

教養科目／**B** 自然の摂理の探求

**2026** 年度 統計学（集中1期）

第10回

# 復習 推定法

母数: 母集団分布の特性値

- ・2項分布の母比率. (例 支持率)
- ・正規分布の母平均, 母分散. (例 身長の平均, 分散)

## ・点推定

母数を1つの値で推測する.

日本の首相の支持率は 75% である.

## ・区間推定

母数を区間で推測する.

日本の首相の支持率は 70% から 80% である  
確率は 90% である.

## 3.1. 2項分布についての点推定

母集団分布が2項分布  $B(1,p)$  の母比率  $p$  は標本平均で推定される.

母集団からの無作為標本の実現値  $x_1, x_2, \dots, x_n$  を用いて, 母比率  $p$  は

$$\hat{p} = \bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n}$$

と推定される.

## 3.2. 正規分布についての点推定

母集団分布: 正規分布  $N(\mu, \sigma^2)$

母集団からの無作為標本の実現値:  $x_1, x_2, \dots, x_n$

- ・母平均  $\mu$  は標本平均で推定する:

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n}$$

- ・母分散  $\sigma^2$  は不偏分散で推定する:

$$u_x^2 = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}{n-1}$$

### 3.3. 正規分布についての区間推定

階級	実験 1
168.0~168.5	0.002
168.5~169.0	0.024
169.0~169.5	0.078
169.5~170.0	0.136
170.0~170.5	0.166
170.5~171.0	0.174
171.0~171.5	0.186
171.5~172.0	0.126
172.0~172.5	0.072
172.5~173.0	0.024
173.0~173.5	0.008
173.5~174.0	0.002
174.0~174.5	0.002
合計	1

最初に得られた標本平均

$$\hat{\mu} = 169.8$$

区間

$$164.8 \sim 174.8$$

を考え、この区間に比較的大きい確率で母数  $\mu$  が含まれるだろうと考える。

$\mu$  の本当の値 : 171.0

### 3.3. 正規分布についての区間推定

母集団分布: 正規分布  $N(\mu, \sigma^2)$

母平均  $\mu$  を区間推定する.

