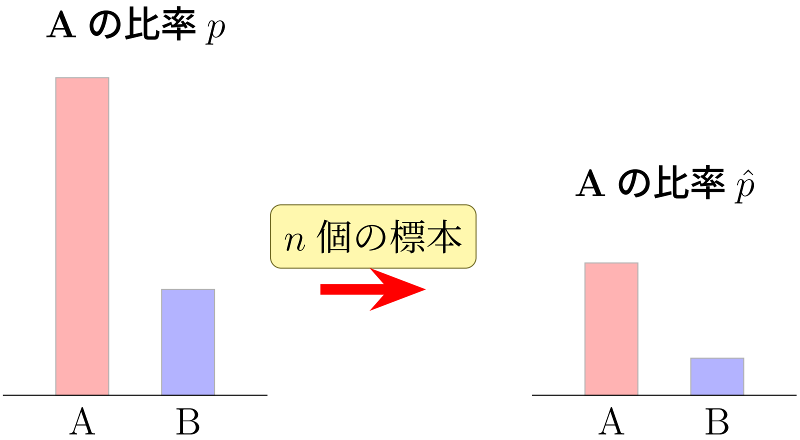
ある県で、800 世帯を無作為抽出して、A という意見の賛否を調べたところ、500 世帯が賛成であった。
全世帯における賛成の母比率 p を信頼度 95% で推定せよ。

math-support.jp

今回の学習目標

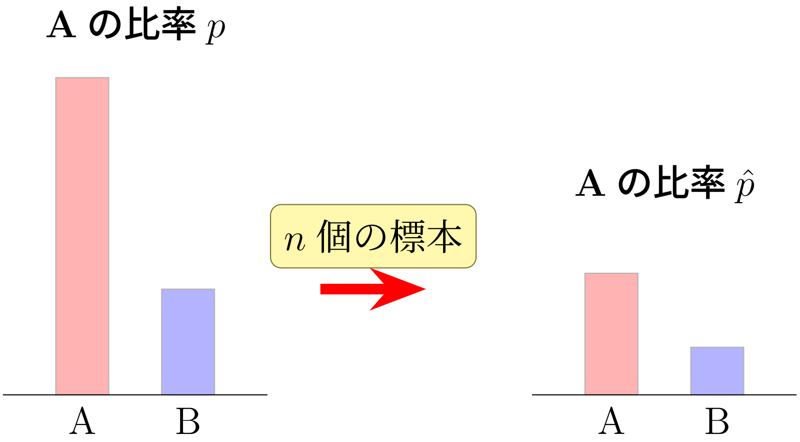
標本の比率から母比率を推定する。

政権を支持しているかどうかを問うような場合、母集団は「A：支持」「B：不支持」の２つに分かれる。母集団の「A」の比率を求めるには、n 個の標本を取り出し、母集団の比率を推定する。

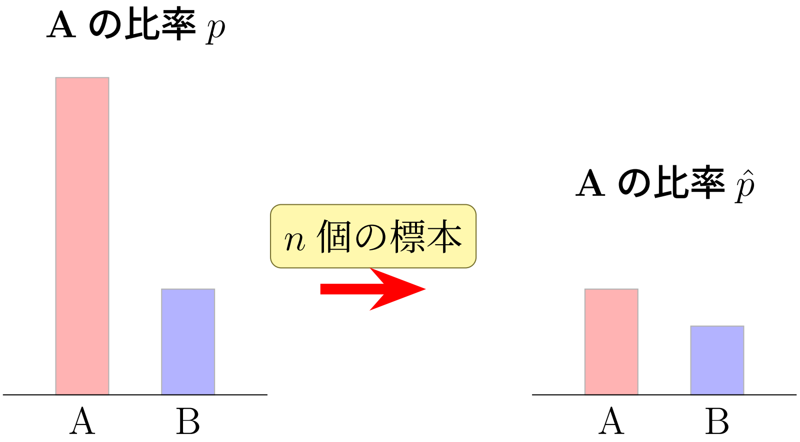


math-support.jp

何度も標本をとっていくと、標本比率 $\hat{p}$ は毎回違う値をとる。中心極限定理によると、母集団の分布がいかなるものでも、ある程度の大きさの標本を取り出せば、それは正規分布になる。

