

第5章

推定

統計学講義 (P.97~109)

稲垣宣生 吉田光雄 山根芳知 地道正行 共著

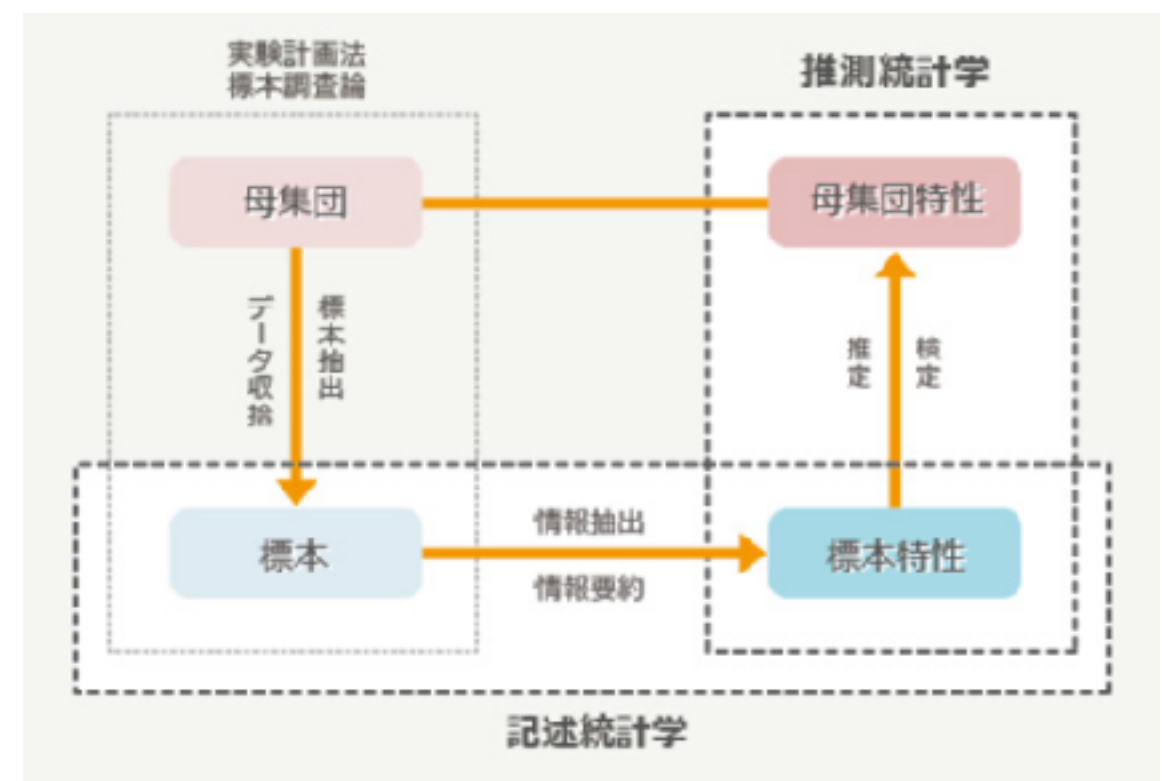
発表者 吉富 楓雅

統計学の種類(再掲)

- 統計学の種類

1. 記述統計学 (古典統計学)

2. 推測統計学 (近代統計学)



統計学の分類

推定とは？

- **母集団の情報を得ることが目的**
 - * データの背後により大きい母集団があると考え、データはそこから無作為に抽出された標本とみなし、データ処理により母集団の性質を統計的に推測
- **母数の値を「データの値から推測する」**
 - * 母数：母集団の平均や分散などの特性値

推定量とその性質

確率関数と密度関数

- 確率変数 X の従う分布が、母数 θ によって特徴付けられるものである
 - * 離散型分布：確率関数 $p(x | \theta)$
 - * 連続型分布：密度関数 $f(x | \theta)$
- n 個の無作為標本の組をベクトル $X = (X_1, X_2, \dots, X_n)$ で表す.

推定量と推定値

- **推定量**(estimator): 母数 θ の推定のために用いられる統計量

$$T = T(X) = T(X_1, \dots, X_n)$$

- **推定値**(estimate): データ $x = (x_1, \dots, x_n)$ に対する推定量の値

$$T = T(X) = T(x_1, \dots, x_n)$$

- 一般に, 推定量 T は母数 θ の周りに分布している
- **偏り**(bias): 推定量 T の平均 $E(T)$ と母数 θ との誤差

$$d = E(T) - \theta$$

- $E(T) = \theta$ の時は不偏. その推定量を不偏推定量(unbiased estimator; UE)という

信頼できる推定を行う

信頼度100%はダメ？

- 研究では信頼係数**99%**, **95%**, **90%**がよく使用される
 - ▶ データ数や状況によって使用する信頼係数を変える
- 信頼係数100%は使用されない
 - ▶ 100%信頼できるのに、なぜ使用しないのか？

