最尤推定に基づく検定

確率・統計 - 第 *12* 講

村田 昇

前回のおさらい

- 分散未知の正規分布を用いた検定
  - 平均値の検定
  - 平均値の差の検定
  - 分散の検定
  - 分散の比の検定

統計的仮説検定

- ある現象・母集団に対して仮定された仮説の真偽をデータに基づいて統計的に検証する方法
- 推定と大きく異なるのは，母集団の分布に対して何らかの仮説を考えるところ

検定における仮説

- **帰無仮説** H_0

  検定統計量の分布を予想するために立てる仮説
- **対立仮説** H_1

  "帰無仮説が誤っているときに起こりうるシナリオ" として想定する仮説

検定の基本的手続き

1. 仮説を立てる
2. 仮説のもとで **検定統計量** が従う **帰無分布** を調べる
3. 実際のデータから検定統計量の値を計算する
4. 計算された検定統計量の値が仮説が正しいときに十分高い確率で得られるかどうかを判断する
   - **棄却域** を用いる方法
   - p-**値** を計算する方法

検定の用語

- 仮説の判定
  - 帰無仮説を **棄却**: 帰無仮説は誤っていると判断すること
  - 帰無仮説を **受容**: 帰無仮説を積極的に棄却できないこと
- 検定の誤り
  - **第一種過誤**: "正しい帰無仮説を棄却する" 誤り
  - **第二種過誤**: "誤った帰無仮説を受容する" 誤り

- 検定の設計
  - サイズ: “第一種過誤が起きる確率” を小さく
  - 検出力: “第二種過誤が起きない確率” を大きく

有意水準と p-値

- **有意水準**

  第一種過誤が起きる確率 (サイズ) として許容する上限

- p-値 (**有意確率**): (検定統計量 T, 棄却域 R_α)

  検定統計量の値が棄却域に含まれる有意水準の最小値

  $$(p\text{-値}) = \min\{\alpha \in (0,1) | T \text{ が } R_\alpha \text{ に含まれる}\}$$

- 有意水準と p-値の関係

  p-値が有意水準未満のときに帰無仮説を棄却する

両側検定と片側検定

- 対立仮説によって棄却域の形は変わりうる
  - 2 つの薬の治験結果に対する仮説:
    * 古い薬 (高価) と新しい薬 (安価) の効能が変わらない
    * 古い薬に比べて新しい薬の効能が改善した
  - 両側検定:

    棄却域がある定数 $a < b$ によって

    $$(-\infty, a) \cup (b, \infty)$$

  - 片側検定

    棄却域がある定数 a によって

    $$(a, \infty) \qquad \text{(右片側検定)}$$
    $$(-\infty, a) \qquad \text{(左片側検定)}$$

平均値の検定

- 問題

  確率変数列の平均値が μ_0 と等しいか検定せよ.

  $$X_1, X_2, \ldots, X_n$$

- 検定問題

  $$X_i = \mu + \varepsilon_i, \quad i = 1, \ldots, n \qquad \varepsilon_i \sim \mathcal{N}(0, \sigma^2)$$

  を観測値の確率モデル (σ^2 は **未知**) とするとき

  $$H_0 : \mu = \mu_0 \quad \text{vs} \quad H_1 : \mu \neq \mu_0$$

- 平均と分散の推定量
  - 標本平均

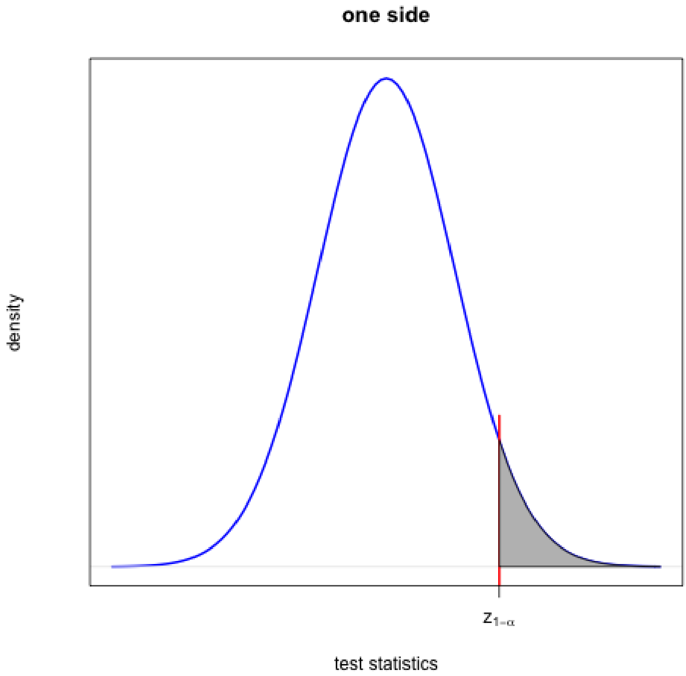

図 1: 片側検定の棄却域 (帰無分布の右の場合)

