

統計的検定入門

資料：Rによる統計的検定と推定・第2章

ISBN978-4-274-06878-2 オーム社

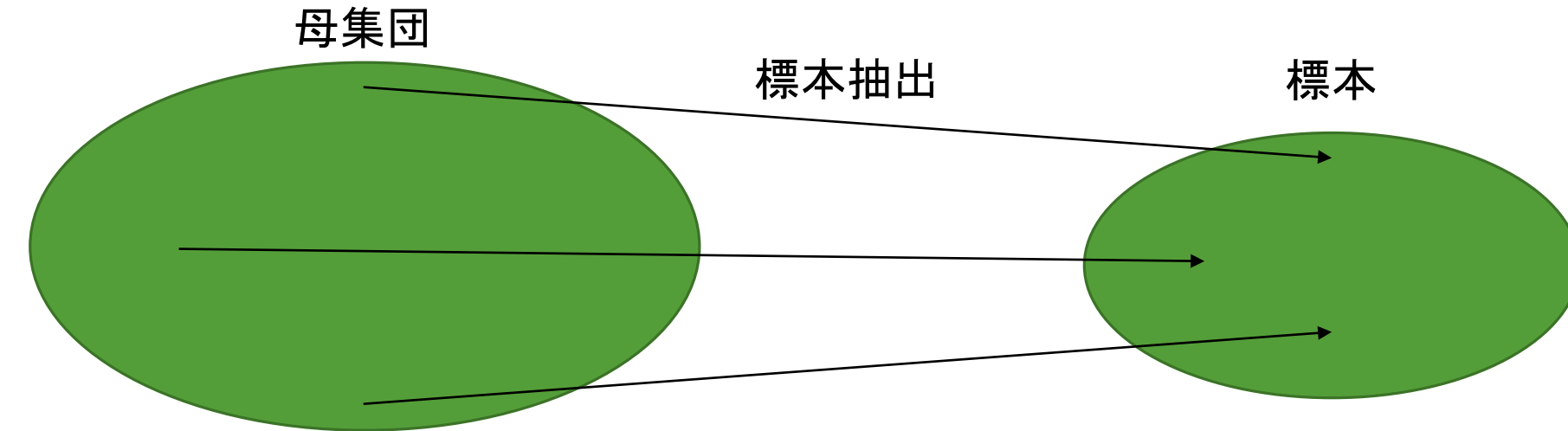
著者：内田 治 西澤 英子

60171050 松野祐介

検定の役割

- ◆ 比較した際の差が統計的に有意であるか判定
- ◆ 標本の差が母集団に差があるから生じたとは限らない

母集団と標本



母数

- ◆ 母平均
- ◆ 母分散
- ◆ 母標準偏差
- ◆ 母相関係数
- ◆ 母比率

標本統計量

- ◆ 標本平均
- ◆ 標本分散, 不偏分散
- ◆ 標本標準偏差
- ◆ 標本相関係数
- ◆ 標本比率

検証される仮説

◆ 帰無仮説 $H_0: \mu = 80$

◆ 対立仮説 $H_1: \mu \neq 80$

◆ 対立仮説の立て方

➤ 両側仮説: $\mu \neq 80$

➤ 片側仮説: $\mu > 80$ 、または $\mu < 80$

◆ 帰無仮説 H_0 が正しいと仮定して検証を行う

有意確率・有意水準

◆p値(有意確率)

- 帰無仮説が正しいと仮定した場合に偏った値が生じる確率
- * この値が小さい時帰無仮説は棄却される

◆有意水準 α

- p値が小さいかどうかの判定基準で、0.05とするのが習慣

◆p値が有意水準 α より小さければ帰無仮説は棄却

判断の誤りと検出力

		検定の結果	
		帰無仮説	対立仮説
本当の状態	帰無仮説	○	第1種の過誤
	対立仮説	第2種の過誤	○

- ◆ 第1種の過誤: 帰無仮説が真である時に、帰無仮説を棄却
- ◆ 第2種の過誤: 帰無仮説が偽である時に、帰無仮説を棄却しない
- ◆ 第2種の過誤の確率は β で表され、 $1 - \beta$ を検出力と呼ぶ

区間推定

- ◆ 検定と併用されることが多い統計的手法
- ◆ 信頼率：通常は0.95という数値を用いる
 - 信頼率0.95の信頼区間は95%信頼区間と呼ばれる
- ◆ 例： $76.5 \leq \mu \leq 84.5$

検定の種類1

◆正規性に関する検定

- シャピロ-ウィルク(Shapiro-Wilk)検定
- コルモゴロフ-スミルノフ(Kolmogorov-Smirnov)検定
- 適合度のカイ2乗検定

* 上記以外にヒストグラムや正規確率プロットを併用するとよい

◆母平均に関する検定

- z検定(母標準偏差が既知のときに用いる)
- t検定(母標準偏差が未知のときに用いる)
- 独立した2つの標本に関するt検定(2つの平均値の差に関する検定)
- 対応のある標本に関するt検定(対応のある平均値の差に関する検定)
- 分散分析(2つ以上の平均値に関する検定)

検定の種類2

◆母分散に関する検定

- F検定(2つの分散の比に関する検定)
- レーベン(Levene)検定(2つ以上の分散に関する検定)
- バートレット(Bartlett)検定(2つ以上の分散に関する検定)
- ハートレー(Hartley)検定(2つ以上の分散に関する検定)

◆割合に関する検定

- 二項検定
- 分割表(クロス集計表)の独立性に関するカイ2乗検定

◆相関に関する検定

- 母相関係数に関するt検定

検定の選択

◆ 測定 of 尺度

- 名義尺度
- 順序尺度
- 連続尺度

◆ データの種類

- 計数值
- 計量値
- 順位値

◆ グループの数

- 1つ
- 2つ
- 3つ以上

* また、データの数も検定方法の選択に影響する

母割合に関する検定

◆ 2つの商品A, Bを30人に評価させたとき

➤ Aのほうが好き: 22人

➤ Bのほうが好き: 8人

◆ 帰無仮説 H_0 : AとBの好きな人の割合に差は無い

◆ 対立仮説 H_1 : AとBの好きな人の割合に差がある

→ 帰無仮説のもとで、Aが22人以上、または、8人以下となる確率

→ $p\text{値} = 0.01612 < 0.05$

→ 結果: 帰無仮説は棄却される

まとめ

◆統計的検定の入門として以下の内容に触れた

- 母集団と標本
- 帰無仮説・対立仮説
- 有意確率・有意水準
- 判定の誤り・検出力
- 検定の種類
- 検定の選択