

【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、

$$\text{これらの平均 } \bar{X} = \frac{X_1 + X_2 + \dots + X_n}{n}$$

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。

また、

$$S = \sqrt{\frac{1}{n} \sum_{k=1}^n (X_k - \bar{X})^2}$$

標本平均との差の 2 乗の平均に
ルートを付けたもの



標本平均

を **標本標準偏差** という。

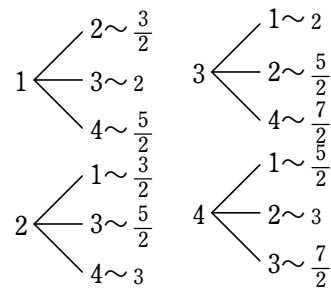
X_1, X_2, \dots, X_n は、標本を抽出するという試行の結果により値の定まる確率変数である。したがって、 \bar{X}, S も、標本を抽出するという試行の結果により値の定まる確率変数である。

問 7) 4 枚のカードに、それぞれ 1, 2, 3, 4 の数字が書いてある。この 4 枚のカードからなる母集団から、次のような方法で大きさ 2 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 とする。標本平均 $\bar{X} = \frac{1}{2}(X_1 + X_2)$ の確率分布を、次のそれぞれの場合に求めよ。

(1) 復元抽出

\bar{X}	1	2	3	4
1	1	$\frac{3}{2}$	2	$\frac{5}{2}$
2	$\frac{3}{2}$	2	$\frac{5}{2}$	3
3	2	$\frac{5}{2}$	3	$\frac{7}{2}$
4	$\frac{5}{2}$	3	$\frac{7}{2}$	4

(2) 非復元抽出

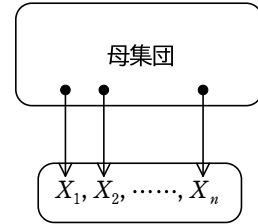


\bar{X}	1	$\frac{3}{2}$	2	$\frac{5}{2}$	3	$\frac{7}{2}$	4	計
P	$\frac{1}{16}$	$\frac{2}{16}$	$\frac{3}{16}$	$\frac{4}{16}$	$\frac{3}{16}$	$\frac{2}{16}$	$\frac{1}{16}$	1

\bar{X}	$\frac{3}{2}$	2	$\frac{5}{2}$	3	$\frac{7}{2}$	計
P	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{2}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	1

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 の母集団から大きさ n の無作為標本を抽出し, その標本のもつ変数 x の値として定まる確率変数を X_1, X_2, \dots, X_n とする。この抽出が復元抽出の場合, これらの各変数は, 大きさ 1 の標本の確率変数とみなされ, それぞれ母集団分布に従うから



$$E(X_1) = E(X_2) = \dots = E(X_n) = m$$

$$\sigma(X_1) = \sigma(X_2) = \dots = \sigma(X_n) = \sigma$$

したがって, \bar{X} の期待値と標準偏差は,

$$\begin{aligned} E(\bar{X}) &= E\left(\frac{X_1 + X_2 + \dots + X_n}{n}\right) \\ &= \frac{1}{n}\{E(X_1) + E(X_2) + \dots + E(X_n)\} \\ &= \frac{1}{n} \cdot nm = m \end{aligned}$$

X_1, X_2, \dots, X_n は互いに独立な確率変数であるから

$$\begin{aligned} \sigma(\bar{X}) &= \sqrt{V\left(\frac{X_1 + X_2 + \dots + X_n}{n}\right)} \\ &= \sqrt{V\left(\frac{X_1}{n}\right) + V\left(\frac{X_2}{n}\right) + \dots + V\left(\frac{X_n}{n}\right)} \\ &= \sqrt{\frac{1}{n^2}\{V(X_1) + V(X_2) + \dots + V(X_n)\}} \\ &= \sqrt{\frac{1}{n^2} \cdot n\sigma^2} = \sqrt{\frac{\sigma^2}{n}} = \frac{\sigma}{\sqrt{n}} \end{aligned}$$

