

# データ分布と予測

## 仮説検定

堀田敬介  
2009/12/24, Tue.

## Contents

- **仮説検定とは？**
  - 仮説検定の思想
  - 仮説検定における仮説: 帰無仮説と対立仮説
  - 対立仮説のたて方: 片側検定と両側検定
  - 棄却域・有意水準, 仮説検定における誤りと検出力
- **母集団の母数に対する仮説検定**
  - 母平均の検定 (母分散が**既知**の場合)
  - 母平均の検定 (母分散が**未知**の場合)
  - 母分散の検定
  - 母比率の検定
- **2つの母集団に対する仮説検定**
  - 平均値の差の検定 (母分散が**既知**の場合)
  - 平均値の差の検定 (母分散が**未知**だが**等しい**場合)
  - 平均値の差の検定 (母分散が**未知**で**等しくない**場合)
  - 分散の比の検定
- **適合度検定と独立性の検定**

## 仮説検定とは？

推測統計  
statistical inference

標本データの平均値と分散 (標準偏差) から, 母集団の平均値と分散 (標準偏差), 母比率などの母数 (パラメータ) を推定する

仮説検定  
hypothesis testing

標本データの平均値と分散 (標準偏差) などをもとに, 母集団に対する「**ある仮説**」が**間違いかどうか**判定する

母数  $\mu, \sigma^2$  **母集団** 母集団に対する仮説  $H_0$

↑ 標本から母集団を推定      ↓ 無作為抽出 (ランダムサンプリング)      ↓ 標本から仮説を検定

標本

## 仮説検定とは？

• 例: 日本人成人女性の平均身長

仮説 「日本人成人女性の平均身長は160cmである」

無作為抽出 (ランダムサンプリング)

**母集団**  
日本人 成人女性

標本	1	2	3	4	5	6	7	8	9	10
身長	150	165	155	170	150	145	175	160	165	140

標本平均の値: 157.5cm

仮説は間違っているのか? それとも正しいのか?

### 仮説検定とは？

標本身長	1	2	3	4	5	6	7	8	9	10
	150	165	155	170	150	145	175	160	165	140

標本平均の値: 157.5cm

- 例: 日本人成人女性の平均身長
- 仮説 「日本人成人女性の平均身長は160cmだ」
- 否定 「いや違うわ! 160cmより低いはずよ」

たまたま測った10人の平均が160cmでないからと言って、仮説を否定はできないよ

でも無作為抽出(ランダムサンプリング)してるのだから、「仮説が正しい」なら「160cmを大幅に下回ることなんて**滅多に無い**」はずよ

じゃあ、標本平均が160cmから大きく外れてなければ、仮説は間違っていないと判断することにしようよ

そうね、「仮説が間違っていると**見えそうな範囲**」を決めましょう

### 仮説検定とは？

- 例: 日本人女性の平均身長

仮説が間違ってる**っぽい範囲**

$\bar{X} \leq ???$

仮説が間違ってる**っぽい範囲**

$t$ 分布の下側5%

$T \leq -1.833$

5% 自由度9の**t**分布  $T = -1.833$

95%

標本平均  $\bar{X}$ 、標本分散  $S^2$  に対し、確率変数  $T$  が自由度  $n-1$  の  $t$  分布に従う  $T = \frac{\bar{X} - \mu}{S/\sqrt{n-1}} \sim t_{\alpha}(n-1)$

棄却域 rejection region (臨界域 critical region)

採択域 region of acceptance

### 仮説検定とは？

- 例: 日本人女性の平均身長

仮説 「日本人成人女性の平均身長は160cmである」

標本平均の値:  $\bar{X} = 157.5cm$

標本分散の値:  $S^2 = 113.3$  (標本標準偏差の値:  $S = 10.8$ )

自由度9の**t**分布の下側5%棄却域

$T \leq -1.833$   $T = \frac{\bar{X} - \mu}{S/\sqrt{n-1}} \leq -1.833$

$\bar{X} \leq \mu - 1.833 \frac{S}{\sqrt{n-1}}$

$\bar{X} \leq 160 - 1.833 \frac{10.8}{\sqrt{10-1}} = 153.4$

標本平均の**棄却域**  $\bar{X} \leq 153.4$

標本平均値157.5cmで棄却域にないため、仮説は棄却されないつまり、仮説が間違っているとは言えない

### 仮説検定とは？

- 仮説検定の手順

- 母集団に対する「仮説」を立てる
- 「仮説」の棄却域を設定する
- 結論を述べる

棄却域を5%とするということは、残り95%信頼区間から外れていれば「仮説」を「棄却」するということ

棄却域: 小 信頼区間: 広 ↔ 棄却域: 大 信頼区間: 狭

★ 棄却域を5%と設定するということは、検定の結論が間違っている危険性が5%はあるということ

5%: **有意水準** (significance level) [危険率 (risk)]

## 仮説検定とは？

- 母集団に対する「仮説」について
  - 【**帰無仮説**】日本人成人男性の平均体重は60kgである
  - 【**対立仮説**】日本人成人男性の平均体重は60kgではない

↓

- ★ 仮説1を統計的に検定したとき、その起こる確率が棄却域にあれば、この仮説が棄却される。
- ★ 仮説1が棄却されない場合、「**仮説1が正しいという結論は出せない**」!
- ★ 仮説1は「**棄却されてはじめて意味を持つ**」
- ★ 仮説2は仮説1が棄却された場合に採択される

**帰無仮説 (null hypothesis)**  
棄却されてはじめて意味を持つ

**対立仮説 (alternative hypothesis)**  
本当に示したいこと

統計的検定の目的は「**対立仮説の正しさを示す**」こと!

