

統計学の基礎

生物を研究する上で、統計の知識は非常に重要です。

ここでは統計学の基本中の基本を紹介します。

なぜ生物を研究するのに統計学が必要なのでしょう。それは、刺激に対して生物が常に同じ行動をとるとは限らないからです。

たとえば、ある生物の前に餌Aと餌Bが置いてあるとします。どちらか一方を選ばせたとき、その生物が餌Aを選んだとしたら、「この生物は餌Aが好きだ」と言えるでしょうか？この場合断言はできません。本当は餌Bが好きなのに、その時だけたまたま餌Aを選んだのかもしれないからです。ではどのようにすれば「この生物は餌Aが好きだ」と判定できるのでしょうか？方法は次の通りです。

この実験を10回繰り返したとします。確かめたい仮説(対立仮説といいます)は「この生物は餌Aが好きである」ですが、ここでは「この生物に餌の好みはない」と仮定します(帰無仮説といいます)。つまり餌Aを選ぶ確率と餌Bを選ぶ確率はそれぞれ1/2であると仮定するわけです。

ここで、もし餌Aを9回選んだとしたら、この現象が偶然起こる確率はどうなるのでしょうか？高校数学を学んだ人であれば、

$${}_{10}C_9 \times (1/2)^9 \times (1/2)^1 = 10/1024 = 0.00976562$$

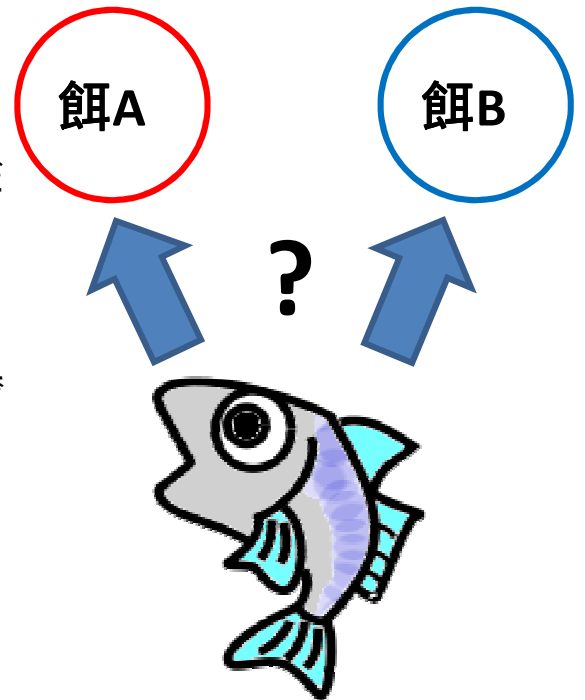
という式から確率は約0.01であるとわかります。

さらに、統計では「実際に起こった結果よりも起こりにくい結果」の確率も合計するので、餌Aを10回選んだ確率も計算します。上の式と同様に、

$${}_{10}C_{10} \times (1/2)^{10} \times (1/2)^0 = 0.00097656$$

という式から確率は約0.001であるとわかります。

これらを合計すると0.011ですから(この確率を有意確率とよび、 $p=0.011$ などと表します)、100回に約1回しか起きないというかなり低い確率になります。つまり、「この生物に餌の好みがない」と仮定



すると、非常に低い確率でしか発生しない現象が起きたわけです。

これほど低確率の現象が偶然起こったとは考えにくいので(一般に0.05より小さい確率だと、偶然起こったとは考えにくいとされます。この場合、 $p < 0.05$ などと書かれます)、「この生物に餌の好みはない」という仮説は棄却されます。したがって「この生物は餌Aが好きだ」ということができるわけです。(このように有意確率が十分に小さいときは「有意な差がある」といいます)。

では餌Aを7回選んだとしたら確率はどうなるでしょうか？先ほどと同じように計算すると、

$${}_{10}C_{10} \times (1/2)^{10} \times (1/2)^0 + {}_{10}C_9 \times (1/2)^9 \times (1/2)^1 + {}_{10}C_8 \times (1/2)^8 \times (1/2)^2 + {}_{10}C_7 \times (1/2)^7 \times (1/2)^3 = 0.1714$$

という式から確率は約0.17であるとわかります。

この場合、偶然起こることもありえるので、「この生物に餌の好みはない」という仮説を棄却することはできません。よって、「この生物は餌Aが好きだ」とはいえないわけです。

このように、生物における統計学では、ある仮説を立てて実験した場合、その仮説が間違っていると仮定して、その実験結果が偶然起こる確率を計算しています。さまざまな検定法が考案されていますが、すべての検定法は「その実験結果が偶然発生する確率」を計算しているのです。

ちなみに、今回用いた方法は「二項検定」と呼ばれています。今回のように二つのうちどちらを選ぶかという実験のほか、性比など「カテゴリAの数とカテゴリBの数」に差があるかどうかを検定するときにも用います。

検定の方法は他にもたくさんあり、実験の方式によって選ぶ必要があります。生物学全般、特に生態学や医学に興味のある人は、様々な統計法をがんばって勉強してください。

おすすめ図書:「生物統計学入門」 石居 進 (培風館)

「生物学を学ぶ人のための統計のはなし」 粕谷 英一 (文一総合出版)

仮説を立て、実験する



その仮説が間違っていた場合を想定して、実験結果が偶然起こる確率を計算する



確率が十分に小さく、偶然起こったとは考えにくいならば、最初の仮説は正しい