

母割合と母平均の推定について

高木 廣文
東邦大学医学部看護学科
国際保健看護学研究室

点推定と区間推定

- 点推定値
標本割合 → 母割合の点推定値
標本平均 → 母平均の点推定値

● 信頼区間は？

母割合(比率)の区間推定

- ・ 人数： N
- ・ ある特性をもつ人： x
 - ・ 標本比率： $p = x / N$
 - ・ 母集団の母割合(比率) π の点推定値
 - ・ p の標準誤差： $SE = \sqrt{p(1-p) / N}$

● 標準化： 平均 0， 分散 1 に変換

$$z = \frac{p - \pi}{SE}$$

標準正規分布と区間推定

z 値の分布は標準正規分布に従うと仮定：

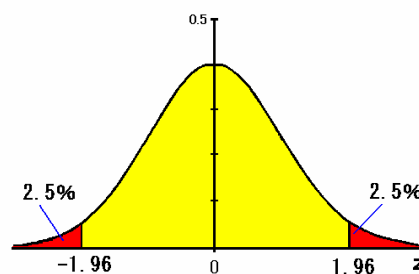


図 標準正規分布

母割合の信頼区間推定

母割合 π の95%信頼区間： $-1.96 < z < 1.96$



$$p - 1.96 \times SE < \pi < p + 1.96 \times SE$$

- ・ 99%信頼区間： 1.96 → 2.58

母割合の区間推定の例：

ある町の住民100人の血液型を無作為に調査したところ、12人がAB型であった。

この町の住民のAB型の割合の95%信頼区間を求めてみよう。

母割合の区間推定の例の計算

- ・ 標本比率 $p = 12/100 = 0.12$
- ・ 標準誤差 :

$$SE = \sqrt{0.12 \times 0.88/100} = 0.0325$$

● 母割合 π の95%信頼区間 :

$$0.12 - 1.96 \times 0.0325 < \pi < 0.12 + 1.96 \times 0.0325$$
$$0.054 = 0.12 - 0.066 < \pi < 0.12 + 0.066 = 0.186$$

(丸目の誤差がある)

母割合の区間推定—正確な推定

- 前述の信頼区間の推定式 :
 $p = 0$ または $p = 1$ の場合, 推定不可能
- ・ p の標準誤差 SE のより正確な式を用いた標準化を行う。

$$z = \frac{p - \pi}{\sqrt{\pi(1 - \pi) / N}}$$

母割合の2次式による区間推定

標準正規分布の上側確率 α 点 : z_α
100(1-2 α)%信頼区間の推定 :

$$-z_\alpha < \frac{p - \pi}{\sqrt{\pi(1 - \pi) / N}} < z_\alpha$$

2次式による信頼区間

$$(N + z_\alpha^2)\pi^2 + 2(pN + z_\alpha^2/2)\pi + Np^2 = 0$$

100(1-2 α)%信頼区間の上限 π_+ と下限 π_- :

$$\pi_\pm = \frac{p + z_\alpha^2/2N \pm z_\alpha \sqrt{p(1-p)/N + (z_\alpha/2N)^2}}{1 + z_\alpha^2/N}$$

2次式による信頼区間の例:

- 前述の血液型の例 :
標本数 : $N = 100$
標本割合 : $p = 12/100 = 0.12$
標準正規分布の上側2.5%点 : $z_{0.025} = 1.96$

$$\text{分母} = 1 + z_\alpha^2/N = 1 + 1.96^2/100 = 1.038416$$

$$\text{分子の前半} = p + z_\alpha^2/2N = 0.12 + 1.96^2/200$$
$$= 0.139208$$

2次式による信頼区間の例(2) :

$$\text{分子の後半} = z_\alpha \sqrt{p(1-p)/N + (z_\alpha/2N)^2}$$
$$= 1.96 \times \sqrt{0.12 \times 0.88/100 + (1.96/200)^2}$$
$$= 1.96 \times \sqrt{0.001056 + 0.00009604}$$
$$= 0.066525759$$