## 仮説検定とは？

- 対立仮説の立て方
  - 【**帰無仮説**】  
 $\mu = 60$  : 日本人男性の平均体重は60kgである
  - 【**対立仮説**】
    - $\mu \neq 60$  : 日本人男性の平均体重は60kgではない **両側検定 two tailed test**
    - $\mu > 60$  : 日本人男性の平均体重は60kgより重い **片側(右側)検定 one tailed test**
    - $\mu < 60$  : 日本人男性の平均体重は60kgより軽い **片側(左側)検定 one tailed test**

3つある対立仮説のうちのどれか1つを、検定をする人が選ぶ(実施する検定に対して最も適当と思われるものを選ぶ)

## 仮説検定とは？

- 棄却域・有意水準, 検定における誤りと検出力
  - 第1種の誤り error of the first kind
    - 帰無仮説 $H_0$ が正しいのにそれを棄却してしまう
      - 例: 品質管理において「合格するはずの良製品を不合格判定」する
      - 例: 刑事犯罪において「無罪の人を有罪」にする
  - 第2種の誤り error of the second kind
    - 帰無仮説 $H_0$ が誤っているのにそれを採択してしまう
      - 例: 品質管理において「不合格のはずの不良品を合格判定」する
      - 例: 刑事犯罪において「有罪の人を無罪」にする

		本当に成り立っているのは	
		帰無仮説 $H_0$	対立仮説 $H_1$
検定結果	$H_0$	正しい (その確率: $1-\alpha$ )	第2種の誤り (その確率: $\beta$ )
	$H_1$	第1種の誤り (その確率: $\alpha$ )	正しい (確率: $1-\beta$ = <b>検出力</b> )

$\alpha$ : 大  $\Leftrightarrow$   $\beta$ : 小  
 $\alpha$ : 小  $\Leftrightarrow$   $\beta$ : 大  
となるので、共に小さくはできない

- 帰無仮説が限定的なので $\alpha$ (有意水準)は1つに定まる。
- サンプル数が大きければ $\beta$ は小さくなる。

## 参考: 仮説検定における誤りと検出力

- 例: 『男にセクハラが多いのは何故か?』  
(出展: 『週刊東洋経済 2007/12/15号 一経済学ってこんなにももしろい!』 - p.66『男はなぜセクハラするのか』)
- 帰無仮説  $H_0$ : 目の前の異性は自分に気がある(→手を出す)
- 対立仮説  $H_1$ : 目の前の異性は自分に気がない(→スルー)

		本当に成り立っているのは	
		帰無仮説 $H_0$	対立仮説 $H_1$
異性が自分に気があるかどうかの検定結果	$H_0$	カップル(セクハラ不成立) (その確率: $1-\alpha$ )	第2種の誤り(セクハラ) (その確率: $\beta$ )
	$H_1$	第1種の誤り (その確率: $\alpha$ )	カップル不成立 (確率: $1-\beta$ = <b>検出力</b> )

**男**: 子孫を多く残したい  
→ 数多くの相手を見つけないと  
→ 女が自分の相手になっても良いと思ってるの  
会を逃すのは損失  
→ 気のない相手に手を出して断られても恥をかき済む

**女**: 妊娠・子育ては大きな負担  
ける男が必要  
分と子供を養育する意のある男を選びたい  
気の男を見つけたとしても、気のない相手に  
より良い

男: 第1種の誤り <<< 第2種の誤り >>> 女: 第1種の誤り >>> 第2種の誤り

### 母数に関する仮説検定

- 母平均に関する仮説検定 (母分散  $\sigma^2$  既知) Z検定

統計量  $\bar{X} \rightarrow Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$  が標準正規分布  $N(0,1)$  に従うことを利用
- 母平均に関する仮説検定 (母分散  $\sigma^2$  未知) t検定

統計量  $\bar{X} \rightarrow T = \frac{\bar{X} - \mu}{S/\sqrt{n-1}}$  が自由度  $n-1$  の  $t$  分布に従うことを利用
- 母分散に関する仮説検定  $\chi^2$ 検定

統計量  $S^2 \rightarrow \chi^2 = \frac{nS^2}{\sigma^2}$  が自由度  $n-1$  の  $\chi^2$  分布に従うことを利用

### 母数に関する仮説検定

- 母平均の検定 [Z検定] (母分散が既知の場合)
  - 例: BMIによる肥満検査
 

BMI値	判定
~20	やせ
~24	普通
~26.5	太気味
~∞	太過

ある会社の無作為抽出100人の社員のBMIが平均値=22.35だった。  
肥満の程度に問題があるといえるか? 有意水準5%で検定  
[出展:『図解雑学 統計解析』p.192]

↓

[帰無仮説] 母平均  $\mu=22$   
[対立仮説] 母平均  $\mu \neq 22$   
で両側検定。ただし、母標準偏差は過去の経験から2.5とする。

$\bar{X} \rightarrow Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$

確率変数Zは標準正規分布  $N(0,1)$  に従う

### 母数に関する仮説検定

- 母平均の検定 [Z検定] (母分散が既知の場合)
  - 例: BMIによる肥満検査
 

$Z \leq -1.96, Z \geq 1.96$ : 両側5%棄却域

$$\begin{cases} Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \leq -1.96 \\ Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \geq 1.96 \end{cases} \Leftrightarrow \begin{cases} \bar{X} \leq \mu - 1.96 \frac{\sigma}{\sqrt{n}} \\ \bar{X} \geq \mu + 1.96 \frac{\sigma}{\sqrt{n}} \end{cases}$$

標準正規分布: 両側5%棄却域

$\bar{X} \leq 22 - 1.96 \times \frac{2.5}{\sqrt{100}} = 21.51$   
 $\bar{X} \geq 22 + 1.96 \times \frac{2.5}{\sqrt{100}} = 22.49$