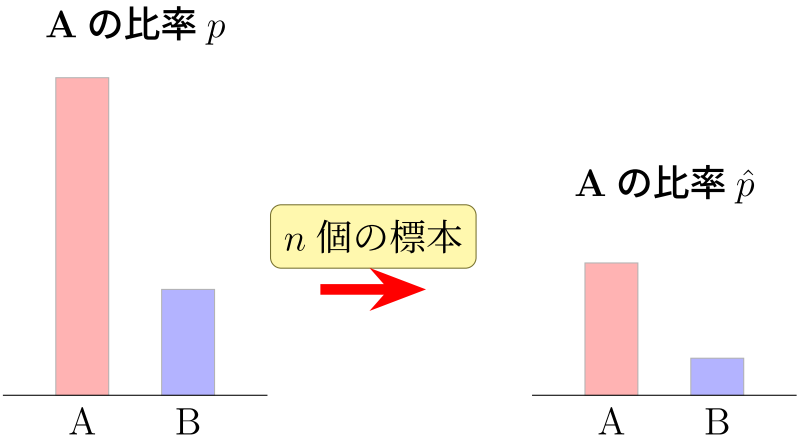


何度も標本をとっていくと、標本比率 $\hat{p}$ は毎回違う値をとる。中心極限定理によると、母集団の分布がいかなるものでも、ある程度の大きさの標本を取り出せば、それは正規分布になる。



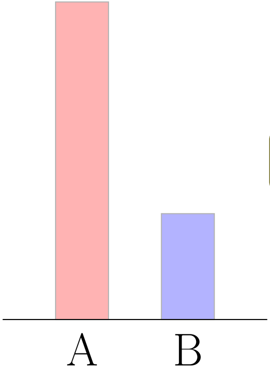
math-support.jp

何度も標本をとっていくと、標本比率 $\hat{p}$ は毎回違う値をとる。中心極限定理によると、母集団の分布がいかなるものでも、ある程度の大きさの標本を取り出せば、それは正規分布になる。