信頼度100%はダメ！

- 信頼区間の幅は%が大きくなるにつれ広くなり, 100%では無限大

* つまり**100%は無意味**

- 高校生の身長の信頼区間が信頼度100%で[1m, 2m]となった

* 常識で考えて当たり前の結果を示しただけ

- **できるだけ大きい信頼係数で可能な限り狭い範囲の推定区間を得る必要がある！**

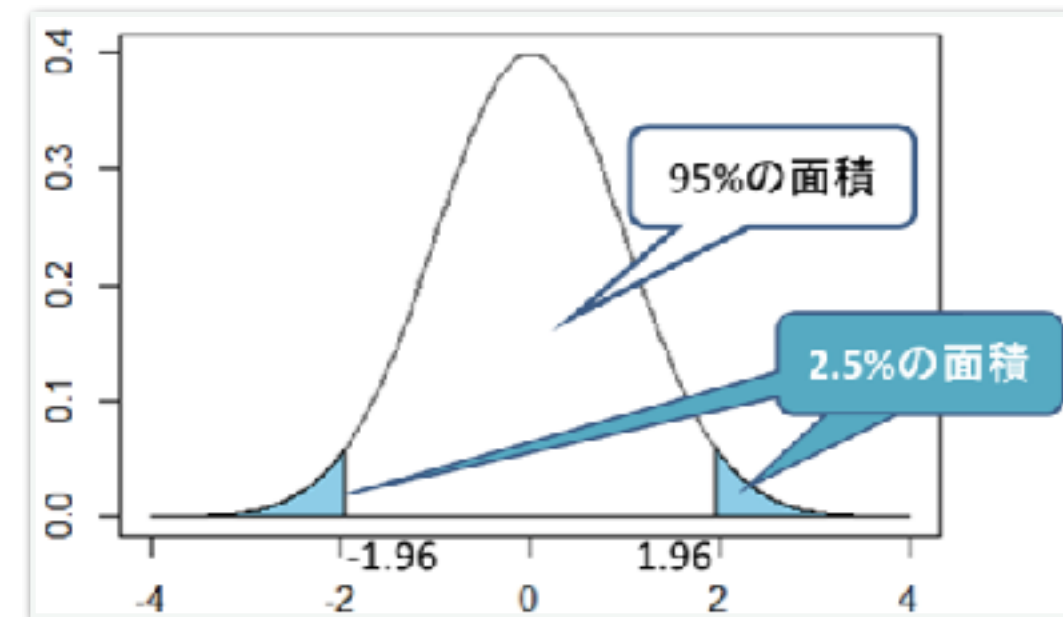
例題

- 2015年12月末時点の各都道府県内にある映画館のスクリーンの数
 - * 無作為に10都道府県のデータを抽出
 - ▶ 母分散 = 5560
 - ▶ スクリーン数の分布は正規分布に従う
- このデータから母平均の95%信頼区間を求めよ

No.	都道府県	全スクリーン数
1	兵庫	126
2	大阪	224
3	奈良	34
4	岩手	25
5	千葉	199
6	茨城	89
7	福岡	178
8	山梨	14
9	滋賀	38
10	鳥取	11

信頼区間95%とは？

- 標準化した値が95%の面積内にあれば良い
 - * 複数の信頼区間の中で真値の含まれる信頼区間は95%
- 信頼区間96%だったら正確になる？
 - * 信頼区間にも多少の不確かさがあるため、95%と96%を比較することにはあまり意味はない



母分散が判明している時の区間推定

1. 標本平均を求める

2. 標本平均の標準化を求める

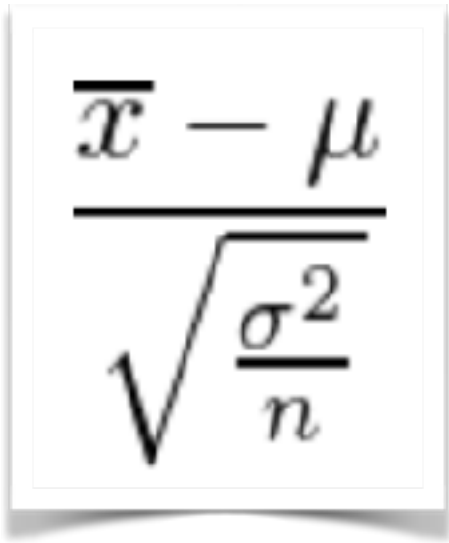
* 母平均を μ ，母分散を σ^2 ，抽出したサンプル数を n とすると標本平均の標準化は右式の通り

• 母平均の信頼区間の算出には**中心極限定理**を利用

* 標準正規分布を利用するため，**標本平均を標準化**

3. 標準正規分布表から，上側2.5%部分となる Z の値を求める

* 今回は $z = 1.96$


$$\frac{\bar{x} - \mu}{\sqrt{\frac{\sigma^2}{n}}}$$

母分散が判明している時の区間推定

- よって標準化した値の範囲は図1の通り

$$-1.96 \leq \frac{\bar{x} - \mu}{\sqrt{\frac{\sigma^2}{n}}} \leq 1.96$$

(図1) 標準化した値の範囲

- * 1.96はよく出てくるので覚えておく和良好的

$$\bar{x} - 1.96 \times \sqrt{\frac{\sigma^2}{n}} \leq \mu \leq \bar{x} + 1.96 \times \sqrt{\frac{\sigma^2}{n}}$$

(図2) μ について変換

- μ について式を変形(図2)
- 値を代入(図3)

$$93.8 - 1.96 \times \sqrt{\frac{5560}{10}} \leq \mu \leq 93.8 + 1.96 \times \sqrt{\frac{5560}{10}}$$

(図3) それぞれの値を代入

$$47.6 \leq \mu \leq 140.0$$

解析ツールで信頼区間
を求めましょう！