→  $\bar{X} \leq 21.51, \bar{X} \geq 22.49$ : 標本平均の両側5%棄却域

→  $\bar{X} = 22.35$ だったので、棄却域にない。即ち、帰無仮説を棄却できない

→ この会社の社員について肥満の程度に問題があるとは言えない (注:「問題がない」と積極的にいえるわけではない)

### 母数に関する仮説検定

- 母平均の検定 [t検定] (母分散が未知の場合)
  - 例: 酒屋の不正疑惑
 

ある酒屋では酒の量をごまかして売っているという噂があったので実際に1合(=180cc)のお酒を5本買って調べてみた。

酒量(cc)	175	180	165	170	170
標本平均値	172.0cc				

標本平均値は172.0ccであり、180ccより8ccも少ないが、果たしてこの店は不当表示で訴えられるか? [出展:『なるほど統計学』p.125]

↓

[帰無仮説] 母平均  $\mu=180$   
[対立仮説] 母平均  $\mu < 180$   
とし、有意水準5%で片側検定(左側)。

$\bar{X} \rightarrow T = \frac{\bar{X} - \mu}{S/\sqrt{n-1}}$

確率変数Tは自由度  $n-1$  の  $t$  分布  $t_{\alpha}(n-1)$  に従う

## 母数に関する仮説検定

- 母平均の検定〔t検定〕(母分散が未知の場合)
  - 例: 酒屋の不正疑惑

$$T \leq -2.132: \text{片側5\%棄却域}$$

$$\Leftrightarrow T = \frac{\bar{X} - \mu}{S/\sqrt{n-1}} \leq -2.132$$

$$\Leftrightarrow \bar{X} \leq \mu - 2.132 \frac{S}{\sqrt{n-1}}$$

$$\Leftrightarrow \bar{X} \leq 180.0 - 2.132 \frac{5.099}{\sqrt{5-1}} = 174.6$$

→  $\bar{X} \leq 174.6$  : 標本平均の片側5%棄却域

→  $\bar{X} = 172.0$  だったので、棄却域にある。即ち、**帰無仮説は棄却される**

→ この酒屋は酒量をごまかしているらしい

t分布: 片側(下側)5%棄却域

## 母数に関する仮説検定

- 母平均の検定〔t検定〕(母分散が未知の場合)
  - 演習1: 空調システムの作動状況検査

設定温度を25℃とし、7日間室内温度測定し、このシステムが正しく動いているかどうか5%有意水準で両側検定せよ

24.2	25.3	26.2	25.7	24.4	25.1	25.6
------	------	------	------	------	------	------

〔出展:『統計学入門』p.241〕

〔帰無仮説〕  $\mu = 25.0$   
〔対立仮説〕  $\mu \neq 25.0$

標本平均:  $\bar{X} = 25.21$ , 標本分散:  $S^2 = 0.438$  (標本標準偏差:  $S = 0.662$ )  
 $t_{\pm 0.025}(6) = \pm 2.447$  より  
 $T \leq -2.447, T \geq 2.447$ : 両側5%棄却域

$$\Leftrightarrow T = \frac{\bar{X} - \mu}{S/\sqrt{n-1}} \leq -2.447, \frac{\bar{X} - \mu}{S/\sqrt{n-1}} \geq 2.447$$

$$\Leftrightarrow \begin{cases} \bar{X} \leq \mu - 2.447 \frac{S}{\sqrt{n-1}} = 25 - 2.447 \frac{0.662}{\sqrt{7-1}} = 24.34, \\ \bar{X} \geq \mu + 2.447 \frac{S}{\sqrt{n-1}} = 25 + 2.447 \frac{0.662}{\sqrt{7-1}} = 25.66 \end{cases}$$

$\bar{X} \leq 24.34, \bar{X} \geq 25.66$ : 標本平均の両側5%棄却域

空調システムが正しく動いていないとは言えない

有意水準5%で帰無仮説を棄却できない

## 母数に関する仮説検定

- 母平均の検定〔t検定〕(母分散が未知の場合)
  - 演習2: 補習授業の効果測定

英語の補習を行った後の試験成績は上がったか? 5%有意水準で効果を検定せよ。10名の対象学生に対する補習前後の得点差は下表

-1	3	4	5	3	0	7	4	2	-2
----	---	---	---	---	---	---	---	---	----

〔出展:『統計学入門』p.241〕

〔帰無仮説〕  $\mu = 0$   
〔対立仮説〕  $\mu > 0$

標本平均:  $\bar{X} = 2.50$ , 標本分散:  $S^2 = 7.050$  (標本標準偏差:  $S = 2.655$ )  
 $t_{0.025}(9) = 1.833$  より  
 $T \geq 1.833$ : 片側5%棄却域

$$\Leftrightarrow T = \frac{\bar{X} - \mu}{S/\sqrt{n-1}} \geq 1.833$$

$$\Leftrightarrow \bar{X} \geq \mu + 1.833 \frac{S}{\sqrt{n-1}} = 0 + 1.833 \frac{2.655}{\sqrt{10-1}} = 1.622$$

$\bar{X} \geq 1.622$  : 標本平均の片側5%棄却域

補習の効果はなかったとは言えない(あったらしい)

有意水準5%で帰無仮説は棄却される

## 母数に関する仮説検定

- 母平均の検定〔t検定〕(母分散が未知の場合)
  - 演習3: 今年の生徒は出来がよいか?