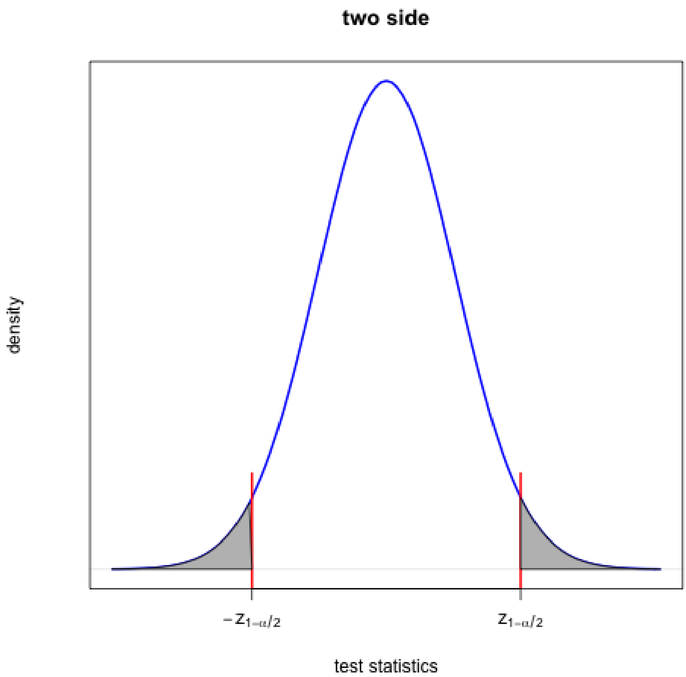

図 2: 両側検定の棄却域

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

– 不偏分散

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

- 帰無仮説のもとでの推定量の性質
  – 標本平均と不偏分散は互いに独立
  – 標本平均 (標準正規分布)

$$\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \sim \mathcal{N}(0,1)$$

  – 不偏分散 (χ^2-分布)

$$\frac{(n-1)s^2}{\sigma^2} \sim \chi^2(n-1)$$

- 検定統計量

$$T = \frac{\sqrt{n}(\bar{X} - \mu_0)}{s}$$

  の帰無分布は自由度 $n-1$ の t-分布
  – (t-分布) = (標準正規分布)/ $\sqrt{(\chi^2\text{-分布})/(自由度)}$
- 棄却域 (両側検定の場合)

$$R_\alpha = \bigl(-\infty, -t_{1-\alpha/2}(n-1)\bigr) \cup \bigl(t_{1-\alpha/2}(n-1), \infty\bigr)$$

平均値の差の検定

- 問題
  2 つの確率変数列の平均値が等しいか検定せよ.

$$X_1, X_2, \ldots, X_m, \qquad Y_1, Y_2, \ldots, Y_n$$

- 検定問題

$$\begin{aligned} X_i &= \mu_1 + \varepsilon_{1i}, \quad i = 1, \ldots, m & \varepsilon_{1i} &\sim \mathcal{N}(0, \sigma_1^2) \\ Y_j &= \mu_2 + \varepsilon_{2j}, \quad j = 1, \ldots, n & \varepsilon_{2j} &\sim \mathcal{N}(0, \sigma_2^2) \end{aligned}$$

  を観測値の確率モデル (σ_i^2 は **未知**) とするとき

$$H_0 : \mu_1 = \mu_2 \quad \text{vs} \quad H_1 : \mu_1 \neq \mu_2$$

- 検定統計量

  s_1^2, s_2^2 をそれぞれ X, Y の不偏分散として

  $$T = \frac{\bar{X} - \bar{Y}}{\sqrt{s_1^2/m + s_2^2/n}}$$

  の帰無分布は近似的に自由度 $\hat{\nu}$ の t-分布 (**Welch の近似**)

  $$\hat{\nu} = \frac{(s_1^2/m + s_2^2/n)^2}{(s_1^2/m)^2/(m-1) + (s_2^2/n)^2/(n-1)}$$

- 棄却域 (両側検定の場合)

  $$R_\alpha = \big(-\infty, -t_{1-\alpha/2}(\hat{\nu})\big) \cup \big(t_{1-\alpha/2}(\hat{\nu}), \infty\big)$$

分散の検定

- 問題

  確率変数列の分散が σ_0^2 と等しいか検定せよ.

  $$X_1, X_2, \dots, X_n$$

- 検定問題

  $$X_i = \mu + \varepsilon_i, \quad i = 1, \dots, n \qquad \varepsilon_i \sim \mathcal{N}(0, \sigma^2)$$

  を観測値の確率モデルとするとき

  $$H_0 : \sigma^2 = \sigma_0^2 \quad \text{vs} \quad H_1 : \sigma^2 \neq \sigma_0^2$$

- 検定統計量

  $$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2}$$

  の帰無分布は自由度 $n-1$ の χ^2-分布

- 棄却域 (両側検定の場合)

  $$R_\alpha = \Big(0, \chi^2_{\alpha/2}(n-1)\Big) \cup \Big(\chi^2_{1-\alpha/2}(n-1), \infty\Big)$$

