

円周率の統計学

円周率を 50,000 桁計算して、0, 1, 2, ..., 9 の出現頻度を求めると、

数字	頻度
0	5033
1	5054
2	4867
3	4948
4	5011
5	5052
6	5018
7	4977
8	5030
9	5010
合計	50000

これを見ると、多少のばらつきはあるが、0 から 9 までの数字は、同じ数だけ現れているように見える。そこで、0 から 9 までの数字は、全て等しい割合で現れると仮定してみよう。すなわち、

$$\text{帰無仮説 } H_0: p_0 = p_1 = p_2 = \cdots = p_9 = \frac{1}{10}$$

をたてて、適合度の検定を行う。総数を $N = 50000$ とすると、

$$\chi^2 = \sum_{i=0}^9 \frac{(N_i - p_i N)^2}{p_i N}$$

は、自由度 9 の χ^2 分布に従う。上の表の値を代入すると、

$$\begin{aligned}\chi^2 &= \frac{1}{5000}(33^2 + 54^2 + 133^2 + 52^2 + 11^2 + 52^2 + 18^2 + 23^2 + 30^2 + 10^2) \\ &= \frac{29076}{5000} \\ &= 5.8152\end{aligned}$$

自由度が 9 の場合、危険率 5% のときの棄却域は、 $\chi^2 > 16.92$ である。 $\chi^2 = 5.8152$ は、採択域にある。従って、仮説 H_0 は、危険率 5% で棄却できない。