数学の定期試験で10人の採点を終えた所、平均が71点で昨年度(65.7点)より5.3点も良い! 今年の生徒は出来がよいのだろうか? 有意水準5%で検定せよ。

70	62	82	73	67	75	85	71	60	65
----	----	----	----	----	----	----	----	----	----

〔出展:『図解雑学 統計解析』p.202〕

〔帰無仮説〕  $\mu = 65.7$   
〔対立仮説〕  $\mu > 65.7$

標本平均:  $\bar{X} = 71.0$ , 標本分散:  $S^2 = 59.20$  (標本標準偏差:  $S = 7.694$ )  
 $t_{0.025}(9) = 1.833$  より  
 $T \geq 1.833$ : 片側5%棄却域

$$\Leftrightarrow T = \frac{\bar{X} - \mu}{S/\sqrt{n-1}} \geq 1.833$$

$$\Leftrightarrow \bar{X} \geq \mu + 1.833 \frac{S}{\sqrt{n-1}} = 65.7 + 1.833 \frac{7.694}{\sqrt{10-1}} = 70.40$$

$\bar{X} \geq 70.40$  : 標本平均の片側5%棄却域

今年の生徒は昨年より出来が悪いとは言えない(良いらしい)

有意水準5%で帰無仮説は棄却される

## 母数に関する仮説検定

### 母分散の検定 [ $\chi^2$ 検定]

例: 製品のばらつき検査  
 目標重量25kgの製品の重量のばらつきが大きいことがわかり修理した。修理後の製品を無作為抽出した結果が以下。修理前の分散が9kg<sup>2</sup>のとき、この製品は性能が向上したといえるか？

24   26   27   22   26

[出展:『なるほど統計学』p.132]

修理後は性能が向上したと考えられる  
 ⇒ 分散は小さくなったはず  
 ⇒ 片側検定(左側)  
 [帰無仮説]  $\sigma^2 = 9$   
 [対立仮説]  $\sigma^2 < 9$

確率変数 $\chi^2$ は  
 自由度n-1の  
 $\chi^2$ 分布に従う

標本平均:  $\bar{X} = 25.0$   
 標本分散:  $S^2 = 3.20$   
 (標本標準偏差:  $S = 1.789$ )

$$S^2 \rightarrow \chi^2 = \frac{nS^2}{\sigma^2}$$

## 母数に関する仮説検定

### 母分散の検定 [ $\chi^2$ 検定]

例: 製品のばらつき検査

$\chi^2 \leq 0.7107$ : 片側5%棄却域  
 $\Leftrightarrow \chi^2 = \frac{nS^2}{\sigma^2} \leq 0.7107$   
 $\Leftrightarrow S^2 \leq 0.7107 \frac{\sigma^2}{n}$   
 $\Leftrightarrow S^2 \leq 0.7107 \frac{9}{5} = 1.279$

$\chi^2$ 分布: 片側(下側)5%棄却域

→  $S^2 \leq 1.279$ : 標本分散の片側5%棄却域  
 →  $S^2 = 3.20$ だったので、棄却域にない。即ち、帰無仮説は棄却されない  
 → 修理後に性能があがったとは言えない

## 母数に関する仮説検定

### 母分散の検定 [ $\chi^2$ 検定]

演習: 小学校の知能テスト  
 ある小学校の入学時知能テストの結果は平均50、分散36だった。本年度入学児25名を無作為抽出したところ平均53、分散48だった。本年度入学児の揃い方は例年と違うか？ 有意水準5%で検定せよ  
 [出展:『統計学入門』p.242]

[帰無仮説]  $\sigma^2 = 36$   
 [対立仮説]  $\sigma^2 \neq 36$

標本分散:  $S^2 = 48$   
 $\chi^2_{0.975}(24) = 12.4012, \chi^2_{0.025}(24) = 39.3641$  より  
 $\chi^2 \leq 12.4012, \chi^2 \geq 39.3641$ : 両側5%棄却域  
 $\Leftrightarrow \chi^2 = \frac{nS^2}{\sigma^2} \leq 12.4012, \chi^2 = \frac{nS^2}{\sigma^2} \geq 39.3641$   
 $\Leftrightarrow \begin{cases} S^2 \leq 12.4012 \frac{\sigma^2}{n} = 12.4012 \frac{36}{25} = 17.8578, \\ S^2 \geq 39.3641 \frac{\sigma^2}{n} = 39.3641 \frac{36}{25} = 56.6843 \end{cases}$   
 $S^2 \leq 17.86, S^2 \geq 56.68$ : 標本分散の両側5%棄却域

本年度入学児は揃い方が例年と違うとは言えない  
 有意水準5%で帰無仮説は棄却されない

## Coffee Break!

### 両側検定と片側検定の使い分け

例: 母平均の両側検定      例: 母平均の片側検定(右側)

[帰無仮説]  $\mu = \mu_0$       [帰無仮説]  $\mu = \mu_0$   
 [対立仮説]  $\mu \neq \mu_0$       [対立仮説]  $\mu > \mu_0$

$\mu_0$ は既知の値

<両側検定を実施する理由>  
 母数の値がある目標値と等しいかどうかを調べたい

<片側検定を実施する理由>

- $\mu < \mu_0$ が起り得ない
- $\mu > \mu_0$ を積極的に見いだせばそれでよい
- 母数の大きさが理論的・経験的に予測される

例: 生産ラインの機械が正しく動いているか      例: 補習後の成績は上がった? (上がったと積極的に知りたい)  
 例: 今年の夏は寒い気がする。本当か? (「寒い」という経験から平均気温が低いことを予測)

## 2標本検定 two-sample test

- 母平均の差の検定
  - 2つの正規母集団について母平均の差の検定
    - 母分散が**既知**の場合 Z検定
    - 母分散が**未知**だが**等しい**場合 t検定
    - 母分散が**未知**で**等しくない**場合 ウェルチの検定  
 $\sigma_1^2 = \sigma_2^2 = \sigma$   
 $\sigma_1^2 \neq \sigma_2^2$
- 母分散の比の検定
  - 2つの正規母集団について母分散の比の検定 F検定