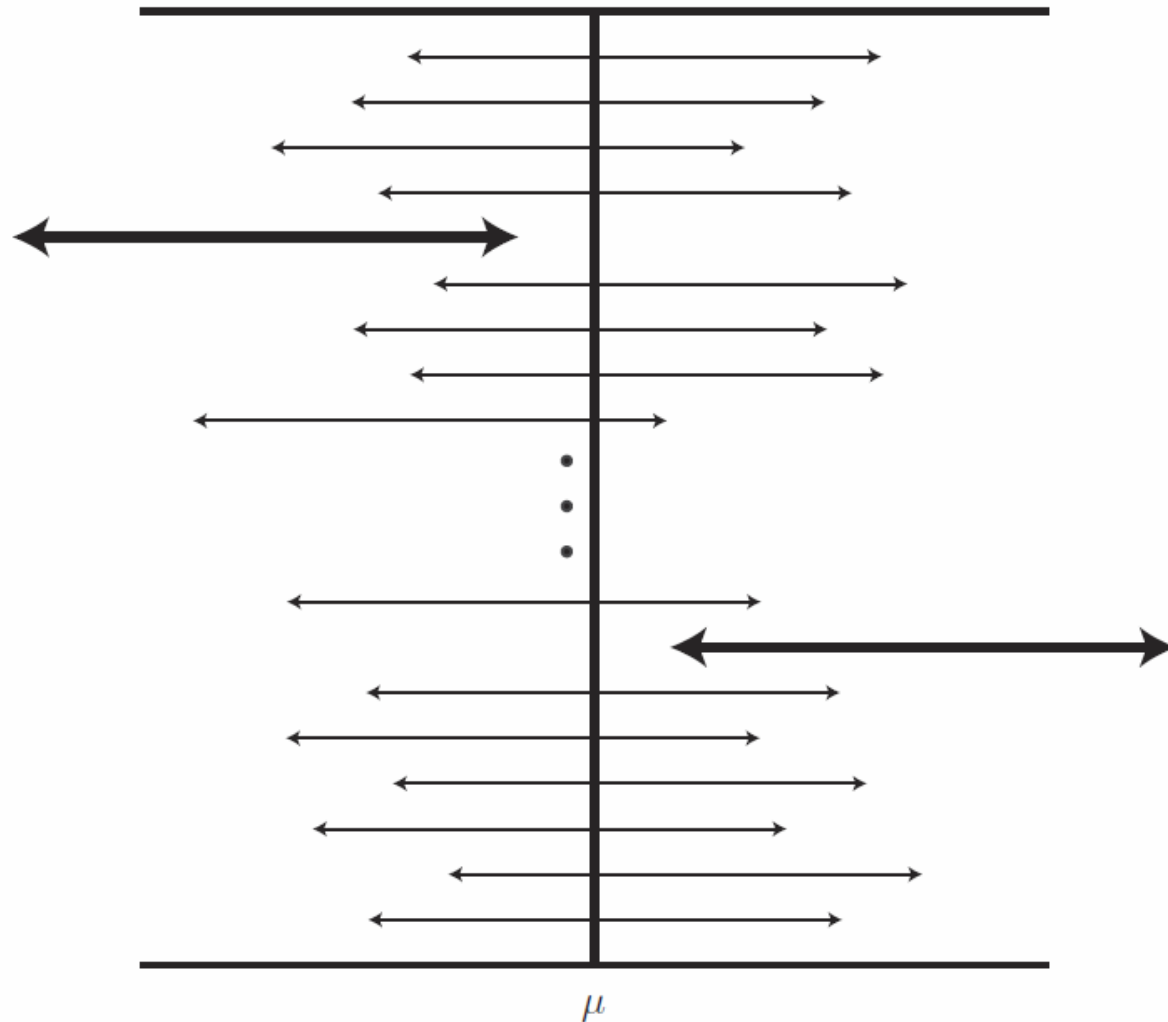
95% 信頼区間

$a < \mu < b$  あるいは  $a \sim b$  と表す.

ただし,  $a, b$  は無作為標本から決める.

95% を信頼係数あるいは信頼度という.

90%, 95%, 99% が用いられるが, 以下では, 95% の場合を中心に説明する.



95% 信頼区間を 100 個作ったら, そのうち 95個 (95%) が  $\mu$  を含んでいる.

95% 信頼区間を 100 個作ったら, そのうち 95 個 (95%) が  $\mu$  を含んでいる.

## 注意

95% 信頼区間は必ず 95% の割合で母平均  $\mu$  を含むわけではない. 100 個の 95% 信頼区間がすべて母平均  $\mu$  を含むこともあれば, 90 個しか  $\mu$  を含まないこともある.

### 3.3.1. 正規分布 $N(\mu, \sigma^2)$ の母平均 $\mu$ の区間推定 (母分散 $\sigma^2$ が既知の場合)

例 3.3 駅から大学までの距離(km) を 20 回測ってみた結果は、次のようであった。

1.01, 0.97, 1.03, 0.97, 1.03, 1.01, 1.04, 1.01, 1.02, 1.15,  
1.03, 1.00, 0.94, 0.99, 1.06, 1.02, 1.03, 1.01, 1.14, 1.01

測定結果は正規分布  $N(\mu, 0.2^2)$  に従っているものとする。

母平均  $\mu$  の推定値は

$$\text{標本平均} = \frac{1.01 + 0.97 + \dots + 1.01}{20} \doteq 1.0235.$$

母平均  $\mu$  を区間推定する.

標本平均を確率変数と考えると,

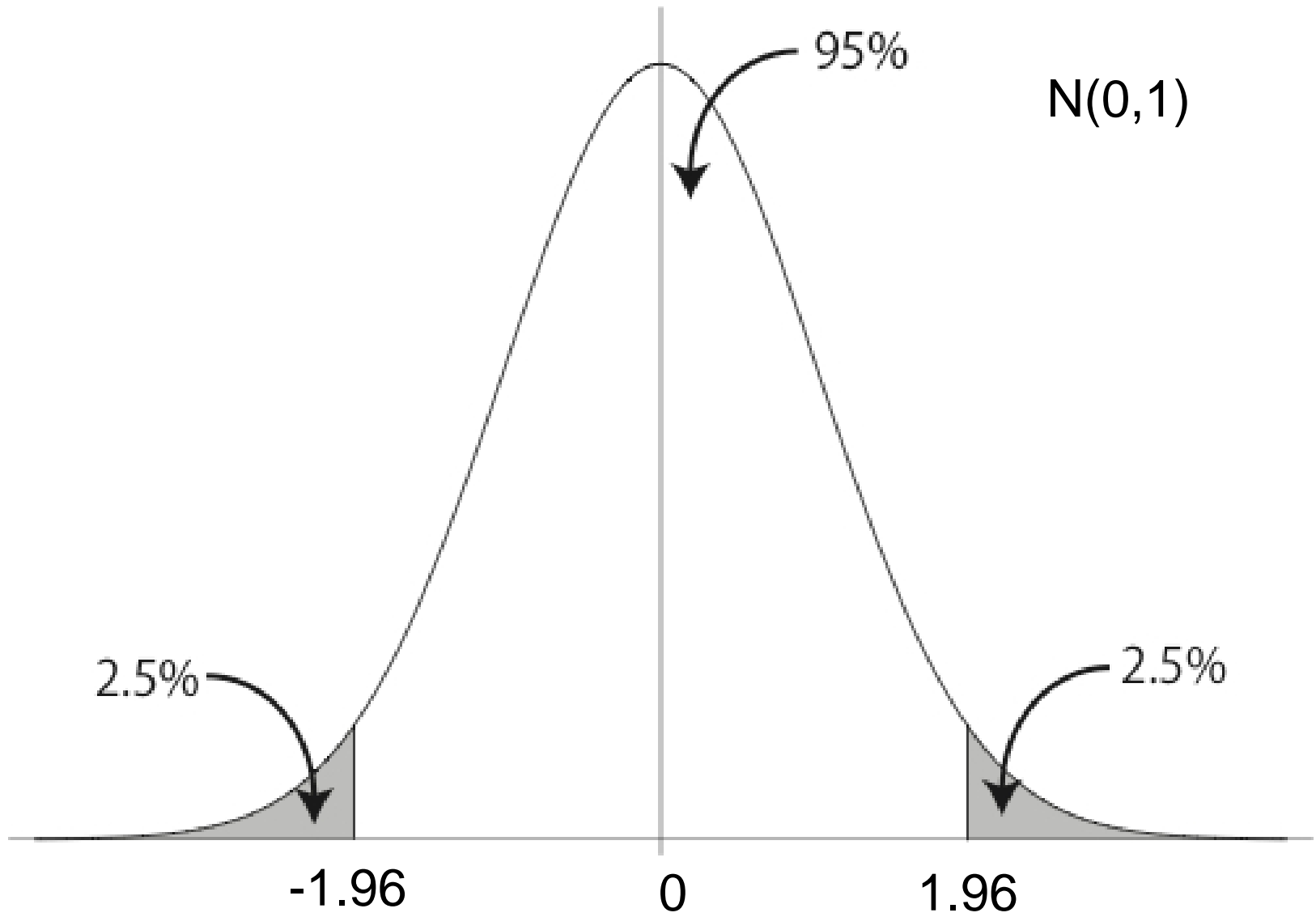
$$\sqrt{\frac{n}{\text{母分散}}} \times (\text{標本平均} - \text{母平均})$$