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

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

$\sigma > 0$ に注意

非復元抽出による無作為標本の場合も, 母集団の大きさが標本の大きさ n に比べて十分大きいときは, 復元抽出の場合と同様に考えてよい。

したがって, 次のことが成り立つ。

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。

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, 標本平均 \bar{X} の期待値と標準偏差は

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

例 1 3) 母平均 60, 母標準偏差 5 の母集団から, 大きさ 25 の標本を無作為抽出するとき, その標本平均 \bar{X} の期待値と標準偏差は

$$E(\bar{X}) = 60, \quad \sigma(\bar{X}) = \frac{5}{\sqrt{25}} = \frac{5}{5} = 1 \quad \text{終}$$

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
 \Rightarrow 母平均 m が不明でも標本平均を m の推測値とする

例 5人の身長が
 170, 174, 166, 168, 177 (cm)

母集団	標本	標本の平均値	標本の平均値の平均
170	(174, 166)	$\rightarrow 170.0$	$\rightarrow 171.0$ ← 一致する!
174	(174, 168)	$\rightarrow 171.0$	
174	(174, 170)	$\rightarrow 172.0$	母平均の分散 $\rightarrow 8.0$
166	(174, 174)	$\rightarrow 174.0$	
168	(174, 177)	$\rightarrow 175.5$	母分散の $\frac{1}{n}$ 倍 (無限母集団: 復元抽出) 標準偏差はルートを付けて $\frac{1}{\sqrt{n}}$ 倍に
177	(166, 166)	$\rightarrow 166.0$	
	

母集団数 $N=5$
 母平均 $m=171.0$
 母分散 $\sigma^2=16.0$
 母標準偏差 $\sigma=4.0$

標本数 $n=2$

n が小さいとたまたま偏った標本を選ぶ確率が高い (影響が大きい)。

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

例題 8) 箱の中に製品が多数入っていて、その中の不良品の割合は $\frac{1}{6}$ である。この箱の中から、標本として無作為に 30 個の製品を抽出するとき、 k 番目に抽出された製品が不良品なら 1、良品なら 0 の値を対応させる確率変数を X_k とする。

標本平均 $\bar{X} = \frac{1}{30}(X_1 + X_2 + \dots + X_{30})$ の期待値と標準偏差を求めよ。

解答 母集団は製品全体であるが、母集団における変量は、不良品のとき 1、良品のとき 0 という 2 つの値をとる。その場合の母平均 m と母標準偏差 σ は

X_k	0	1	計
P	$\frac{5}{6}$	$\frac{1}{6}$	1

$$m = 1 \cdot \frac{1}{6} + 0 \cdot \frac{5}{6} = \frac{1}{6}$$

$$\sigma = \sqrt{\left(1^2 \cdot \frac{1}{6} + 0^2 \cdot \frac{5}{6}\right) - m^2} = \sqrt{\frac{1}{6} - \left(\frac{1}{6}\right)^2} = \sqrt{\frac{5}{36}} = \frac{\sqrt{5}}{6}$$

したがって、 \bar{X} の期待値と標準偏差は

$$E(\bar{X}) = m = \frac{1}{6} \quad \sigma(\bar{X}) = \frac{\sigma}{\sqrt{30}} = \frac{\sqrt{6}}{36}$$

母集団	標本 \rightarrow 標本の平均値 \bar{X}	標本の平均値の平均
母集団数 $N=?$	(X_1, X_2, \dots, X_{30})	$\rightarrow E(X) = m = \frac{1}{6}$ ← 一致!
母平均 $m = \frac{1}{6}$	(X_1, X_2, \dots, X_{30})	標本の平均値の分散・標準偏差 $V(\bar{X}) = \frac{\sigma^2}{30} = \frac{1}{30} \cdot \frac{5}{36} = \frac{1}{6^3}$
母分散 $\sigma^2 = \frac{5}{36}$	(X_1, X_2, \dots, X_{30})	
母標準偏差 $\sigma = \frac{\sqrt{5}}{6}$	(X_1, X_2, \dots, X_{30})	$\sigma(\bar{X}) = \frac{\sigma}{\sqrt{30}} = \frac{\sqrt{5}}{6\sqrt{30}} = \frac{\sqrt{6}}{36}$