## 2標本検定 two-sample test

2つの母平均に「差がある」か「ない」かの検定

- 母平均の差の検定
  - 2つの正規母集団について母平均の差の検定

$N(\mu_1, \sigma_1^2)$

↓  $m$ 個 無作為抽出  
 $X_1, X_2, \dots, X_m$

標本平均:  $\bar{X} \sim N(\mu_1, \sigma_1^2/m)$

$N(\mu_2, \sigma_2^2)$

↓  $n$ 個 無作為抽出  
 $Y_1, Y_2, \dots, Y_n$

標本平均:  $\bar{Y} \sim N(\mu_2, \sigma_2^2/n)$

両側検定の場合 [帰無仮説]  $\mu_1 = \mu_2$  ← 2つの正規母集団の母平均に差がない  
 [対立仮説]  $\mu_1 \neq \mu_2$  ← 2つの正規母集団の母平均に差がある

片側検定の場合 [帰無仮説]  $\mu_1 = \mu_2$  ← 2つの正規母集団の母平均に差がない  
 [対立仮説]  $\mu_1 > \mu_2$  ← 2つの正規母集団の母平均に差がある (or  $\mu_1 < \mu_2$ )

## 2標本検定 two-sample test

- 母平均の差の検定
  - 2標本の標本平均の差の標本分布
 
$$\begin{cases} \bar{X} \sim N(\mu_1, \sigma_1^2/m) \\ \bar{Y} \sim N(\mu_2, \sigma_2^2/n) \end{cases} \Leftrightarrow \begin{cases} E(\bar{X} - \bar{Y}) = E(\bar{X}) - E(\bar{Y}) = \mu_1 - \mu_2 \\ V(\bar{X} - \bar{Y}) = V(\bar{X}) + (-1)^2 V(\bar{Y}) = V(\bar{X}) + V(\bar{Y}) = \frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n} \end{cases}$$

$$\Leftrightarrow \bar{X} - \bar{Y} \sim N(\mu_1 - \mu_2, \frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n})$$
  - この検定の例: 医薬の効果の検証
    - 患者を新薬を使った治療を行うグループ(処理群)とそれ以外(対照群)に分け、グループで結果に差があるかどうか(新薬の効果があるかどうか)を検定する。

## 2標本検定 two-sample test

- 母平均の差の検定 (母分散が**既知**のとき)
  - 標本平均の差の標本分布 Z検定

$$Z = \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\sigma_1^2/m + \sigma_2^2/n}} \sim N(0,1)$$

## 2標本検定 two-sample test

- 母平均の差の検定 (母分散が未知だが等しい  $\sigma_1^2 = \sigma_2^2 = \sigma^2$ )
  - 標本平均の差の標本分布
 
$$\bar{X} - \bar{Y} \sim N(\mu_1 - \mu_2, (\frac{1}{m} + \frac{1}{n})\sigma^2)$$
  - 合併分散 pooled variance
 
$$s^2 = \frac{(m-1)s_1^2 + (n-1)s_2^2}{m+n-2}$$
 ただし、 $s_1^2, s_2^2$  は不偏推定量
 
$$\begin{cases} s_1^2 = \frac{1}{m-1} \sum_{i=1}^m (X_i - \bar{X})^2, \\ s_2^2 = \frac{1}{n-1} \sum_{j=1}^n (Y_j - \bar{Y})^2 \end{cases}$$
  - 補足: 『合併分散は不偏推定量である』
 
$$\begin{aligned} (\because) E(s^2) &= E\left(\frac{(m-1)s_1^2 + (n-1)s_2^2}{m+n-2}\right) = \frac{(m-1)E(s_1^2) + (n-1)E(s_2^2)}{m+n-2} \\ &= \frac{(m-1)\sigma^2 + (n-1)\sigma^2}{m+n-2} = \sigma^2 \end{aligned}$$

## 2標本検定 two-sample test

- 母平均の差の検定 (母分散が未知だが等しい  $\sigma_1^2 = \sigma_2^2 = \sigma^2$ )
  - $$\begin{cases} Z = \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sigma\sqrt{1/m+1/n}} \sim N(0,1), \\ \chi^2 = \frac{(m+n-2)s^2}{\sigma^2} \sim \chi^2(m+n-2) \end{cases}$$
 (ただし、 $Z$ と $\chi^2$ は独立)
  - $$T = \frac{Z}{\sqrt{\chi^2 / (m+n-2)}}$$

$$= \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sigma\sqrt{1/m+1/n}} \bigg/ \sqrt{\frac{s^2}{\sigma^2}}$$

$$= \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{s\sqrt{1/m+1/n}} \sim t_{\alpha}(m+n-2)$$

## 2標本検定 two-sample test

- 母平均の差の検定 (母分散が未知だが等しい  $\sigma_1^2 = \sigma_2^2 = \sigma^2$ )
  - 例: 苗木の生長における2社の肥料の違い
 

ある苗木の生長について2社の肥料に違いがあるか?

A社の肥料	24.3	25.2	20.4	26.1	22.1	23.4	24.2	20.9	24.7	23.7	21.6	23.4	20.2
B社の肥料	21.3	19.4	22.3	17.2	18.3	20.3	21.4	23.6	21.1	21.3	20.3	19.5	

標本平均値 =  $\frac{A社の肥料}{23.092}$     不偏分散値 =  $\frac{A社の肥料}{3.579}$   
 $\frac{B社の肥料}{20.500}$      $\frac{B社の肥料}{3.029}$
  - 5%有意水準で両側検定
 

(帰無仮説)  $\mu_1 = \mu_2$  [対立仮説]  $\mu_1 \neq \mu_2$

$$s^2 = \frac{(m-1)s_1^2 + (n-1)s_2^2}{m+n-2} = \frac{12 \times 3.579 + 11 \times 3.029}{13+12-2} = 3.316$$

$$T = \frac{\bar{X} - \bar{Y}}{s\sqrt{1/m+1/n}} = \frac{4.500 - 5.389}{1.821\sqrt{1/13+1/12}} = 3.556$$