は標準正規分布  $N(0,1)$  に従うことが知られている.

確率 95% で

$$-1.96 < \sqrt{\frac{n}{\text{母分散}}} \times (\text{標本平均} - \text{母平均}) < 1.96$$

が成り立つ.



## 公式 3.1.

母平均  $\mu$  の 95% 信頼区間は

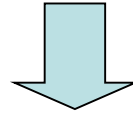
$$\text{標本平均} - 1.96 \times \sqrt{\frac{\text{母分散}}{n}} < \text{母平均} < \text{標本平均} + 1.96 \times \sqrt{\frac{\text{母分散}}{n}}$$

数式で書くと,

$$\bar{x} - 1.96\sqrt{\frac{\sigma^2}{n}} < \mu < \bar{x} + 1.96\sqrt{\frac{\sigma^2}{n}}$$

例 3.3 では,

$n = 20$ , 標本平均 = 1.0235, 母分散  $\sigma^2 = 0.2^2$



$$\text{標本平均} - 1.96 \times \sqrt{\frac{\text{母分散}}{n}} = 1.0235 - 1.96 \sqrt{\frac{0.2^2}{20}} \doteq 0.94$$

$$\text{標本平均} + 1.96 \times \sqrt{\frac{\text{母分散}}{n}} = 1.0235 + 1.96 \sqrt{\frac{0.2^2}{20}} \doteq 1.11$$

母平均  $\mu$  の 95% 信頼区間は 0.94 ~ 1.11

### 3.3.2. 正規分布 $N(\mu, \sigma^2)$ の母平均 $\mu$ の区間推定 (母分散 $\sigma^2$ が未知の場合)

例 3.4. ある花の種を蒔いてから発芽するまでの時間は正規分布  $N(\mu, \sigma^2)$  に従っているとす。10 粒の種を蒔いて発芽するまでの時間を測ったところ,

標本平均 = 163.77, 不偏分散 = 138.185

であった。

母平均  $\mu$  の 95% 信頼区間を求める。

データの個数が 10 の場合を考える.