分散の比の検定

- 問題

  2 つの確率変数列の分散が等しいか検定せよ.

  $$X_1, X_2, \dots, X_m, \qquad Y_1, Y_2, \dots, Y_n$$

- 検定問題

$$
\begin{aligned}
X_i &= \mu_1 + \varepsilon_{1i}, \quad i = 1, \ldots, m & \varepsilon_{1i} &\sim \mathcal{N}(0, \sigma_1^2) \\
Y_j &= \mu_2 + \varepsilon_{2j}, \quad j = 1, \ldots, n & \varepsilon_{2j} &\sim \mathcal{N}(0, \sigma_2^2)
\end{aligned}
$$

を観測値の確率モデルとするとき

$$
H_0 : \sigma_1^2 = \sigma_2^2 \quad \text{vs} \quad H_1 : \sigma_1^2 \neq \sigma_2^2
$$

- 検定統計量

  s_1^2, s_2^2 をそれぞれ X, Y の不偏分散として

  $$
  F = \frac{s_1^2}{s_2^2}
  $$

  の帰無分布は自由度 $m-1, n-1$ の F-分布

- 棄却域 (両側検定の場合)

  $$
  R_\alpha = \bigl(0, F_{\alpha/2}(m-1, n-1)\bigr) \cup \bigl(F_{1-\alpha/2}(m-1, n-1), \infty\bigr)
  $$

漸近正規性にもとづく検定

推定量の漸近正規性

- 漸近正規性 (データ数が多いときの性質)

  多くの推定量 $\hat{\theta}$ の分布は正規分布で近似できる

  - モーメントに基づく記述統計量は漸近正規性をもつ
  - 最尤推定量は広い範囲の確率分布に対して漸近正規性をもつ
  - いずれも中心極限定理にもとづく

- 信頼区間と同様に正規分布を用いて検定を考えることができる

最尤推定量の漸近正規性

- 定理

  $f(x) > 0$ が連続で 2 階微分可能ならば $\sqrt{n}(\hat{\theta}^* - \theta_0)$ は $n \to \infty$ で正規分布 $\mathcal{N}(0, I(\theta_0)^{-1})$ に近づく.

  - 観測データが十分多ければ，最尤推定量の誤差 (分散) は Cramer-Rao 下界に一致する
  - Fisher 情報量 (f は確率質量関数または確率密度関数)

  $$
  \begin{aligned}
  I(\theta_0) &= \mathbb{E}_{\theta_0}\left[-\frac{\partial^2}{\partial\theta^2}\log f(X; \theta_0)\right] \\
  &= \mathbb{E}_{\theta_0}\left[\left(\frac{\partial}{\partial\theta}\log f(X; \theta_0)\right)^2\right]
  \end{aligned}
  $$

最尤推定量の検定

- 問題

  θ_0 を既知の定数として，母数 θ が真の値 θ_0 であるか否かを検定する

  $$
  H_0 : \theta = \theta_0 \quad \text{vs} \quad H_1 : \theta \neq \theta_0
  $$

  - 上記は両側検定
  - 片側検定も同様に考えることができる

考え方

- 最尤推定量の性質

  観測データ数 n が十分大きいとき，1 次元母数 θ を含む連続分布の最尤推定量 $\hat{\theta}$ は

  $$\mathbb{E}[\hat{\theta}] = \theta_0, \quad \mathrm{Var}(\hat{\theta}) = \frac{1}{nI(\theta_0)}$$

  の正規分布で近似できる．

Z-検定 (正規分布による検定)

- 検定統計量

$$Z = \sqrt{nI(\theta_0)}(\hat{\theta} - \theta_0)$$

- 帰無分布は標準正規分布
- $z_{1-\alpha/2}$:

  標準正規分布の $1-\alpha/2$ 分位点

- 棄却域:

$$R_\alpha = \left(-\infty, -z_{1-\alpha/2}\right) \cup \left(z_{1-\alpha/2}, \infty\right)$$

例題 (McNemar 検定)

- 問題

  A 社と B 社の開発した 2 つの文字認識機械がある．n 個の文字に対してその性能を調べたところ

	1	2	3	...	n	
A 社	○	○	×	...	○	98.1%
B 社	×	○	○	...	○	98.0%

  のような正答率を示した．このとき A 社の機械は B 社より優れていると言えるだろうか？

- 試験した結果によって判断は異なるべき
  (n=10,000)

	A が正解	A が誤り
B が正解	9800	0
B が誤り	10	190

	A が正解	A が誤り
B が正解	9610	190
B が誤り	200	0

演習

練習問題

- 以下の問に答えよ

  A 社と B 社の開発した 2 つの文字認識機械がある．10,000 文字に対してその性能を調べたところ

	A が正解	A が誤り
B が正解	9500	180
B が誤り	220	100

  のような正答率を示した．このとき A 社の機械は B 社より優れていると言えるだろうか？

今回のまとめ

- 最尤推定量を用いた検定
    - 漸近正規性
    - 正規分布を用いた検定に帰着