

P.93~

基本事項

- 一般に、母集団の中である特性 A をもつものの割合を、その特性 A の_____という。
- 抽出された標本の中で特性 A をもつものの割合を_____という。

全体課題 (練習 30)

不良品が全体の 10% 含まれる大量の製品の山から大きさ 100 の無作為標本を抽出するとき、不良品の標本比率 R について、次の問いに答えよ。

- (1) R は近似的にどのような正規分布に従うとみなすことができるか。
- (2) $0.07 \leq R \leq 0.13$ となる確率を求めよ。

解答

♠ エキスパート A 「標本比率は標本平均」

目標 A

標本比率が従う分布について、二項分布から正規分布に至る議論の流れをもとに下の性質が成り立つことを理解し説明できる。

性質

特性 A の母比率 p の母集団から抽出された大きさ n の無作為標本について、 n が大きいとき、標本比率 R は近似的に正規分布 $N\left(p, \frac{pq}{n}\right)$ に従うとみなすことができる。 (つまり $R \sim N\left(p, \frac{pq}{n}\right)$)

説明

- 大きさ n の無作為標本を抽出し、それらの値 X_1, X_2, \dots, X_n を

$$\text{特性を持つなら } X_k = 1$$

$$\text{特性を持たないなら } X_k = 0 \ (k = 1, 2, \dots, n)$$

と定める (← 特性をもつものの個数を数えるダミー変数。)

- $T = X_1 + X_2 + \dots + X_n$ とおけば、 T は「特性 A をもつ」【成功回数 / 成功確率】を表しているので、二項分布 $B(n, p)$ に従う。つまり、

$$T \sim B(n, p).$$

- n が大きいとき、正規近似が可能だから、 T は正規分布 $N(np, npq)$ に従う。つまり、

$$T \sim N\left(\underbrace{np}_{=E(T)}, \underbrace{npq}_{=V(T)}\right).$$

- ところで、標本比率 R とは、相対度数 $\frac{T}{n}$ のことだった。

- T が正規分布に従うとき、その定数倍 $\frac{T}{n}$ も正規分布に従う (P.78 をみよ。)

- 今、

$$E\left(\frac{T}{n}\right) = \frac{1}{n}E(T) = \frac{np}{n} = p,$$

$$V\left(\frac{T}{n}\right) = \underbrace{\frac{1}{n^2}}_{\text{分散は2乗量}} V(T) = \frac{npq}{n^2} = \frac{pq}{n}.$$

だから、結局、

- $R = \frac{T}{n}$ は、正規分布 $N\left(p, \frac{pq}{n}\right)$ に従う。つまり、

$$R \sim N\left(p, \frac{pq}{n}\right)$$

♠ エキスパート B 「標本比率の従う分布」

目標 B

下の性質を用いて、標本比率 R が従う正規分布を間違うことなく書き出すことができる。(性質の証明はエキスパート A がしてくれる)

性質

特性 A の母比率 p の母集団から抽出された大きさ n の無作為標本について、 n が大きいとき、標本比率 R は近似的に正規分布 $N\left(p, \frac{pq}{n}\right)$ に従うとみなすことができる。 (つまり $R \sim N\left(p, \frac{pq}{n}\right)$)

特訓

- あるアンケートにおいて、特定の意見に賛成する人の母比率 p が 50% である。この母集団から大きさ $n = 400$ の標本を抽出したとき、標本比率 R は、正規分布 $N\left(\quad, \quad\right)$ に従うとみなせる。
- ある地域で実施された大規模な予防接種において、副作用が出る人の母比率 p が 20% である。この母集団から大きさ $n = 400$ の標本を抽出したとき、標本比率 R は、正規分布 $N\left(\quad, \quad\right)$ に従うとみなせる。
- 工場で製造される製品のうち、高品質と評価されるものの母比率 p が 40% である。この母集団から大きさ $n = 600$ の標本を抽出したとき、標本比率 R は、正規分布 $N\left(\quad, \quad\right)$ に従うとみなせる。
- あるスポーツチームのファンにおける熱心なサポーターの母比率 p が 75% である。この母集団から大きさ $n = 300$ の標本を抽出したとき、標本比率 R は、正規分布 $N\left(\quad, \quad\right)$ に従うとみなせる。
- ある希少疾患を持つ人の母比率 p が 1% である。この母集団から大きさ $n = 9900$ の標本を抽出したとき、標本比率 R は、正規分布 $N\left(\quad, \quad\right)$ に従うとみなせる。

余談

標本比率 R はダミー変数を用いることで、 $R = \frac{X_1 + X_2 + \dots + X_n}{n}$ とかける。(これはエキスパート A が説明してくれる。)

「これって標本平均と似てない??」というわけで、下の表を読んで標本平均と標本比率の違いと類似性を見つけよう。

	定義	母集団分布	従う分布
標本平均 \bar{X}	$\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{n}$	$N(m, \sigma^2)$	$N\left(m, \frac{\sigma^2}{n}\right)$
標本比率 R	$R = \frac{X_1 + X_2 + \dots + X_n}{n}$	$N(np, npq)$	$N\left(p, \frac{pq}{n}\right)$