標本平均, 不偏分散を確率変数と考えると,

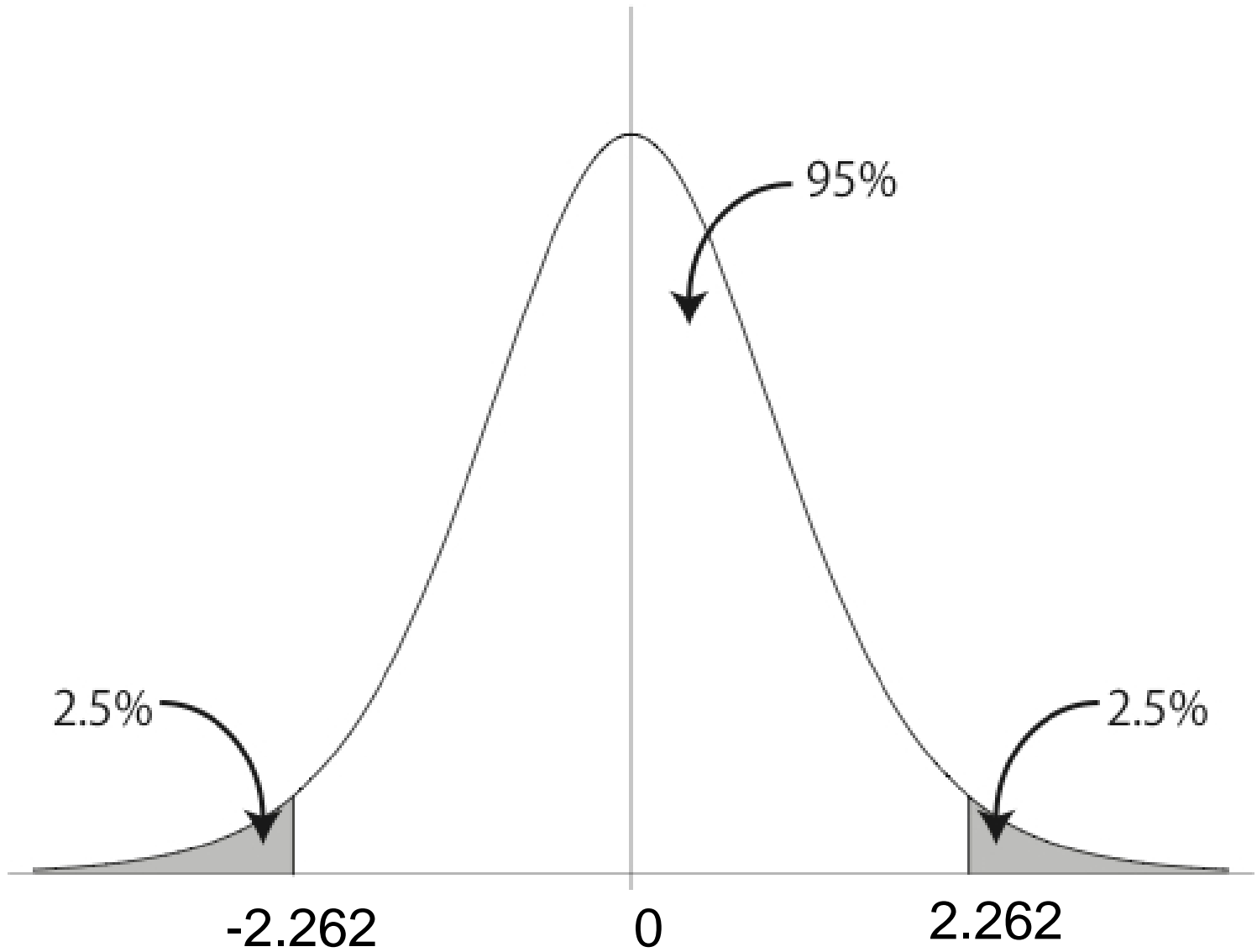
$$\sqrt{\frac{10}{\text{不偏分散}}} \times (\text{標本平均} - \text{母平均})$$

は自由度 9 の t 分布という分布に従うことが知られている.

確率 95% で

$$-2.262 < \sqrt{\frac{10}{\text{不偏分散}}} \times (\text{標本平均} - \text{母平均}) < 2.262$$

が成り立つ.



自由度 9 の t 分布の確率密度関数

確率 95% で

$$-2.262 < \sqrt{\frac{10}{\text{不偏分散}}} \times (\text{標本平均} - \text{母平均}) < 2.262$$

が成り立つ. この不等式を母平均について解くことにより

データの個数が 10 の場合, 母平均  $\mu$  の 95% 信頼区間

$$\text{標本平均} - 2.262 \times \sqrt{\frac{\text{不偏分散}}{10}} < \text{母平均} < \text{標本平均} + 2.262 \times \sqrt{\frac{\text{不偏分散}}{10}}$$

が導かれる.

注意: 2.262 はデータの個数によって変わる.

データの個数が  $n$  の場合, 母平均  $\mu$  の 95% 信頼区間

## 公式 3.2.

母平均  $\mu$  の 95% 信頼区間は

$$\text{標本平均} - t_{[n-1]}(0.025) \sqrt{\frac{\text{不偏分散}}{n}} < \text{母平均}$$

$$< \text{標本平均} + t_{[n-1]}(0.025) \sqrt{\frac{\text{不偏分散}}{n}}$$

ただし,  $t_{[n-1]}(0.025)$  は自由度  $n-1$  の  $t$  分布の上側 2.5% 点である.

数式で書くと,

$$\bar{x} - t_{[n-1]}(0.025) \sqrt{\frac{u_x^2}{n}} < \mu < \bar{x} + t_{[n-1]}(0.025) \sqrt{\frac{u_x^2}{n}}$$

ただし,  $\bar{x}$  は標本平均,  $u_x^2$  は不偏分散であり,

$t_{[n-1]}(0.025)$  は自由度  $n-1$  の  $t$  分布の上側 2.5% 点である.

上側 2.5% 点  $t_{[m]}(0.025)$

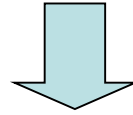
m	4	5	6	7	8
$t_{[m]}(0.025)$	2.776	2.571	2.447	2.365	2.306

m	9	10	11	12	13
$t_{[m]}(0.025)$	2.262	2.228	2.201	2.179	2.160

m	14	15	16	17	18
$t_{[m]}(0.025)$	2.145	2.131	2.120	2.110	2.101

例 3.4 では,

データの個数 = 10, 標本平均 = 163.77, 不偏分散 = 138.185



$$\text{標本平均} - 2.262 \times \sqrt{\frac{\text{不偏分散}}{10}} = 163.77 - 2.262 \sqrt{\frac{138.185}{10}} \doteq 155.4$$

$$\text{標本平均} + 2.262 \times \sqrt{\frac{\text{不偏分散}}{10}} = 163.77 + 2.262 \sqrt{\frac{138.185}{10}} \doteq 172.2$$

母平均  $\mu$  の 95% 信頼区間は 155.4 ~ 172.2

### 3.3.3. 正規分布 $N(\mu, \sigma^2)$ の母分散 $\sigma^2$ の区間推定 (母平均 $\mu$ が未知の場合)

例 3.5. ある花の種を蒔いてから発芽するまでの時間は正規分布  $N(\mu, \sigma^2)$  に従っているとす。10 粒の種を蒔いて発芽するまでの時間を測ったところ,

$$\text{不偏分散} = 138.185$$

であった。

母分散  $\sigma^2$  の 95% 信頼区間を求める

データの個数が 10 の場合を考える.