math-support.jp

A の比率 p

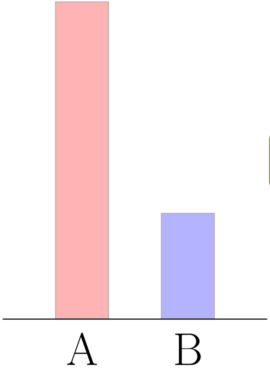


n 個の標本

標本比率 $\hat{p}$ の分布

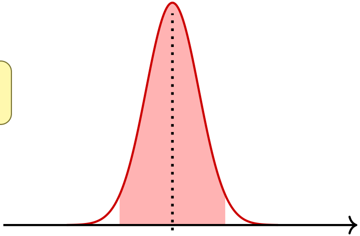
math-support.jp

A の比率 p



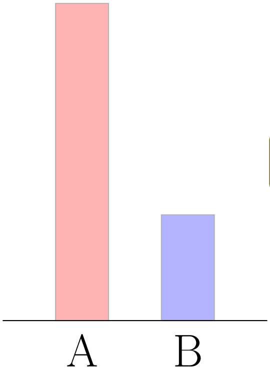
n 個の標本

標本比率 $\hat{p}$ の分布

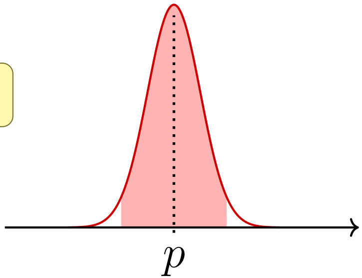


math-support.jp

A の比率 p



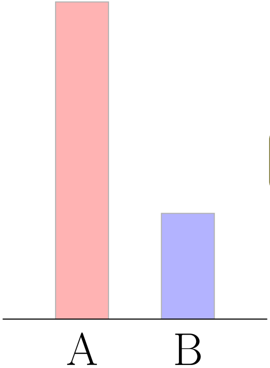
標本比率 $\hat{p}$ の分布



標本比率の平均：$\mathrm{E}(\hat{p}) = p$

math-support.jp

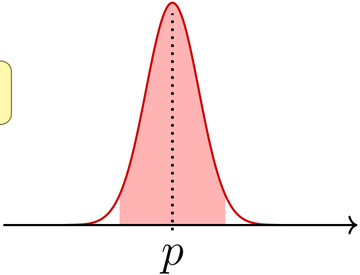
A の比率 p



母分散：$\sigma^2 = p(1-p)$

n 個の標本

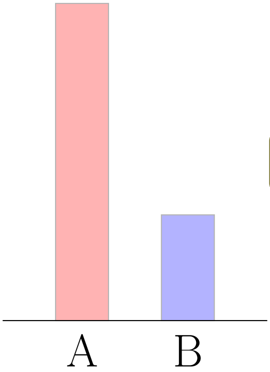
標本比率 $\hat{p}$ の分布



標本比率の平均：$\mathrm{E}(\hat{p}) = p$

math-support.jp

A の比率 p

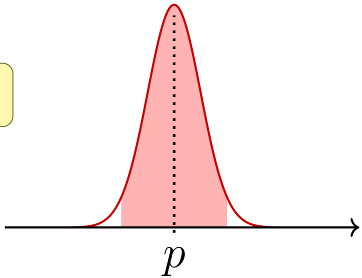


母分散：$\sigma^2 = p(1-p)$

母標準偏差：$\sigma = \sqrt{p(1-p)}$

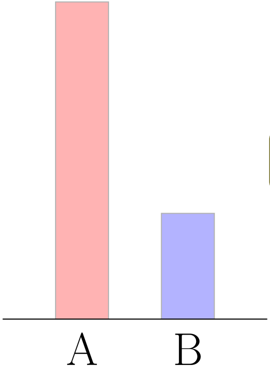
n 個の標本

標本比率 $\hat{p}$ の分布



標本比率の平均：$\mathrm{E}(\hat{p}) = p$

A の比率 p

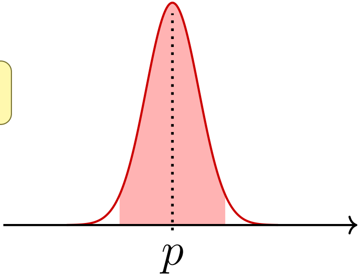


母分散：$\sigma^2 = p(1-p)$

母標準偏差：$\sigma = \sqrt{p(1-p)}$

n 個の標本

標本比率 $\hat{p}$ の分布



標本比率の平均：$\mathrm{E}(\hat{p}) = p$

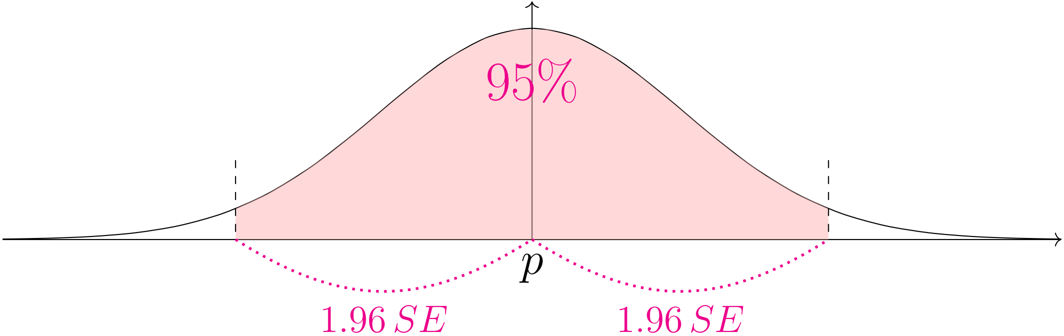
標本比率の標準偏差：$\sigma(\hat{p}) = \sqrt{\dfrac{p(1-p)}{n}}$

