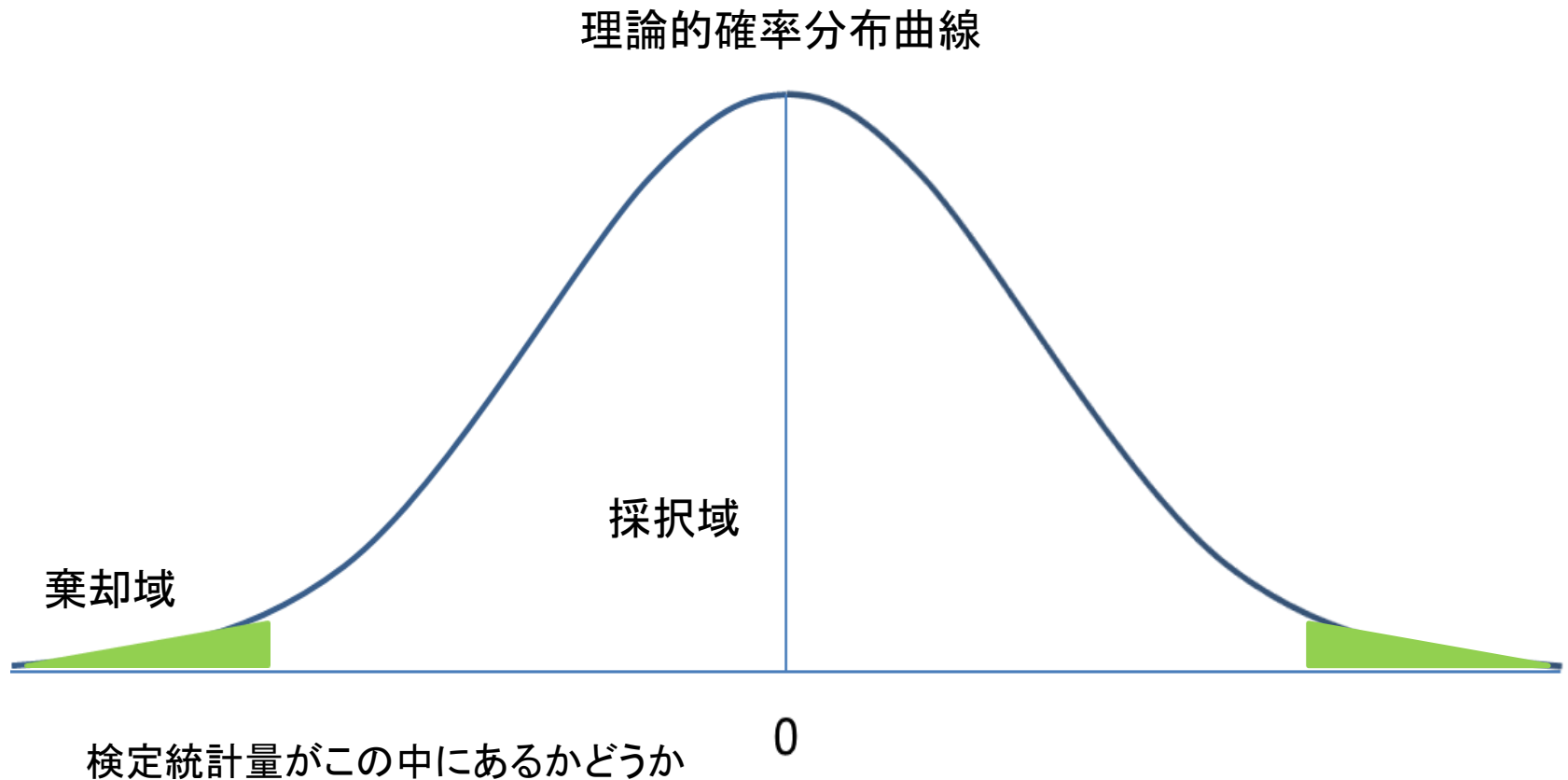
- 母標準偏差 σ が未知数の時、標本標準偏差を用いて代用する方法
- $t = \frac{\bar{X} - \mu}{\frac{SD}{\sqrt{n}}}$
- 自由度 $n-1$ が無限大の時の t 値が z 値である
(標本数が100近くあれば標準正規分布とほぼ重なる)