不偏分散を確率変数と考えると,

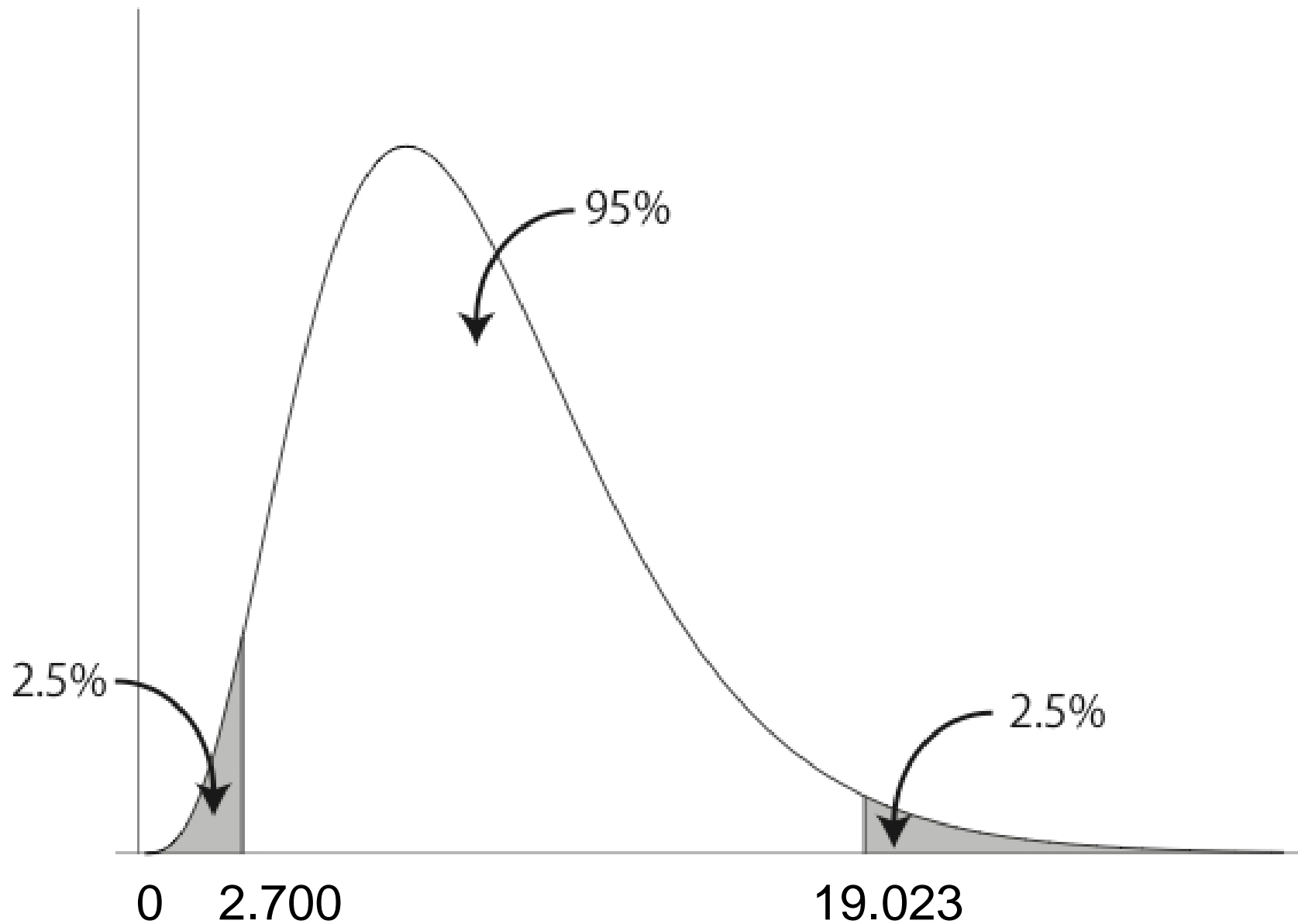
$$9 \times \frac{\text{不偏分散}}{\text{母分散}}$$

は自由度 9 のカイ2乗分布という分布に従うことが知られている.

確率 95% で

$$2.700 < 9 \times \frac{\text{不偏分散}}{\text{母分散}} < 19.023$$

が成り立つ.



自由度 9 のカイ2乗分布の確率密度関数

確率 95% で

$$2.700 < 9 \times \frac{\text{不偏分散}}{\text{母分散}} < 19.023$$

が成り立つ. この不等式を母分散について解くことにより

データの個数が 10 の場合, 母分散  $\sigma^2$  の 95% 信頼区間

$$\frac{9}{19.023} \times \text{不偏分散} < \text{母分散} < \frac{9}{2.700} \times \text{不偏分散}$$

が導かれる.