$T = 3.556 > 2.069 = t_{0.025}(23)$  より、仮説を棄却

## 2標本検定 two-sample test

- 母平均の差の検定 (母分散が未知で等しくない  $\sigma_1^2 \neq \sigma_2^2$ )
  - 標本平均の差の標本分布
    - どのように工夫しても、2つの母分散に拠らない統計量を作れない
    - ↓
    - 標本平均の差の正確な分布を求められない
    - ↓
    - 近似的に分布を求める[ウェルチの近似法]
  - $$\hat{T} = \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{s_1^2/m + s_2^2/n}} \sim t_{\alpha}(v^*)$$
  - ただし、 $v^*$ は  $v = \frac{(s_1^2/m + s_2^2/n)^2}{\frac{(s_1^2/m)^2}{m-1} + \frac{(s_2^2/n)^2}{n-1}}$  に最も近い整数

## 2標本検定 two-sample test

- 母平均の差の検定 (母分散が未知で等しくない  $\sigma_1^2 \neq \sigma_2^2$ )
  - 例: 2つの鉱山から採れる鉱石に含まれる物質の含有量  
各鉱山から採れる鉱石の物質含有量に差があるか?  
[出展:『パソコンによるデータマイニング』p.132]

第1鉱区	4.9	3.9	4.7	4.3	5.8	4.2	4.4	3.3	5.5	4.0	
第2鉱区	5.2	5.0	5.3	6.9	5.0	4.9	4.4	6.5	5.3		
標本平均値	第1鉱区 4.500				第2鉱区 5.389				不偏分散値	第1鉱区 0.564	第2鉱区 0.636

5%有意水準で両側検定  
 ([帰無仮説]  $\mu_1 = \mu_2$  [対立仮説]  $\mu_1 \neq \mu_2$ )

**ウェルチの検定**

$$T = \frac{\bar{X} - \bar{Y}}{\sqrt{s_1^2/m + s_2^2/n}} = \frac{4.500 - 5.389}{\sqrt{0.564/10 + 0.636/9}} = -2.493$$

$\nu = \frac{(s_1^2/m + s_2^2/n)^2}{\frac{(s_1^2/m)^2}{m-1} + \frac{(s_2^2/n)^2}{n-1}} = 16.517$  より, 自由度17のt分布を考える  
 $T = -2.493 < -2.110 = -t_{0.025}(17)$  より, 仮説を棄却

## 2標本検定 two-sample test

- 母分散の比の検定 [F検定]
  - 2つの正規母集団について母分散の比の検定

$N(\mu_1, \sigma_1^2)$

↓  $m$ 個 無作為抽出  
 $X_1, X_2, \dots, X_m$

 $\chi_1^2 = \frac{mS_1^2}{\sigma_1^2} \sim \chi^2(m-1)$

$N(\mu_2, \sigma_2^2)$

↓  $n$ 個 無作為抽出  
 $Y_1, Y_2, \dots, Y_n$

 $\chi_2^2 = \frac{nS_2^2}{\sigma_2^2} \sim \chi^2(n-1)$

両側検定の場合 [帰無仮説]  $\sigma_1^2 = \sigma_2^2$  ← 2つの正規母集団の母分散に差がない  
 [対立仮説]  $\sigma_1^2 \neq \sigma_2^2$  ← 2つの正規母集団の母分散に差がある

片側検定の場合 [帰無仮説]  $\sigma_1^2 = \sigma_2^2$  ← 2つの正規母集団の母分散に差がない  
 [対立仮説]  $\sigma_1^2 > \sigma_2^2$  ← 2つの正規母集団の母分散に差がある  
 (or  $\sigma_1^2 < \sigma_2^2$ )

## 2標本検定 two-sample test

- 母分散の比の検定 [F検定]
  - F分布とフィッシャーの分散比

$$\chi_1^2 = \frac{mS_1^2}{\sigma_1^2} \sim \chi^2(m-1)$$

$$\chi_2^2 = \frac{nS_2^2}{\sigma_2^2} \sim \chi^2(n-1)$$

(ただし  $\chi_1^2$  と  $\chi_2^2$  は独立)

$$F = \frac{\chi_1^2/m - 1}{\chi_2^2/n - 1} = \frac{mS_1^2/(m-1)\sigma_1^2}{nS_2^2/(n-1)\sigma_2^2}$$

フィッシャーの分散比  $= \frac{\sigma_2^2}{\sigma_1^2} \cdot \frac{s_1^2}{s_2^2} \sim F(m-1, n-1)$

自由度(m-1, n-1)のF分布に従う

F分布: 両側5%棄却域

## 2標本検定 two-sample test

- 母分散の比の検定 [F検定]
  - 例題: 2つの工作機械の製品バラツキ検査

機械 I	100.46	100.35	100.36	100.48	100.39	100.72	100.42	100.68	100.86
	100.57	100.59	100.46	100.32	100.46	100.72	100.62		
機械 II	100.33	100.12	100.35	100.89	100.90	100.31	100.46	100.12	100.43
	100.88	100.28	100.08	100.42	100.11	100.16	100.71	100.26	100.18
不偏分散値	機械 I 0.02475				機械 II 0.07752				

5%有意水準で両側検定  
 ([帰無仮説]  $\sigma_1^2 = \sigma_2^2$ , [対立仮説]  $\sigma_1^2 \neq \sigma_2^2$ )

$$F = \frac{\sigma_2^2}{\sigma_1^2} \cdot \frac{s_1^2}{s_2^2} = 1 \cdot \frac{0.02475}{0.07752} = 0.319$$

$$F_{0.975}(15,17) = \frac{1}{F_{0.025}(17,15)} = \frac{1}{2.723} = 0.367$$

より, 仮説は棄却される

## 2標本検定 two-sample test

- 母分散の比の検定 [F検定]
  - F分布表からのF値の求め方
    - F分布表はαが通常「0.1, 0.05, 0.25, 0.01, 0.005」のときの表のみ表示されていて、下側(左側)パーセント点は表示されていない。そこで...

