

1 学習内容の説明 ⇒ 2 問題演習 ⇒ 3 振り返り(確認テスト・相互採点・リフレクションの記入)

【態度目標】しゃべる、質問する、説明する、動く、協力する、貢献する

【内容目標】標本平均から母集団を推測する方法を整理していこう

母平均と標本平均、母標準偏差と標本の標準偏差の違いを理解しよう

□標本平均の期待値と標準偏差

母集団から大きさ n の無作為標本を抽出し、それらの変数 x の値を X_1, X_2, \dots, X_n とするとき、

$$\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{n}$$

を **標本平均** という。 n を固定すると、標本平均 \bar{X} は1つの確率変数になる。



標本平均

例21) 標本平均の確率分布と期待値

1, 2の数字が書いてあるカードが50枚ずつある。

この合計100枚のカードからなる母集団から、復元抽出によって大きさ3の無作為標本を抽出し、そのカードの数字を順に X_1, X_2, X_3 とする。

このとき、 $X_1 + X_2 + X_3$ のとりうる値は3, 4, 5, 6であるから、

標本平均 $\bar{X} = \frac{X_1 + X_2 + X_3}{3}$ のとりうる値は $1, \frac{4}{3}, \frac{5}{3}, 2$ である。

\bar{X} がそれぞれの値をとる確率を求めて、

\bar{X} の確率分布は右の表のようになる。

また、 \bar{X} の期待値 $E(\bar{X})$ は

$$E(\bar{X}) = 1 \cdot \frac{1}{8} + \frac{4}{3} \cdot \frac{3}{8} + \frac{5}{3} \cdot \frac{3}{8} + 2 \cdot \frac{1}{8} = \frac{3}{2}$$

\bar{X}	1	$\frac{4}{3}$	$\frac{5}{3}$	2	計
P	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$	1

<補足> 例21の100枚のカードからなる母集団の母平均は $\frac{3}{2}$ である。

例 5人の身長が
170, 174, 166, 168, 177 (cm)

母集団

170
174
166
168
177

母集団数 $N=5$
母平均 $m=171.0$
母分散 $\sigma^2=16.0$

標本数 $n=2$

標本

(174, 166) → 170.0
(174, 168) → 171.0
(174, 170) → 172.0
(174, 174) → 174.0
(174, 177) → 175.5
(166, 166) → 166.0
...

標本の平均値

→ 171.0
→ 171.0
→ 172.0
→ 174.0
→ 175.5
→ 166.0
...

標本の平均値の平均

→ 171.0 ← 一致する!

標本の平均値の分散

→ 8.0

平均はいろいろな値をとるが、それらの値は母平均 m を中心にして周りを動いている
⇒ 母平均 m が不明でも標本平均を m の推測値とする

n が小さいとたまたま偏った標本を選ぶ確率が多い(影響が大きい)。 n が大きくなると母集団との差が小さくなる。

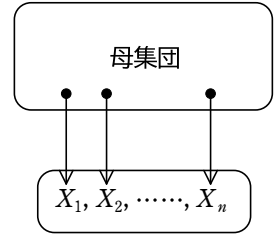
母分散の $\frac{1}{n}$ 倍 (無限母集団: 復元抽出)
母分散の $\frac{N-n}{N-1} \cdot \frac{1}{n}$ 倍 (有限母集団: 非復元抽出)

標本平均 \bar{X} の確率分布と母集団分布の関係を調べてみよう。

母平均 m ，母標準偏差 σ の母集団から，復元抽出によって大きさ n の無作為標本を抽出し，それらの変量 x の値を X_1, X_2, \dots, X_n とする。

各 X_k は，どれも大きさ 1 の標本で，母集団分布に従う確率変数である。

互いに独立である



よって $E(X_1) = E(X_2) = \dots = E(X_n) = m$
 $\sigma(X_1) = \sigma(X_2) = \dots = \sigma(X_n) = \sigma$

したがって

$$E(\bar{X}) = E\left(\frac{X_1 + X_2 + \dots + X_n}{n}\right)$$

$$= \frac{1}{n} \{E(X_1) + E(X_2) + \dots + E(X_n)\}$$

$$= \frac{1}{n} \cdot nm = m$$

また，復元抽出の場合， X_1, X_2, \dots, X_n は互いに独立な確率変数であるから

$$V(\bar{X}) = V\left(\frac{X_1 + X_2 + \dots + X_n}{n}\right)$$

$$= \frac{1}{n^2} \{V(X_1) + V(X_2) + \dots + V(X_n)\}$$

$$= \frac{1}{n^2} \cdot n\sigma^2 = \frac{\sigma^2}{n}$$

【おさらい】
 $E(aX + bY) = aE(X) + bE(Y)$
 特に X_1, X_2, \dots, X_n に対して
 $E(X_1 + \dots + X_n) = E(X_1) + \dots + E(X_n)$
 $V(aX) = a^2V(X)$
 復元抽出のときは X_1, X_2, \dots, X_n は互いに独立であり
 $V(X_1 + X_2 + \dots + X_n)$
 $= V(X_1) + V(X_2) + \dots + V(X_n)$
 $= n\sigma^2$

$V(X_1) = V(X_2) = \dots = V(X_n) = \sigma^2$

よって $\sigma(\bar{X}) = \sqrt{\frac{\sigma^2}{n}} = \frac{\sigma}{\sqrt{n}}$ $\sigma > 0$ に注意

これまでのことをまとめると，次のことがいえる。

標本平均の期待値と標準偏差
 母平均 m ，母標準偏差 σ の母集団から大きさ n の無作為標本を抽出するとき，その標本平均 \bar{X} の期待値 $E(\bar{X})$ と標準偏差 $\sigma(\bar{X})$ は

$$E(\bar{X}) = m, \quad \sigma(\bar{X}) = \frac{\sigma}{\sqrt{n}}$$

非復元抽出の場合も，標本の大きさ n に比べて母集団の大きさが十分大きいときは，復元抽出と同様に扱ってよいことが知られている。

例 2 2) 母平均 60，母標準偏差 5 の十分大きい母集団から，大きさ 25 の標本を抽出するとき，その標本平均 \bar{X} について

期待値は $E(\bar{X}) = 60$

標準偏差は $\sigma(\bar{X}) = \frac{5}{\sqrt{25}} = \frac{5}{5} = 1$

終