

信頼区間

道具としての統計学 - 7章

蘆田 誠人

目次

- 区間推定・信頼区間とは
- 信頼区間の例
- Rで実現
- まとめ

区間推定とは

- 標本の統計量を基に、母集団の平均などを、区間を持たせて推定する統計学手法
- 一部の標本から母平均を推定する際に用いられる

例)

100万本あるネジの長さの平均

→ 10本標本を取り、信頼区間を求める

信頼区間とは

- 母平均をある確率で含む範囲

(例) 信頼区間95%

→ 95%の確率で母平均を含む区間

- 確率が大きくなれば, 信頼区間も広くなる

信頼区間の公式

$$\text{信頼区間} = X \pm t \times \frac{SD}{\sqrt{n}}$$

X : 標本平均, t : t値, SD : 標本平均偏差, n : 標本数

信頼区間の例

問題

100万本あるネジの長さの平均を調べたい

手元には10本しかない

母平均の95%信頼区間を求めよ

ネジ10本の長さ(単位 : mm)

1	2	3	4	5	6	7	8	9	10
31	29	28	35	30	30	32	27	33	29

信頼区間の例

①標本平均を求める

$$\begin{aligned} X &= (31 + 29 + 28 + 35 + 30 + 30 + 32 + 27 + 33 + 29) / 10 \\ &= \underline{30.4} \end{aligned}$$

信頼区間の例

②偏差を求める

ネジ1本の長さと偏差の差

$$31 - 30.4 = \underline{0.6}, 29 - 30.4 = \underline{-1.4}$$

$$28 - 30.4 = \underline{-2.4}, 35 - 30.4 = \underline{4.6}$$

$$30 - 30.4 = \underline{-0.4}, 30 - 30.4 = \underline{-0.4}$$

$$32 - 30.4 = \underline{1.6}, 27 - 30.4 = \underline{-3.4}$$

$$33 - 30.4 = \underline{2.6}, 29 - 30.4 = \underline{-1.4}$$

信頼区間の例

③不偏分散を求める

偏差の二乗合計と標本数-1で割る

$$0.6^2 = 0.36, -1.4^2 = 1.96$$

$$-2.4^2 = 5.76, 4.6^2 = 21.16$$

$$-0.4^2 = 0.16, -0.4^2 = 0.16$$

$$1.6^2 = 2.56, -3.4^2 = 11.56$$

$$2.6^2 = 6.76, -1.4^2 = 1.96$$

$$\text{不偏分散} = 52.4 / (10 - 1) = \underline{5.822}$$

信頼区間の例

④標本標準偏差を求める

不偏分散の平方根を計算する

$$\text{標本標準偏差} = \sqrt{5.822} = \underline{\underline{2.413}}$$

信頼区間の例

⑤ t値を求める

自由度 = 標本数 - 1

95%信頼区間

t値 = 2.262

自由度	確率95%	確率99%	自由度	確率95%	確率99%
1	12.706	63.657	18	2.101	2.878
2	4.303	9.925	19	2.093	2.861
3	3.182	5.841	20	2.086	2.845
4	2.776	4.604	21	2.080	2.831
5	2.571	4.032	22	2.074	2.819
6	2.447	3.707	23	2.069	2.807
7	2.365	3.499	24	2.064	2.797
8	2.306	3.355	25	2.060	2.787
9	2.262	3.250	26	2.056	2.779
10	2.228	3.169	27	2.052	2.771
11	2.201	3.106	28	2.048	2.763
12	2.179	3.055	29	2.045	2.756
13	2.160	3.012	30	2.042	2.750
14	2.145	2.977	40	2.021	2.704
15	2.131	2.947	60	2.000	2.660
16	2.120	2.921	120	1.980	2.617
17	2.110	2.898	∞	1.960	2.576

信頼区間の例

⑥信頼区間を求める

標本平均 $\bar{X} = 30.4$, t 値 = 2.262 ,

標本標準偏差 $SD = 2.413$, 標本数 $n = 10$

$$\text{下限} = 30.4 - 2.262 \times \frac{2.413}{\sqrt{10}} = 28.674$$

$$\text{上限} = 30.4 + 2.262 \times \frac{2.413}{\sqrt{10}} = 32.126$$

95%信頼区間は28.674mm~32.126mm

Rで区間推定

- Rでは二行で実行できる

```
> neji = c(31,29,28,35,30,30,32,27,33,29)
> t.test(neji)

      One Sample t-test

data:  neji
t = 39.841, df = 9, p-value = 1.967e-11
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 28.6739 32.1261
sample estimates:
mean of x
      30.4
```

まとめ

- 信頼区間とは、母平均をある確率で含む範囲
- 確率を大きくすれば、信頼区間の範囲も広がる
- 公式を用いて求める方法やRを用いて求める方法がある