実際の計算では母比率 p は未知であるため、$\sigma(\hat{p})$ の式中の p を標本比率 $\hat{p}$ で代用します。

$$\sigma(\hat{p}) \approx \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

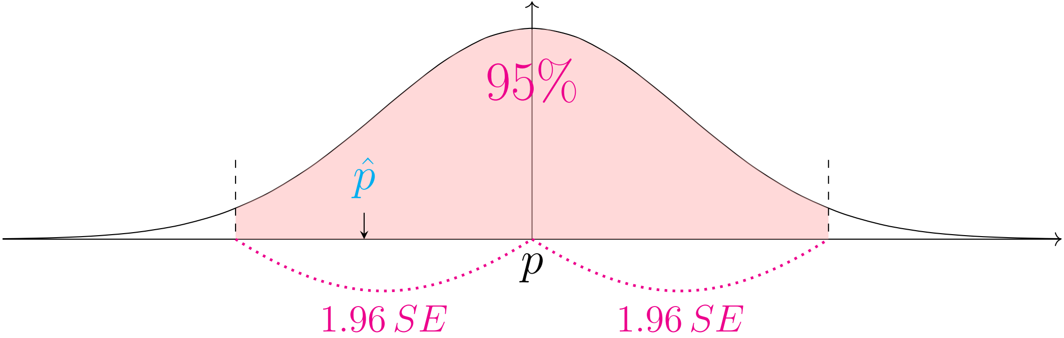
この値を**標準誤差**と呼び、SE (Standard Error) と表します。

math-support.jp



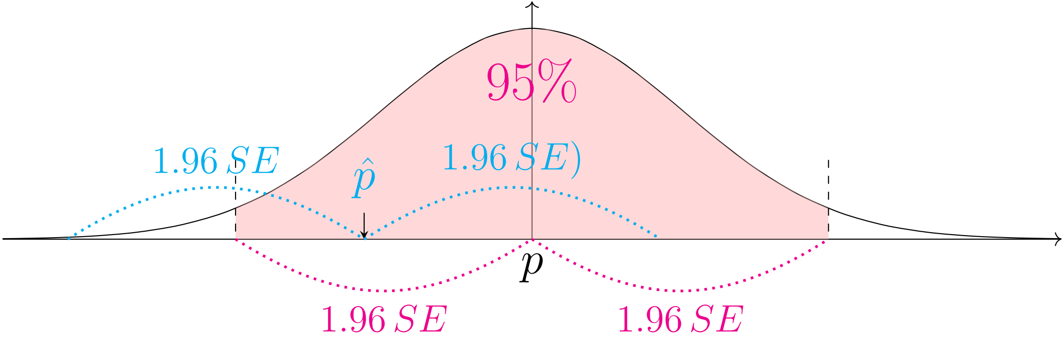
標本比率の標準誤差：$SE = \sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}}$

math-support.jp



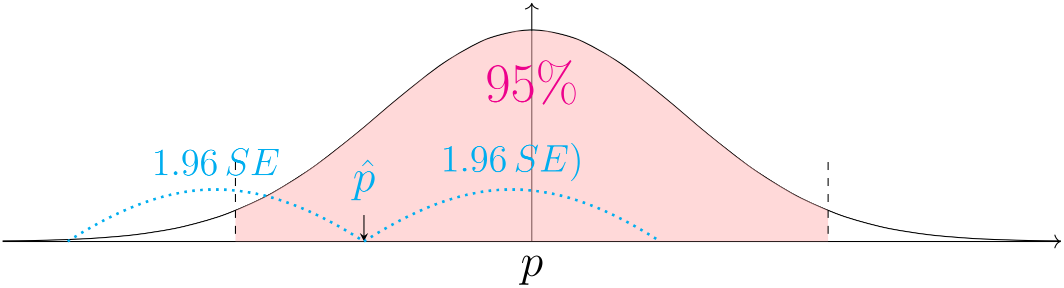
標本比率の標準誤差：$SE = \sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}}$

math-support.jp



標本比率の標準誤差：$SE = \sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}}$

math-support.jp



標本比率の標準誤差：$SE = \sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}}$

母比率 p の推定

標本の大きさ n、標本比率 $\hat{p}$ とするとき、

標準誤差を $SE = \sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}}$ とすると、

信頼度 95% の信頼区間

$$\hat{p} - 1.96 \cdot SE \leqq p \leqq \hat{p} + 1.96 \cdot SE$$

母比率 p の推定

標本の大きさ n、標本比率 $\hat{p}$ とするとき、

標準誤差を $SE=\sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}}$ とすると、

信頼度 95% の信頼区間

$$\hat{p}-1.96\cdot SE \leqq p \leqq \hat{p}+1.96\cdot SE$$

信頼度 99% の信頼区間

$$\hat{p}-2.58\cdot SE \leqq p \leqq \hat{p}+2.58\cdot SE$$

math-support.jp

例 1　ある県で、400 世帯を無作為抽出し、A という意見の賛否を調べたところ、250 世帯が賛成であった。全世帯における賛成の母比率 p を信頼度 95%で推定せよ。