$$F_{1-\alpha}(m, n) = \frac{1}{F_{\alpha}(n, m)}$$

であることを利用して求める。

- 例: 自由度(10, 13)のF分布の下側(左側)5%点
 
$$F_{0.95}(10, 13) = \frac{1}{F_{0.05}(13, 10)} = \frac{1}{2.671} = 0.374$$

## 適合度検定 goodness-of-fit test

- 適合度の $\chi^2$ 検定 [Chi-square goodness-of-fit test]
  - 仮定された理論上の確率分布に対し、標本から求められた度数が適合するか否かを検証する。
  - 例: さいころ投げ
    - さいころを300回投げると理論上は1~6の目が50回ずつ出る(さいころの目は一様分布となる)。実際に300回投げると...

賽の目	1	2	3	4	5	6	合計
観測度数	47	52	46	53	47	55	300

[帰無仮説] さいころの目の出方に偏りが無い(一様分布に従う)  
 [対立仮説] さいころの目の出方に偏りがある(一様分布に従わない)

適合度検定:  
母集団分布の従う確率分布について立てた帰無仮説を検証する

## 適合度検定 goodness-of-fit test

- 適合度の測定
  - 例: さいころ投げ
    - [帰無仮説] さいころの目の出方に偏りが無い(一様分布に従う)
    - [対立仮説] さいころの目の出方に偏りがある(一様分布に従わない)

	賽の目	1	2	3	4	5	6	合計
$f_i$	観測度数	47	52	46	53	47	55	300
$p_i$	理論確率	1/6	1/6	1/6	1/6	1/6	1/6	1
$np_i$	理論度数	50	50	50	50	50	50	300

K. Pearsonの適合度基準

$$\chi^2 = \sum_{i=1}^k \frac{(f_i - np_i)^2}{np_i}$$

(十分大きいnで)自由度k-1の $\chi^2$ 分布に従う

観測度数と理論度数の差(ずれ)のばらつきが許容できる以上に多き過ぎる場合、それは変だよな → 変な範囲を決めよう

$\chi^2$ 分布: 上側5%棄却域

注)  $f_1, f_2, \dots, f_k$ はk個の確率変数だが、 $f_1 + f_2 + \dots + f_k = n$ より、自由度は1減る

## 適合度検定 goodness-of-fit test

- 適合度の測定
  - 例: さいころ投げ
    - 自由度5の $\chi^2$ 分布の片側5%棄却域(右側)

$$\chi^2 \geq 11.0705 = \chi_{0.05}^2(5)$$

K. Pearsonの適合度基準

$$\chi^2 = \frac{(47-50)^2}{50} + \frac{(52-50)^2}{50} + \dots + \frac{(55-50)^2}{50} = 1.44$$

[参考: p値=0.92 > 0.05=有意水準]

→ Pearsonの適合度基準値(1.44)が棄却域( $\chi^2 \geq 11.0705$ )にない  
 → 帰無仮説は棄却できない

このさいころは、目の出方に偏りがあるとは言えない

## 適合度検定 goodness-of-fit test

- 適合度の $\chi^2$ 検定 [Chi-square goodness-of-fit test]

例: さいころ投げ その2

帰無仮説は棄却されたけど、ほんとにこのさいころはいかさまじゃない？ 奇数の目が出にくく、偶数の目が多く出ている気がするけど...

賽の目	1	2	3	4	5	6	合計
観測度数	47	52	46	53	47	55	300

【帰無仮説】 さいころの偶奇の出方に偏りが無い(一様分布に従う)  
 【対立仮説】 さいころの偶奇の出方に偏りがある(一様分布に従わない)

賽の目	奇数目	偶数目	
観測度数	140	160	300
理論確率	1/2	1/2	1
理論度数	150	150	300
差	-10	10	
自乗	100	100	200
$\chi^2$ 値	0.667	0.667	1.333 > ? 3.841

参考:  $p$ 値 = 0.248213 < ? 0.05  
 よって、帰無仮説は棄却されない

## 適合度検定 goodness-of-fit test

- 適合度の $\chi^2$ 検定

演習: メンデルの法則[えんどう豆の形質遺伝]

表現型	黄色・丸い	黄色・しわ	緑色・丸い	緑色・しわ	計
観測度数	315	101	108	32	556
理論確率	9/16	3/16	3/16	1/16	1
理論度数	312.75	104.25	104.25	34.75	556
度数差	2.25	-3.25	3.75	-2.75	

【帰無仮説】 えんどう豆の表現型はメンデルの法則に適合している  
 【対立仮説】 えんどう豆の表現型はメンデルの法則に適合していない

$\chi^2 = 0.470 < 7.815 = \chi^2_{0.05}(3)$  ( $p$ 値 = 0.93 > 0.05 = 有意水準)

より、適合度基準値が片側5%棄却域にないので、帰無仮説は棄却できない

えんどう豆の表現型はメンデルの法則に適合しているようだ

## 独立性の検定 test of independence

- 独立性の $\chi^2$ 検定 [Chi-square test of independence]

2つの属性の間に関係があるかどうかを検証する

例: 支持政党に男女差があるか? [出展: 『図解雑学 統計解析』 p.224]

支持政党	A党	B党	C党	合計
男性	96	74	46	216
女性	54	46	44	144
合計	150	120	90	360

支持政党に男女差があるか?  
 それとも、差がない(支持政党は男女差に独立)?

