確率・統計　後期　第5回

母平均の区間推定

稲積　泰宏（いなづみ　やすひろ）

今日の内容

- 区間推定とは
- 信頼区間の考え方
- 母分散が未知の場合（t分布）
- 母分散が既知の場合（正規分布）

前回の復習

点推定

母数を1つの値で推定する方法

例：母平均 μ を $\overline{x} = 50$ と推定

> **問題点**
>
> 推定値がどれくらい信頼できるか分からない

今回学ぶこと

区間推定で推定の精度を表現する

区間推定とは

定義

母数が含まれる範囲（区間）を推定する方法

例

母平均 μ は 95% の確率で 48〜52 の範囲にある

利点

- 推定の不確実性を表現できる
- 推定の信頼性が分かる

信頼区間と信頼係数

信頼区間

母数が含まれると期待される区間

信頼係数

その区間に母数が含まれる確率
記号：$1 - \alpha$（よく使われる値：0.95, 0.99）

> **95% 信頼区間**
>
> 100回推定したら、約95回は真の母数を含む区間

信頼区間の解釈

正しい解釈

同じ方法で何度も標本を取って信頼区間を作ると、
そのうち 95% の区間が真の母数を含む

注意

「真の母数が 95% の確率でこの区間にある」
という解釈は厳密には正しくない

母数は固定された値であり、確率変数ではない

母平均の区間推定の2つの場合

標本：$X_1, X_2, \ldots, X_n$ は正規分布 $N(\mu, \sigma^2)$ に従う

ケース1：母分散 σ^2 が未知

→ t 分布を使用（現実的な状況）

ケース2：母分散 σ^2 が未知（nが大きい）

→ 標準正規分布（Z 分布）を使用

ケース3：母分散 σ^2 が既知

→ 標準正規分布（Z 分布）を使用

ケース1：母分散が未知の場合

現実的な状況

通常、母分散 σ^2 は未知

→ 不偏分散 U^2 で推定する

問題

$$\frac{\overline{X} - \mu}{\sqrt{U^2/n}}$$

は標準正規分布に従わない

t 分布の登場

定義

母分散が未知のとき

$$T = \frac{\overline{X} - \mu}{\sqrt{U^2/n}} \sim t(n-1)$$

は自由度 $n-1$ の t 分布に従う

t 分布の性質

- 標準正規分布に似た形状
- 自由度が小さいと裾が厚い
- $n \to \infty$ で標準正規分布に近づく

t 分布と標準正規分布の比較

共通点

- 平均0、左右対称
- 釣鐘型の分布

相違点

- t 分布の方が裾が厚い（ばらつきが大きい）
- 自由度が大きくなると標準正規分布に近づく

信頼区間の構成（母分散未知）

信頼係数 $1-\alpha$ の信頼区間を求める

t 分布の上側 $\alpha/2$ 点を $t_{n-1}(\alpha/2)$ とすると

$$P\left(-t_{n-1}(\alpha/2) \leq \frac{\overline{X}-\mu}{\sqrt{U^2/n}} \leq t_{n-1}(\alpha/2)\right) = 1-\alpha$$

信頼区間の公式（母分散未知）

母平均 μ の信頼係数 $1-\alpha$ の信頼区間

$$\overline{x} - t_{n-1}(\alpha/2)\sqrt{u^2/n} \leq \mu \leq \overline{x} + t_{n-1}(\alpha/2)\sqrt{u^2/n}$$

ポイント

- t 分布の上側確率点を使用

例題1）

問3)

ケース2：nが大きいとき

前提

- 母集団：正規分布 $N(\mu, \sigma^2)$
- 母分散 σ^2 は未知
- 標本サイズ：nが大きい

信頼区間の構成（nが大きいとき）

信頼係数 $1-\alpha$ の信頼区間を求める

ステップ1

標準正規分布の上側 $\alpha/2$ 点を $z_{\alpha/2}$ とすると

$$P\left(-z_{\alpha/2} \leq \frac{\overline{X}-\mu}{\sqrt{U^2/n}} \leq z_{\alpha/2}\right) = 1-\alpha$$

信頼区間の構成（続き）

ステップ2

不等式を μ について解く

$$P\left(\overline{x} - z_{\alpha/2}\sqrt{\frac{U^2}{n}} \leq \mu \leq \overline{x} + z_{\alpha/2}\sqrt{\frac{U^2}{n}}\right) = 1 - \alpha$$

母平均 μ の信頼係数 $1-\alpha$ の信頼区間

$$\overline{x} - z_{\alpha/2}\sqrt{\frac{u^2}{n}} \leq \mu \leq \overline{x} + z_{\alpha/2}\sqrt{\frac{u^2}{n}}$$

よく使われる信頼係数と z 値

信頼係数	α	$z_{\alpha/2}$
95%	0.05	1.960
99%	0.01	2.576

95% 信頼区間では $z_{0.025} = 1.960$ を使用
99% 信頼区間では $z_{0.005} = 2.576$ を使用

例題2）

問4）

信頼区間の幅に影響する要因

信頼区間の幅

$$2 \times t_{n-1}(\alpha/2) \times \sqrt{\frac{u^2}{n}}$$

幅を狭くする方法

1. 標本サイズ n を大きくする → n が大きくなる
2. 信頼係数を下げる → $t_{n-1}(\alpha/2)$ が小さくなる

> 信頼係数を保ちながら精度を上げるには、標本サイズを増やす

信頼係数と区間の幅のトレードオフ

信頼係数	区間の幅
高い（99%）	広い
中程度（95%）	中程度

バランス

信頼性と精度のバランスを考えて信頼係数を選ぶ
実務では 95% がよく使われる

まとめ

区間推定

- 母数が含まれる範囲を推定
- 推定の不確実性を表現できる

母平均の信頼区間

- 母分散未知 → t 分布（通常はこちら）
- 母分散既知 → Z 分布（標準正規分布）

実務での注意点

- 通常は母分散未知なので t 分布を使用
- 標本サイズを大きくすると精度が向上

次回までに問題集187から189を解いておくこと

感想を会議のチャット欄へ