math-support.jp

例 1　ある県で、400 世帯を無作為抽出し、A という意見の賛否を調べたところ、250 世帯が賛成であった。全世帯における賛成の母比率 p を信頼度 95%で推定せよ。

$$n = 400, \quad \hat{p} = \frac{250}{400} = 0.625 \text{ であるので、}$$

math-support.jp

例 1　ある県で、400 世帯を無作為抽出し、A という意見の賛否を調べたところ、250 世帯が賛成であった。全世帯における賛成の母比率 p を信頼度 95%で推定せよ。

$n = 400, \quad \hat{p} = \dfrac{250}{400} = 0.625$ であるので、

$$SE = \sqrt{\frac{0.625(1-0.625)}{400}} = 0.0242$$

math-support.jp

例 1　ある県で、400 世帯を無作為抽出し、A という意見の賛否を調べたところ、250 世帯が賛成であった。全世帯における賛成の母比率 p を信頼度 95%で推定せよ。

$n = 400, \quad \hat{p} = \dfrac{250}{400} = 0.625$ であるので、

$$SE = \sqrt{\frac{0.625(1-0.625)}{400}} = 0.0242$$

95%の信頼区間は

math-support.jp

例 1　ある県で、400 世帯を無作為抽出し、A という意見の賛否を調べたところ、250 世帯が賛成であった。全世帯における賛成の母比率 p を信頼度 95%で推定せよ。

$n = 400, \quad \hat{p} = \dfrac{250}{400} = 0.625$ であるので、

$$SE = \sqrt{\frac{0.625(1-0.625)}{400}} = 0.0242$$

95%の信頼区間は

$$0.625 - 1.96 \cdot 0.0242 \leq p \leq 0.625 + 1.96 \cdot 0.0242$$

math-support.jp

例 1　ある県で、400 世帯を無作為抽出し、A という意見の賛否を調べたところ、250 世帯が賛成であった。全世帯における賛成の母比率 p を信頼度 95%で推定せよ。

$n = 400,\quad \hat{p} = \dfrac{250}{400} = 0.625$ であるので、

$$SE = \sqrt{\frac{0.625(1-0.625)}{400}} = 0.0242$$

95%の信頼区間は

$$0.625 - 1.96 \cdot 0.0242 \leqq p \leqq 0.625 + 1.96 \cdot 0.0242$$

$$0.578 \leqq p \leqq 0.672$$

math-support.jp

例 1　ある県で、400 世帯を無作為抽出し、A という意見の賛否を調べたところ、250 世帯が賛成であった。全世帯における賛成の母比率 p を信頼度 95%で推定せよ。

$$n = 400, \quad \hat{p} = \frac{250}{400} = 0.625 \text{ であるので、}$$

$$SE = \sqrt{\frac{0.625(1-0.625)}{400}} = 0.0242$$

95%の信頼区間は

$$0.625 - 1.96 \cdot 0.0242 \leqq p \leqq 0.625 + 1.96 \cdot 0.0242$$

$$0.578 \leqq p \leqq 0.672$$

答　57.8%から 67.2%の間

math-support.jp

ビデオを止めて問題を解いてみよう

問 1　ある県で、800 世帯を無作為抽出して、A という意見の賛否を調べたところ、500 世帯が賛成であった。全世帯における賛成の母比率 p を信頼度 95%で推定せよ。

math-support.jp

問 1　ある県で、800 世帯を無作為抽出し、A という意見の賛否を調べたところ、500 世帯が賛成であった。全世帯における賛成の母比率 p を信頼度 95%で推定せよ。

