< 非復元抽出による標本調査 1 >

例 3　2 ページ例 2 で、n 人の標本を非復元抽出した場合を考える。すなわち N 人の身長データ $\{y_1, y_2, \cdots, y_N\}$ の中から順に n 個のデータ $\{X_1, X_2, \cdots, X_n\}$ を (無作為に) 非復元抽出する。このとき N 個のデータ $\{y_1, y_2, \cdots, y_N\}$ の中の任意の n 個のデータ $\{y_{i1}, y_{i2}, \cdots, y_{in}\}$ を選ぶ確率は

$$P(X_1 = y_{i1}, X_2 = y_{i2}, \cdots, X_n = y_{in}) = \frac{1}{N(N-1)(N-2)\cdots(N-n+1)}$$

である。このとき、次の定理が成り立つ。

定理 15

(1) $P(X_k = y_i) = \dfrac{1}{N}$　　$(k = 1, 2, \cdots, n \ ,\ \ y_i \in \{y_1, y_2, \cdots, y_N\})$

(2) $P(X_k = y_i\, , X_\ell = y_i) = \dfrac{1}{N(N-1)}$　　$(k \neq \ell \ ,\ \ i \neq j)$

(3) $E[X_k] = \mu \quad , \quad E\left[(X_k - \mu)^2\right] = \sigma^2$　　$(k = 1, 2, \cdots, n)$

(4) $E\left[\,\overline{X}\,\right] = \mu$

(5) $E\left[\left(\overline{X} - \mu\right)^2\right] = \dfrac{N-n}{N-1} \cdot \dfrac{\sigma^2}{n}$

(6) $E\left[\dfrac{1}{n-1}\displaystyle\sum_{k=1}^{n}\left(X_k - \overline{X}\right)^2\right] = \dfrac{N}{N-1}\sigma^2$

ただし $\mu = \dfrac{1}{N}\displaystyle\sum_{i=1}^{N} y_i \quad , \quad \sigma^2 = \dfrac{1}{N}\displaystyle\sum_{i=1}^{N}(y_i - \mu)^2 \quad , \quad \overline{X} = \dfrac{1}{n}\displaystyle\sum_{k=1}^{n} X_k$ である。

系　$V = \dfrac{N-1}{N(n-1)}\displaystyle\sum_{k=1}^{n}\left(X_k - \overline{X}\right)^2$ とおくと $E[V] = \sigma^2$ が成り立つ。

この $V = \dfrac{N-1}{N(n-1)}\displaystyle\sum_{k=1}^{n}\left(X_k - \overline{X}\right)^2$ を非復元抽出の場合の不偏分散という。

定理 16

$$E\left[\left(V - \sigma^2\right)^2\right]$$

$$= \left(\frac{N-1}{N}\right)^2 \frac{\mu_4}{n} - \frac{n-3}{n(n-1)}\sigma^4 - \frac{N-1}{N^2 n(n-1)}\left\{n^2 - 6n + 7 + \frac{4(n-2)(n-3)}{N-2} + \frac{6(n-2)(n-3)}{(N-2)(N-3)}\right\}\mu_4$$

$$+ \ \frac{1}{Nn(n-1)}\left\{n^2 - 8n + 9 + \frac{(n-2)(n-3)(5N-9)}{(N-2)(N-3)}\right\}\sigma^4$$

ただし $\mu_4 = \dfrac{1}{N}\displaystyle\sum_{i=1}^{N}(y_i - \mu)^4$ である。