$$\pi_- = (0.139208 - 0.066525759) / 1.038416$$
$$= 0.070$$

$$\pi_+ = (0.139208 + 0.066525759) / 1.038416$$
$$= 0.198$$

母平均値の区間推定

- 母平均値の区間推定のための必要な用語 : 「標準化」, 「不偏分散」, 「t分布」
- 標準化 standardization
データ $x \rightarrow z$: 平均0, 分散1へと変換。
母集団 : 平均 μ , 分散 σ^2 の正規分布。

$$z = \frac{x - \mu}{\sigma}$$

標本平均値の標準化

- ・ 母集団 : 平均 μ , 分散 σ^2 の正規分布。
- ・ n 個のデータ x_1, x_2, \dots, x_n
- ・ 平均値 \bar{x} の期待値 = 母平均値 μ
- ・ 平均値の「標準誤差」 = σ / \sqrt{n}

平均値の標準化 :

$$z = \frac{\bar{x} - \mu}{(\sigma / \sqrt{n})}$$

母分散が既知の場合 :

- 母集団の母分散 σ^2 が分かっている場合 :
母平均値 μ の 100(1-2)% 信頼区間 :

$$\bar{x} - z_\alpha \times \frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + z_\alpha \times \frac{\sigma}{\sqrt{n}}$$

- ・ 通常は, 母分散 σ^2 は未知である。
母分散の推定をする必要がある。

母分散の点推定値—不偏分散

母分散の推定値 : 不偏分散 S^2 :

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$$

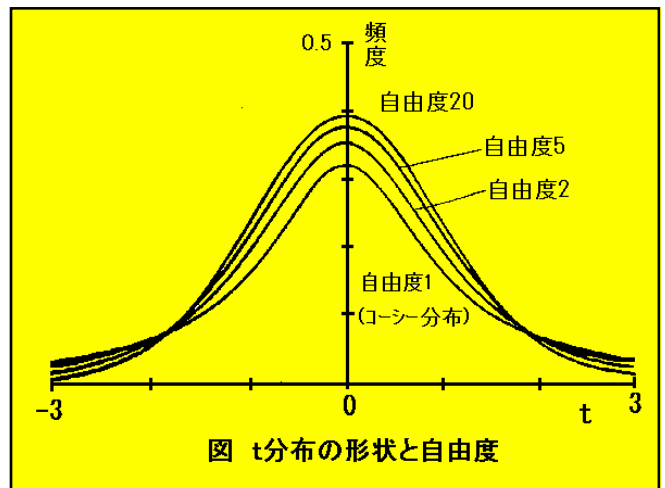
不偏分散による標準化—t分布

母分散 σ^2 のかわりに, 不偏分散 S^2 を使用 :

$$t = \frac{\bar{x} - \mu}{(S / \sqrt{n})}$$

Student (W. G. Gosset) の t 分布に従う。

- ・ 自由度 $n-1$ の t 分布 (Student の t 分布) に従う。
・ 標本数が大 \rightarrow 標準正規分布に近づく。



母平均値の信頼区間の推定

自由度 ν の t 分布の上側 α 点を $t_{\alpha}(\nu)$ 。
母平均値の $100(1-2\alpha)\%$ 信頼区間：

$$\begin{aligned} \text{下限} &: \bar{x} - t_{\alpha}(\nu) \times \frac{S}{\sqrt{n}} \\ \text{上限} &: \bar{x} + t_{\alpha}(\nu) \times \frac{S}{\sqrt{n}} \end{aligned}$$

計算には $t_{\alpha}(\nu)$ の値が必要。
統計数値表を見るかコンピュータで計算。

母平均値の信頼区間の例：

- 小学生123名のHDL-コレステロールを調べたところ、平均値は58.9mg/dl、不偏分散は220.8(mg/dl)²であった。
このときの、母平均値の95%信頼区間を求めよう。
なお、自由度122の t 分布の上側2.5点は1.980である。

母平均値の信頼区間の例の計算

$$\begin{aligned} \text{下限} &= 58.9 - 1.980 \sqrt{220.8/123} \\ &= 56.25 \end{aligned}$$

$$\begin{aligned} \text{上限} &= 58.9 + 1.980 \sqrt{220.8/123} \\ &= 61.55 \end{aligned}$$

- ・母平均値の95%信頼区間：
56.25 < 母平均 < 61.55