

情報で学ぶ統計学③

度数の比較

この授業の目標

- 度数の比較について理解する
- 検定をツールを使って行える

この授業のカリキュラム

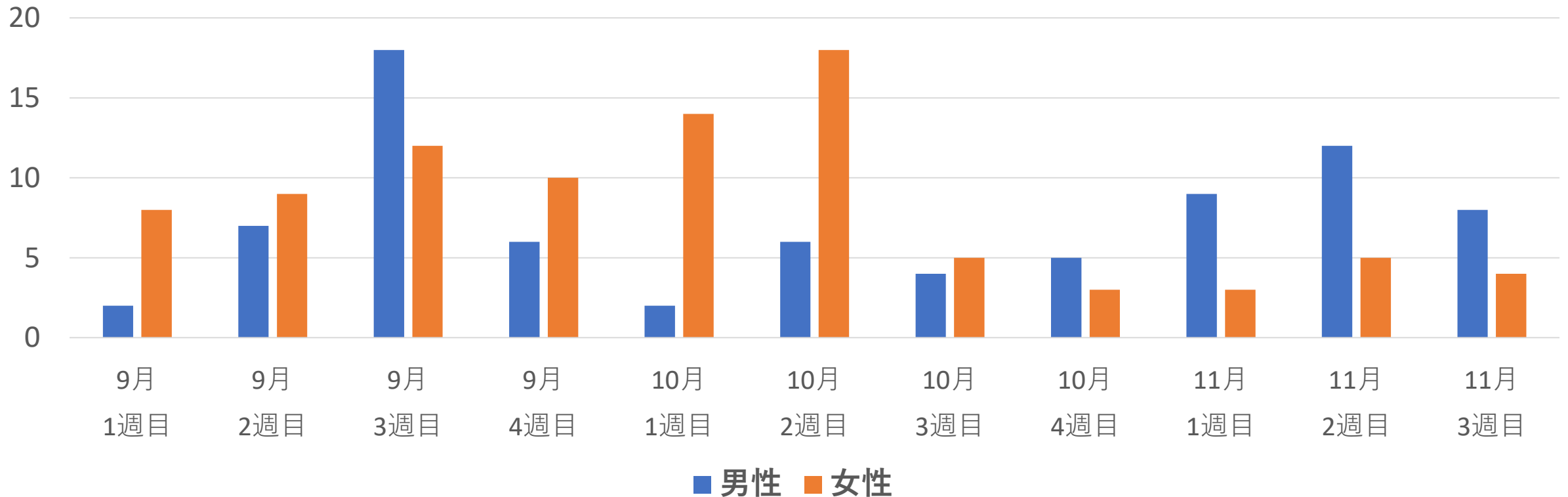
1. 検定の考え方
2. 検定をやってみよう

検定の考え方

この「検定」は英検や漢検の検定と違います。

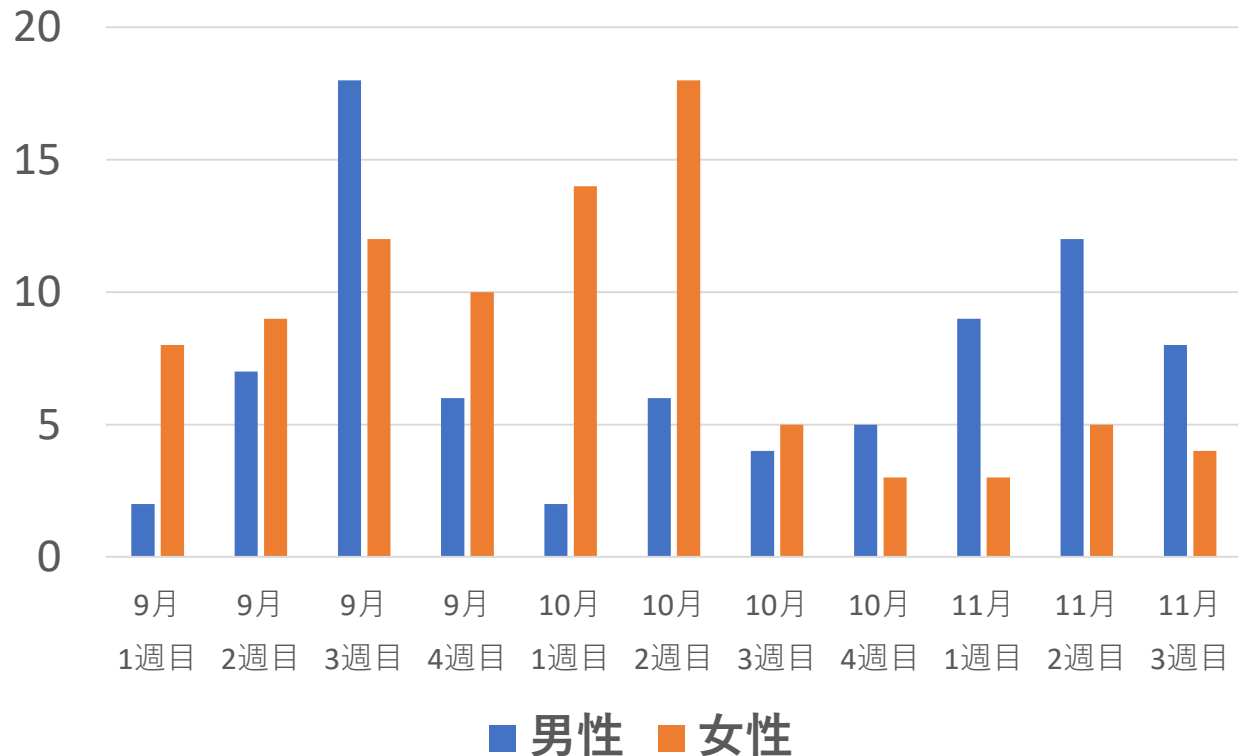
Twitterの長時間利用者数

| | 9月 1週目 | 9月 2週目 | 9月 3週目 | 9月 4週目 | 10月 1週目 | 10月 2週目 | 10月 3週目 | 10月 4週目 | 11月 1週目 | 11月 2週目 | 11月 3週目 | 合計 |
|----|-----------|-----------|-----------|-----------|------------|------------|------------|------------|------------|------------|------------|----|
| 男性 | 2 | 7 | 18 | 6 | 2 | 6 | 4 | 5 | 9 | 12 | 8 | 79 |