注意: 2.700, 19.023 はデータの個数によって変わる.

データの個数が  $n$  の場合, 母分散  $\sigma^2$  の 95% 信頼区間

### 公式 3.3.

母分散  $\sigma^2$  の 95% 信頼区間は

$$\frac{n-1}{\chi_{[n-1]}^2(0.025)} \times \text{不偏分散} < \text{母分散} < \frac{n-1}{\chi_{[n-1]}^2(0.975)} \times \text{不偏分散}$$

ただし,  $\chi_{[n-1]}^2(0.025)$  と  $\chi_{[n-1]}^2(0.975)$  はそれぞれ自由度  $n-1$  のカイ2乗分布の上側 2.5% 点, 下側 2.5% 点 (上側 97.5% 点) である.

数式で書くと,

$$\frac{n-1}{\chi_{[n-1]}^2(0.025)} u_x^2 < \sigma^2 < \frac{n-1}{\chi_{[n-1]}^2(0.975)} u_x^2$$

ただし,  $u_x^2$  は不偏分散であり,

$\chi_{[n-1]}^2(0.025)$  と  $\chi_{[n-1]}^2(0.975)$  はそれぞれ自由度  $n-1$  の

カイ2乗分布の上側 2.5% 点, 下側 2.5% 点 (上側 97.5% 点) である.

下側 2.5% 点  $\chi^2_{[m]}(0.975)$  と上側 2.5% 点  $\chi^2_{[m]}(0.025)$

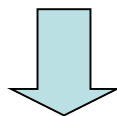
m	4	5	6	7	8
$\chi^2_{[m]}(0.975)$	0.484	0.831	1.237	1.690	2.180
$\chi^2_{[m]}(0.025)$	11.143	12.833	14.449	16.013	17.535

m	9	10	11	12	13
$\chi^2_{[m]}(0.975)$	2.700	3.247	3.816	4.404	5.009
$\chi^2_{[m]}(0.025)$	19.023	20.483	21.920	23.337	24.736

m	14	15	16	17	18
$\chi^2_{[m]}(0.975)$	5.629	6.262	6.908	7.564	8.231
$\chi^2_{[m]}(0.025)$	26.119	27.488	28.845	30.191	31.526

例 3.5 では,

データの個数 = 10, 不偏分散 = 138.185



$$\frac{9}{19.023} \times \text{不偏分散} = \frac{9}{19.023} \times 138.185 \doteq 65.4$$

$$\frac{9}{2.700} \times \text{不偏分散} = \frac{9}{2.700} \times 138.185 \doteq 460.6$$

母分散  $\sigma^2$  の 95% 信頼区間は 65.4 ~ 460.6

## 3.4. 2項分布についての区間推定

日本の首相の支持率調査では、母集団での支持率  $p$   
標本(1000人の有権者)での支持率 0.75 のとき

$p$  の 95% 信頼区間は？

# 母比率 $p$ の区間推定

データの個数が多いとき、中心極限定理より、標本平均を確率変数と考えると、

$$\sqrt{\frac{n}{\text{標本平均}(1-\text{標本平均})}} \times (\text{標本平均} - \text{母比率})$$

は近似的に標準正規分布  $N(0,1)$  に従う。

確率 95%で

$$-1.96 < \sqrt{\frac{n}{\text{標本平均}(1-\text{標本平均})}} \times (\text{標本平均} - \text{母比率}) < 1.96$$

が成り立つ

## 公式 3.4.

母比率  $p$  の 95% 信頼区間は

$$\text{標本平均} - 1.96 \sqrt{\frac{\text{標本平均} (1 - \text{標本平均})}{n}}$$

< 母比率 <

$$\text{標本平均} + 1.96 \sqrt{\frac{\text{標本平均} (1 - \text{標本平均})}{n}}$$

数式で書くと,

$$\bar{x} - 1.96 \sqrt{\frac{\bar{x}(1-\bar{x})}{n}} < p < \bar{x} + 1.96 \sqrt{\frac{\bar{x}(1-\bar{x})}{n}}$$

# 注意

・90%, 99% 信頼区間を求めたい場合には, 1.96 を

1.6449, 2.5758

に置き換える. 1.6449, 2.5758 はそれぞれ標準正規分布の上側 5% 点, 上側 0.5% 点

・近似の条件は

$$np \geq 5 \text{ かつ } n(1 - p) \geq 5$$

ただし, ここで考えている場合には  $p$  は未知母数なので,

$$n \times \text{標本平均} \geq 5 \text{ かつ } n(1 - \text{標本平均}) \geq 5$$

とする

## 例 3.6.

日本の首相の支持率調査で、ランダムに 1000 人の有権者を選び、そのうちの 750 人が支持で、250 人が不支持でした。

このとき、母支持率(母比率)  $p$  の 95% 信頼区間を求める。

$$n = 1000 \quad \text{標本平均} = 0.75$$

から、公式 3.4 を用いて、

$$\text{標本平均} - 1.96 \sqrt{\frac{\text{標本平均} (1 - \text{標本平均})}{n}} = 0.75 - 1.96 \sqrt{\frac{0.75(1-0.75)}{1000}} \doteq 0.72$$

$$\text{標本平均} + 1.96 \sqrt{\frac{\text{標本平均} (1 - \text{標本平均})}{n}} = 0.75 + 1.96 \sqrt{\frac{0.75(1-0.75)}{1000}} \doteq 0.78$$

母比率  $p$  の 95% 信頼区間は **0.72 ~ 0.78**

## 3.5. データの個数の決め方

- ・例 3.6 の母比率  $p$  の 95% 信頼区間は 0.72 ~ 0.78.
- ・もしランダムに選ぶ有権者の人数を 100 人とし、標本平均を 0.75 のままとすると、95% 信頼区間は 0.67 ~ 0.83.

