

超幾何分布

[英] *hypergeometric distribution*

有限個のものの集まりの中から n 個のものを選び出す場合の様に n 回の試行が独立ではないときに用いる。

全体で N 個のものがある集まりの中から n 個を選び出すとする。 N 個のものの中で、ある性質（性質 A とする）を持っているものの割合を p とすると、全体で Np 個のものが性質 A を持っていることになる。選び出される n 個の中で性質 A をもつものの数を x とすれば、問題は x の確率分布を求めることである。

N 個のものから n 個を選ぶ選び方は全体で ${}_N C_n$ 通りである。 n 個のうち x 個は性質 A をもつわけであるから、それらは Np 個の中から選ばれなければならない、その選ばれ方は ${}_{Np} C_x$ 通りである。残りの $(n-x)$ 個は性質 A を持たない $(N-Np)$ 個の中から選ばれなければならない、その選ばれ方は ${}_{N-Np} C_{n-x}$ 通りである。従って、 n 個のうち x 個が Np 個の中から、そして $(n-x)$ 個が $(N-Np)$ 個の中から選ばれる方法は ${}_{Np} C_x \times {}_{N-Np} C_{n-x}$ 通りである。故に求める確率分布 $P(x)$ は

$$P(x) = \frac{{}_{Np} C_x \cdot {}_{N-Np} C_{n-x}}{{}_N C_n}$$

であり、これを超幾何分布という。超幾何分布は N が大きくなると二項分布に近づく。