| 女性 | 8 | 9 | 12 | 10 | 14 | 18 | 5 | 3 | 3 | 5 | 4 | 91 |



Twitterの長時間利用者数

| | 9月 1週 目 | 9月 2週 目 | 9月 3週 目 | 9月 4週 目 | 10月 1週 目 | 10月 2週 目 | 10月 3週 目 | 10月 4週 目 | 11月 1週 目 | 11月 2週 目 | 11月 3週 目 | 合計 |
|----|---------------|---------------|---------------|---------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----|
| 男性 | 2 | 7 | 18 | 6 | 2 | 6 | 4 | 5 | 9 | 12 | 8 | 79 |
| 女性 | 8 | 9 | 12 | 10 | 14 | 18 | 5 | 3 | 3 | 5 | 4 | 91 |

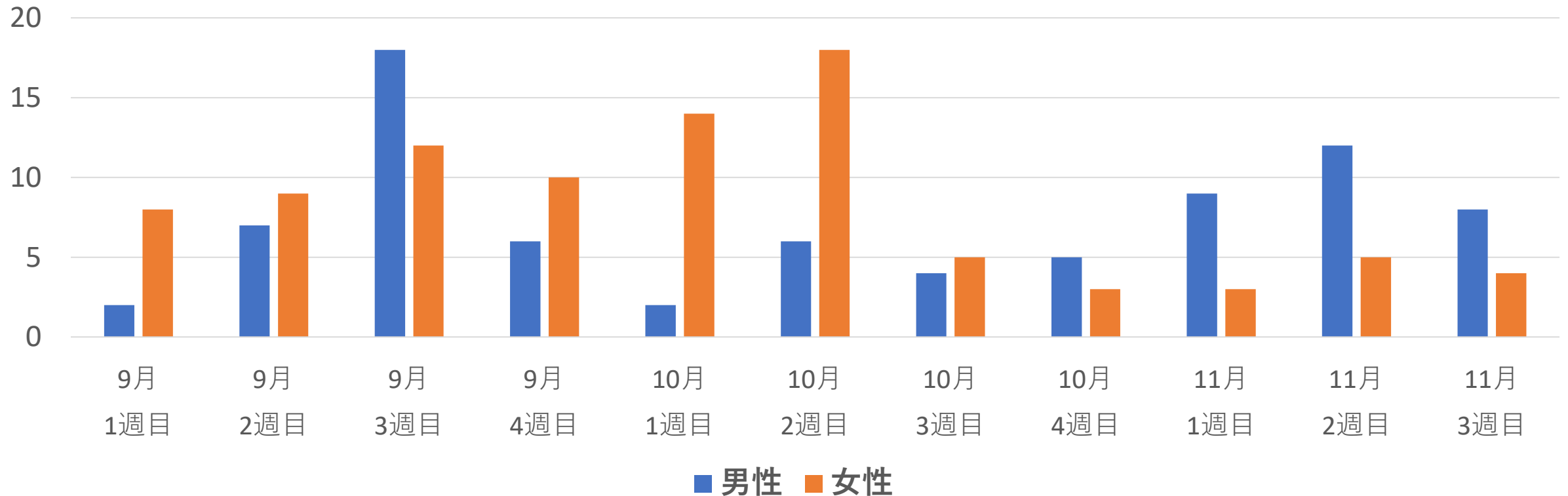


KEYWORD

- 統計では
「男女の人数が79：91であつたときに、人数に違いがあるかを**検定する**」などの言い方をする。
- もし違いがあるときは、「**有意な違いがある**」と表現する。