どちらの信頼区間がいいでしょうか？

- ・信頼区間はその区間幅が小さければ小さいほど推定の精度が高くなり、より効果的.

(95% 信頼区間で、データの個数を大きくすると、信頼区間の幅は小さくなり、推定の精度が上がる.)

- ・データの個数を大きくするためには、コスト、労力等がかかる.

そこで、信頼区間の幅の目標値を設定し、その目標値を達成するために必要なデータの個数を決める.

### 3.5.1. 正規分布 $N(\mu, \sigma^2)$ の母平均 $\mu$ の区間推定 の場合 (母分散 $\sigma^2$ は既知)

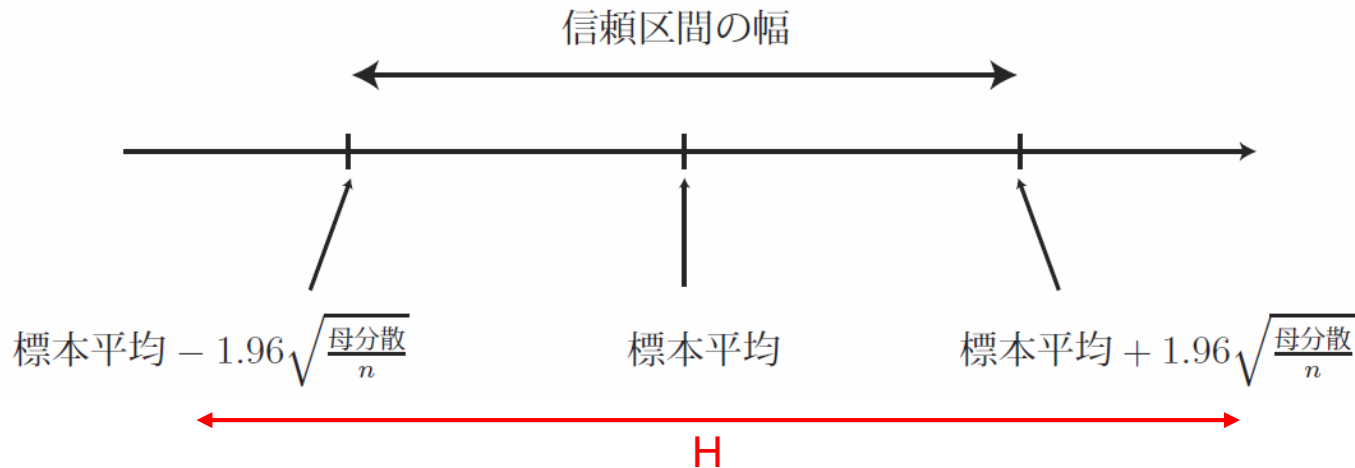
母平均  $\mu$  の 95% 信頼区間は

$$\text{標本平均} - 1.96 \times \sqrt{\frac{\text{母分散}}{n}} < \text{母平均} < \text{標本平均} + 1.96 \times \sqrt{\frac{\text{母分散}}{n}}$$

数式で書くと,

$$\bar{x} - 1.96\sqrt{\frac{\sigma^2}{n}} < \mu < \bar{x} + 1.96\sqrt{\frac{\sigma^2}{n}}$$

信頼区間の幅は  $2 \times 1.96 \times \sqrt{\frac{\text{母分散}}{n}}$



信頼区間の幅の**目標値を H** とし,

$$2 \times 1.96 \times \sqrt{\frac{\text{母分散}}{n}} \leq H$$

となるように n を決める:

$$n \geq \frac{(2 \times 1.96)^2 \times (\text{母分散})}{H^2}$$

例 3.7.

例 3.3 では、データの個数は 20, 母分散は  $0.2^2$ .  
また, 母平均  $\mu$  の 95% 信頼区間

$$0.94 \sim 1.11$$

で, 信頼区間の幅は 0.17.

この信頼区間の幅を  $H=0.15$  以下にするためには,

$$n \geq \frac{(2 \times 1.96)^2 \times (\text{母分散})}{H^2} = \frac{(2 \times 1.96)^2 \times (0.2)^2}{(0.15)^2} \doteq 27.3$$

となり, 必要なデータの個数は 28 となります.