問 1　ある県で、800 世帯を無作為抽出し、A という意見の賛否を調べたところ、500 世帯が賛成であった。全世帯における賛成の母比率 p を信頼度 95%で推定せよ。

$$n = 800, \quad \hat{p} = \frac{500}{800} = 0.625 \text{ であるので、}$$

math-support.jp

問 1　ある県で、800 世帯を無作為抽出し、A という意見の賛否を調べたところ、500 世帯が賛成であった。全世帯における賛成の母比率 p を信頼度 95%で推定せよ。

$n = 800, \quad \hat{p} = \dfrac{500}{800} = 0.625$ であるので、

$$SE = \sqrt{\frac{0.625(1-0.625)}{800}} = 0.0171$$

math-support.jp

問 1 ある県で、800 世帯を無作為抽出し、A という意見の賛否を調べたところ、500 世帯が賛成であった。全世帯における賛成の母比率 p を信頼度 95%で推定せよ。

$n = 800, \quad \hat{p} = \dfrac{500}{800} = 0.625$ であるので、

$$SE = \sqrt{\frac{0.625(1-0.625)}{800}} = 0.0171$$

95%の信頼区間は

math-support.jp

問 1　ある県で、800 世帯を無作為抽出し、A という意見の賛否を調べたところ、500 世帯が賛成であった。全世帯における賛成の母比率 p を信頼度 95%で推定せよ。

$n = 800, \quad \hat{p} = \dfrac{500}{800} = 0.625$ であるので、

$$SE = \sqrt{\frac{0.625(1-0.625)}{800}} = 0.0171$$

95%の信頼区間は

$$0.625 - 1.96 \cdot 0.0171 \leqq p \leqq 0.625 + 1.96 \cdot 0.0171$$

math-support.jp

問 1　ある県で、800 世帯を無作為抽出し、A という意見の賛否を調べたところ、500 世帯が賛成であった。全世帯における賛成の母比率 p を信頼度 95%で推定せよ。

$n = 800, \quad \hat{p} = \dfrac{500}{800} = 0.625$ であるので、

$$SE = \sqrt{\frac{0.625(1-0.625)}{800}} = 0.0171$$

95%の信頼区間は

$$0.625 - 1.96 \cdot 0.0171 \leqq p \leqq 0.625 + 1.96 \cdot 0.0171$$

$$0.591 \leqq p \leqq 0.659$$

math-support.jp

問 1　ある県で、800 世帯を無作為抽出し、A という意見の賛否を調べたところ、500 世帯が賛成であった。全世帯における賛成の母比率 p を信頼度 95%で推定せよ。

$n = 800, \quad \hat{p} = \dfrac{500}{800} = 0.625$ であるので、

$$SE = \sqrt{\frac{0.625(1-0.625)}{800}} = 0.0171$$

95%の信頼区間は

$$0.625 - 1.96 \cdot 0.0171 \leqq p \leqq 0.625 + 1.96 \cdot 0.0171$$

$$0.591 \leqq p \leqq 0.659$$

答　59.1%から 65.9%の間

math-support.jp

[参考]

標本サイズが $400 \to 800$ と 2 倍になり、信頼区間は：

400 世帯：$67.2 - 57.8 = 9.4\%$

800 世帯：$65.9 - 59.1 = 6.8\%$

math-support.jp

[参考]

標本サイズが $400 \to 800$ と 2 倍になり、信頼区間は：

400 世帯：$67.2 - 57.8 = 9.4\%$

800 世帯：$65.9 - 59.1 = 6.8\%$

標本サイズが 2 倍になると信頼区間は $\dfrac{1}{\sqrt{2}} = 0.71$ 倍

math-support.jp

[参考]

標本サイズが $400 \to 800$ と 2 倍になり、信頼区間は：

$$400\text{ 世帯}：67.2 - 57.8 = 9.4\%$$

$$800\text{ 世帯}：65.9 - 59.1 = 6.8\%$$

標本サイズが 2 倍になると信頼区間は $\dfrac{1}{\sqrt{2}} = 0.71$ 倍

信頼区間を小さくして、母平均の値を絞り込みたいときには、標本の数を増やすのが唯一の方法である。

今回の学習目標

標本の比率から母比率を推定する。