統計的仮説検定

- 理論的確率分布曲線下の面積全体に対して裾側の一定の面積に当たる優位水準を事前に設定し検定統計量が棄却域に入っているかどうかによって、得られた結果があれかどうかを判断すること
- 理論的確率分布
 - 標準正規分布、t分布など
- 検定統計量
 - z値、t値など
- P値とは
 - 裾側の面積のこと
 - 通常はツールで計算する

統計的仮説検定

P値 = TDIST(信頼度, 自由度, 片側or両側)



過誤率について

同じ確率

- α 過誤
 - 帰無仮説が正しいのに棄却してしまう誤り
- 有意水準
 - ある事象が起こる確率が偶然とは考えにくい基準
 - 水準 $\alpha\%$ で確率がそれよりも低ければ帰無仮説を棄却する
- 検出力($1-\beta$)
 - 帰無仮説が正しくないときに棄却する確率のこと
 - 正しくないのに棄却しない確率のことを β 過誤という
 - 例えば本当は平均が差があるデータであるのに差がないと判断してしまうなど

→ 背反の考え方

t検定

- t分布を利用する検定
 - 以下のことはパラメトリック検定法である
- 2つの平均値の差の検定であるt検定には主に3種類
 - 等分散を仮定した、対応のないt検定
 - 母分散が等しいかどうかはF検定などを使うこと
 - スチューデントのt検定
 - 等分散を仮定しない対応のないt検定
 - ウェルチのt検定
 - 対応のあるt検定

等分散を仮定した対応のないt検定

- 標本は母数から無作為抽出されている
- 母集団は正規分布に従う
- 標本が互いに独立している
- 2つの母分散が等しい
 - F検定を行って確かめる

対応のあるt検定

- 標本が何らかの形でついになっている
- 対になった標本は母集団から無作為抽出されている
- 対になった標本の差が近似的に正規分布に従う

対応のある、ない検定の違い

- 例:対応のない
 - 日本人男性の20歳と40歳の肺活量の違い
- 例: 対応のある
 - 運動前と運動後の心拍数で違いはあるか
 - PCの起動とシャットダウンの時間で違いがあるのか

最後に

1. データは連続量か離散量か
2. パラメトリック検定か、ノンパラメトリック検定か
3. 群のデータは2つか、3つ以上か
4. 1変数の解析か2変数の関係を求めるのか

索引

- http://bio-info.biz/statistics/test_welch_t.html
- http://bio-info.biz/statistics/test_student_t.html
- http://bio-info.biz/statistics/element_parametric.html
- <https://kotobank.jp/word/%E6%9C%89%E6%84%8F%E6%B0%B4%E6%BA%96-676246>
- <http://www.gen-info.osaka-u.ac.jp/testdocs/tomocom/express/express11.html>

パラメトリックとノンパラメトリック

- パラメトリック検定とは
 - 母集団の分布がある特定の分布に従うことがわかっているデータに対して行う検定法
- ノンパラメトリック検定とは
 - 母数 (パラメーター) に対する一切の前提を仮定しない
 - 得られたデータ数が少なく、データが従う分布を仮定することが困難であり、パラメトリック検定を利用することが不適切であると判断される際に利用

ノンパラメトリックなt検定

- 対応のない2群間の検定
 - マン・ホイットニーの検定（ウィルコクソンの順位和検定）
 - 標本は母集団から無作為抽出されている
 - 標本が互いに独立している
- 対応のある2群間の検定
 - ウィルコクソンの符号付き順位検定
 - 標本が何らかの形で対になっている
 - 対になった標本は母集団から無作為抽出されている