Twitterの長時間利用者数

| | 9月 1週目 | 9月 2週目 | 9月 3週目 | 9月 4週目 | 10月 1週目 | 10月 2週目 | 10月 3週目 | 10月 4週目 | 11月 1週目 | 11月 2週目 | 11月 3週目 | 合計 |
|----|-----------|-----------|-----------|-----------|------------|------------|------------|------------|------------|------------|------------|----|
| 男性 | 2 | 7 | 18 | 6 | 2 | 6 | 4 | 5 | 9 | 12 | 8 | 79 |
| 女性 | 8 | 9 | 12 | 10 | 14 | 18 | 5 | 3 | 3 | 5 | 4 | 91 |



「有意である」場合

長時間利用者の「男性と女性の比率は等しい」と仮定する。

有意水準を決める（ここでは5%）に決める。

長時間利用者が5人であったとして、全員女性である確率（これは約3%）。

3%は有意水準5%よりも小さい。

めったに起こらないことが起こってしまった。それは変なことだ。

そもそも長時間利用者の「男性と女性の比率は等しい」という最初の仮説が間違っていたのだ。

長時間利用者の男性と女性の比率は有意に違う！

「有意ではない」場合

長時間利用者の「男性と女性の比率は等しい」と仮定する。

有意水準を決める（ここでは5%）に決める。

長時間利用者が4人であったとして、全員女性である確率（これは約6%）。

6%は有意水準5%よりも大きい。

起こりうることであると考え。たまたまこういう偶然もあるだろう。

長時間利用者の「男性と女性の比率は等しい」という最初の仮説は否定できない。

長時間利用者の男性と女性の比率は有意に違うとは言えない！

検定を試してみよう

「Js-STAR」と検索してみてください。
すごいWebサイトです。

直接確率検定（正確二項検定）

★お知らせ

- What's new!
- 動作確認・バグ状況
- Tips

★各種分析ツール

度数の分析

- 1 × 2 表(正確二項検定)**
- 1 × 2 表・母比平均値
- 1 × j 表(カイ二乗検定)
- 2 × 2 表(Fisher's exact test)
- i × j 表(カイ二乗検定)
- 2 × 2 × k 表(層化解析)
- i × j × k 表(モデル選択)
- i × j × k × L 表(4元モデル)
- 自動集計検定 2 × 2
- t 検定(Rプログラム)

直接確率計算 1 × 2 表(正確二項検定)

メイン データ形式 グラフ 説明

データ

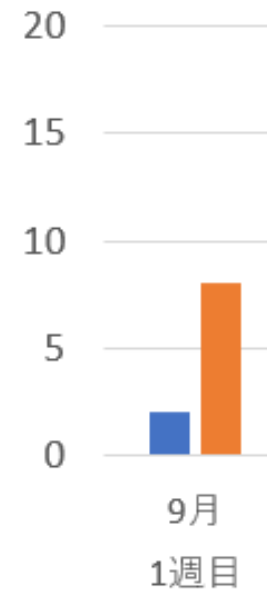
Q&A入力 データ消去

観測値 1 と観測値 2 に度数を入力してください。
(総度数Nは自動的に計算します)

| | |
|----------------------|----------------------|
| 観測値 1 | 観測値 2 |
| <input type="text"/> | <input type="text"/> |

N =

| | 9月 1週目 |
|----|-----------|
| 男性 | 2 |
| 女性 | 8 |



<http://www.kisnet.or.jp/nappa/software/star/>

Js-STAR

直接確率検定（正確二項検定）

結果

保存 コピー 消去 タブ変換 伸▼ ▲縮

[直接確率計算 1 × 2]

| 観測値 1 | 観測値 2 |
|---------------|---------------|
| 2 (0.2000) | 8 (0.8000) |

両側検定 : $p=0.1094$ ns ($.10 < p$)
片側検定 : $p=0.0547$ + ($.05 < p < .10$)

効果量 : $g=0.3000$

/// Analyzed by js-STAR ///

<http://www.kisnet.or.jp/nappa/software/star/>

Js-STAR

効果量は後日説明します。

この授業の振り返り

- 度数の比較について理解する
- 検定をツールを使って行える

参考文献



中野博幸, 田中敏 (2012) フリーソフトjs-STARでかんたん統計データ分析 技術評論社



向後千春, 富永敦子 (2016) 身につく 入門統計学 技術評論社