## 独立性の検定 test of independence

- 独立性の測定

観測度数と理論度数との差が非常に大きいなら、2つの間には差がある(独立でない)ってことだよ

例: 支持政党に男女差があるか?  
 【帰無仮説】 支持政党に男女差がない → 男女共全体比率に等しい  
 【対立仮説】 支持政党に男女差がある

観測度数	支持政党	A党	B党	C党	合計
男性	96	74	46	216	
女性	54	46	44	144	
合計	150	120	90	360	

理論度数	支持政党	A党	B党	C党	合計
男性	90	72	54	216	
女性	60	48	36	144	
合計	150	120	90	360	

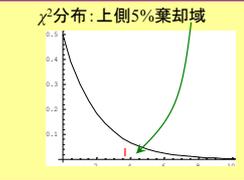
独立性の基準

$\chi^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(f_{ij} - g_{ij})^2}{g_{ij}}$

例) 男性合計の216を合計の党支持率の比 150:120:90で分配し、90:72:54となる

## 独立性の検定 test of independence

- 独立性の $\chi^2$ 検定
- 例：支持政党に男女差があるか？  
自由度2の $\chi^2$ 分布の片側5%棄却域(右側)  
 $= (3-1)(2-1)$   
 $\chi^2 \geq 5.99146 = \chi_{0.05}^2(2)$



独立性の基準  

$$\chi^2 = \frac{(96-90)^2}{\frac{90}{54-60}} + \frac{(74-72)^2}{\frac{72}{46-48}} + \frac{(46-54)^2}{\frac{54}{44-36}}$$

$$= \frac{36}{60} + \frac{4}{48} + \frac{64}{36} = 4.102$$
 [参考:  $p$ 値=0.13 > 0.05=有意水準]

→ 独立性の基準値(4.102)が棄却域( $\chi^2 \geq 5.99146$ )にない  
 → 帰無仮説は棄却できない

支持政党に男女間で差があるとはいえない

## 独立性の検定 test of independence

- 独立性の $\chi^2$ 検定
- 演習：期末試験成績：代数と解析の成績に関係があるか？  
[出展:『統計学入門』p.248]

観測度数	代数\解析			合計	理論度数	代数\解析			合計
A	4	2	3	9	A	3.86	1.93	3.21	9
B	8	4	6	18	B	7.71	3.86	6.43	18
C	6	3	6	15	C	6.43	3.21	5.36	15
合計	18	9	15	42	合計	18	9	15	42

〔帰無仮説〕 代数と解析の成績に関係がない  
 〔対立仮説〕 代数と解析の成績に関係がある

$\chi^2 = 0.19 < 9.488 = \chi_{0.05}^2(4)$  [  $p$ 値=0.996 > 0.05=有意水準 ]  
 より、独立性基準値が片側5%棄却域にないので、帰無仮説は棄却できない

代数と解析の成績には関係があるとはいえない

## 独立性の検定 test of independence

- 独立性の $\chi^2$ 検定
- 例題：男性・女性各看護師に対する恋愛についてのアンケート  
(結果に男女差があるか?) [出展:『マンガでわかるナースの統計学』p.111]

観測度数	好きな異性を前にすると緊張する?			理論度数	好きな異性を前にすると緊張する?		
男性看護師	20	34	54	男性看護師	14.14	39.86	54
女性看護師	35	121	156	女性看護師	40.86	115.14	156
計	55	155	210	計	55	155	210
男	37.0%	63.0%	100.0%	$\chi^2 = 4.42 < 3.84 = \chi_{0.05}^2(1)$ $\chi^2$ 乗値が棄却域にあるので、帰無仮説は棄却される (男女間に差がある→女性の方がより緊張する)			
女	22.4%	77.6%	100.0%				

$p$ 値: 0.035 > 0.05 有意水準

観測度数	異性に対しては自分から話しかける方?			理論度数	異性に対しては自分から話しかける方?		
男性看護師	25	29	54	男性看護師	34.46	19.54	54
女性看護師	109	47	156	女性看護師	99.54	56.46	156
計	134	76	210	計	134	76	210
男	48.3%	53.7%	100.0%	$\chi^2 = 9.65 > 3.84 = \chi_{0.05}^2(1)$ $\chi^2$ 乗値が棄却域にあるので、帰無仮説は棄却される (男女間に差がある→女性の方が異性に話し掛けない)			
女	69.9%	30.1%	100.0%				

$p$ 値: 0.002 < 0.05 有意水準

## 適合度検定・独立性の検定

- 演習1：サイコロの目の出方に偏りがあるか？ 有意水準5%で検定せよ。

賽の目	1	2	3	4	5	6
観測度数	40	65	45			

- 演習2：コーヒーの嗜好に違いがあるか？ 有意水準5%で検定せよ。

モカ\キリマンジャロ	大好き	好き	嫌い	大嫌い	計
大好き	8	5	6	9	28
好き	8	5	9	5	27
嫌い	2	8	1	2	13
大嫌い	2	7	3	0	12
計	20	25	19	16	80

## 参考文献

- 東大教養学部統計学教室編 「統計学入門」東大出版会(1991)
- 東大教養学部統計学教室編 「自然科学の統計学」東大出版会(1992)
- 鈴木達三・高橋宏一「標本抽出の計画と方法」放送大学(1991)
- 永田靖「統計的方法のしくみ」日科技連(1996)
- 永田靖「サンプルサイズの決め方」朝倉書店(2003)
- 村上雅人「なるほど統計学」海鳴社(2002)
- 丹慶勝市「図解雑学 統計解析」ナツメ社(2003)
- 高橋信[著]・トレンドプロ[マンガ]「マンガでわかる統計学」オーム社(2004)
- 田久浩志ほか「マンガでわかるナースの統計学」オーム社(2006)