## 3.5.2. 2項分布の母比率の区間推定の場合

母比率  $p$  の 95% 信頼区間は

$$\text{標本平均} - 1.96 \sqrt{\frac{\text{標本平均} (1 - \text{標本平均})}{n}}$$

< 母比率 <

$$\text{標本平均} + 1.96 \sqrt{\frac{\text{標本平均} (1 - \text{標本平均})}{n}}$$

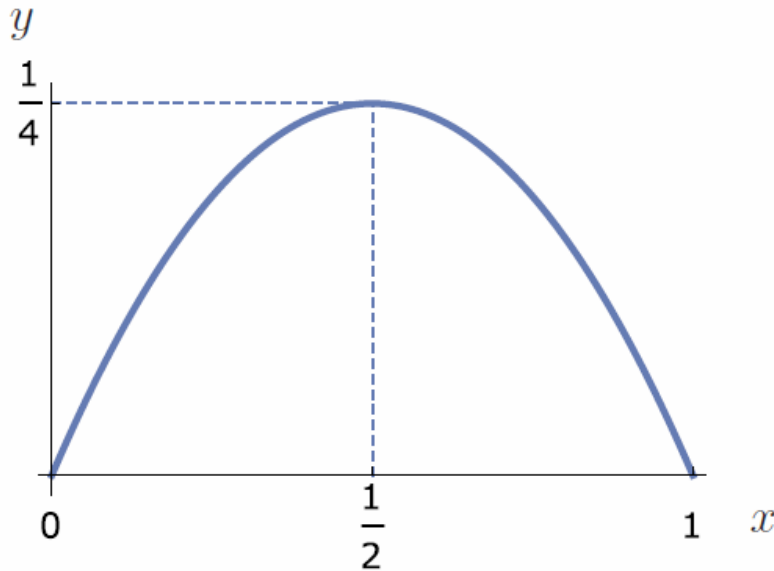
信頼区間の幅は

$$2 \times 1.96 \sqrt{\frac{\text{標本平均} (1 - \text{標本平均})}{n}}$$

## 関数

$$y = x(1 - x) = -\left(x - \frac{1}{2}\right)^2 + \frac{1}{4} \quad (0 \leq x \leq 1)$$

を考える.



$x = \frac{1}{2}$  のとき最大値  $\frac{1}{4}$

$0 \leq x \leq 1$  を満たす  $x$  に対して,  $x(1-x) \leq \frac{1}{4}$

$0 \leq x \leq 1$  を満たす  $x$  に対して,  $x(1-x) \leq \frac{1}{4}$

$0 \leq \text{標本平均} \leq 1 \implies \text{標本平均} (1 - \text{標本平均}) \leq \frac{1}{4}$

よって,

$$2 \times 1.96 \sqrt{\frac{\text{標本平均} (1 - \text{標本平均})}{n}} \leq \frac{1.96}{\sqrt{n}}$$

信頼区間の幅の目標値を  $H$  とし,

$$\left( 2 \times 1.96 \sqrt{\frac{\text{標本平均} (1 - \text{標本平均})}{n}} \leq \right) \frac{1.96}{\sqrt{n}} \leq H$$

となるようにデータの個数  $n$  を決める.

$$n \geq \left( \frac{1.96}{H} \right)^2$$

## 例 3.8.

ある商品のアンケート調査をし、満足するかどうかを尋ねることにしましょう.

満足率(母比率)の 95% 信頼区間を作ることとし、**信頼区間の幅を  $H=0.10$  以下にしたい**ときには、どれくらいのデータが必要でしょうか.

$$\text{データの個数} \geq \left( \frac{1.96}{0.10} \right)^2 \doteq 384.2$$

となり、必要なデータの個数は 385 となる.

## 3.6. 視聴率のはなし

2009年3月24日に行われたワールド・ベースボール・クラシックの決勝戦：韓国 vs 日本戦の視聴率（関東地区）は

	放送開始	分	番組平均世帯視聴率(%)
1回目	9:55	40	19.7
2回目	10:35	253	36.4
3回目	14:48	6	37.6

視聴率は世帯を対象とし、関東地区の調査世帯数は 600 です。  
番組平均世帯視聴率は毎分の視聴率の標本平均です。

1回目, 2回目, 3回目の視聴率(母比率)を  $p_1, p_2, p_3$  の 95% 信頼区間を求めてみましょう。

$p_1$  の 95% 信頼区間 0.165 ~ 0.229

$p_2$  の 95% 信頼区間 0.326 ~ 0.402

$p_3$  の 95% 信頼区間 0.337 ~ 0.415

これらの信頼区間より,

$p_1$  より  $p_2$  が大きい

$p_2$  と  $p_3$  はあまり変わらない

ように思われますが, そのことについては次章の検定で.

# まとめ(1/2)

母集団分布: 正規分布  $N(\mu, \sigma^2)$

母平均  $\mu$  の 95% 信頼区間

母分散  $\sigma^2$  が既知の場合 (公式 3.1)

母分散  $\sigma^2$  が未知の場合 (公式 3.2)

母分散  $\sigma^2$  の 95% 信頼区間

母平均  $\mu$  が未知の場合 (公式 3.3)

# まとめ(2/2)

## 母集団分布: 2項分布 $B(1,p)$

- ・母比率  $p$  の 95% 信頼区間(公式 3.4)

## データの個数の決め方

- ・正規分布  $N(\mu, \sigma^2)$  の母平均  $\mu$  の区間推定の場合  
(ただし, 母分散  $\sigma^2$  は既知)
- ・2項分布  $B(1,p)$  の母比率  $p$  の区間推定の場合