

1 章

統計学の世界へようこそ



1.1 はじめに

生物学を履修している学生は統計学を苦手になっている傾向がある。統計学は数学の一分野であり、通常、生物学者が最初から得意とするものではない。しかし、多くの生物学者は、生物学を勉強するにつれて、統計学の背後にある数学に対する恐怖心を捨て去るべきであると感じてくる。生物統計学に関するどんな本でもまず答えなければならない問いかけは、**なぜ、生物学者は統計学に悩まなければならないのか？**である。また、多くの学生は統計学の背後にある考え方は理解しづらいつ感じている。それは、統計学者が、通常の科学的論理とは違った、ある意味で奇妙な逆転した考え方をするからである。それゆえ、この本でも、**なぜ、統計的論理はそれほど奇妙か？**という問いかけにも答えなければならない。

最後に、学生は、しばしば統計学の本のサイズとその情報量の多さに不満を言うが、それはもっともなことである。これは、非常にたくさんの統計的検定があるからであるが、本書でも、**なぜ、それだけの異なる統計的検定があるか？**という問いかけに答える必要がある。

本章を読めば、読者がこれらの問いかけに対する答えを得たり、考え方を理解する手助を得たり、この本で意欲的に学習したくなるであろう。本章では、この本がどのような論理的構成で書かれてあるかについて紹介しているので、講義あ

るいは自習のときに、本書をどのように使用すれば良いかがわかるであろう。統計学についてより理解し、自信をもちたい場合、特に、実験結果の解析について理解したい場合は、9章の170ページと、さらに、本書の表紙裏側に掲載している**統計的検定を選ぶためのフローチャート**をみるとよい。このチャートはあなたが必要とする統計的検定法を選ぶのに役立つ、後述するそれぞれの検定を行う方法がどの章に書かれているかがわかる。そのため、この本は、必要な時にはすぐ読めるように、研究室に常備するハンドブックとしても有用である。

1.2 なぜ、生物学者が統計を使用しなければならないのか？

なぜ、生物学者は、調査を実施したり実験を解析したりするときに統計を使用しなければならないか？例えば、物理学者が素粒子の質量を比較するときには、統計を必要としない。化学者が異なる酸のpHの値を比較するときには、統計を必要としない。または、エンジニアが異なる形の桁の強度を比較するときには、統計を必要としない。これらは、単一の観察を一般化できるものであるからだ。例えば、単一の中性子が単一の陽子より重い場合、それは他のすべての中子と陽子にもあてはまる。

しかしながら、金髪と黒髪の女性の身長を比較する場合、1人の金髪の女性と1人の黒髪の女性を比較して結論を出すことは明らかに愚かなことである。もしある金髪の女性の背が黒髪の女性より高い場合、この単一の観察から、金髪の女性が黒髪の女性より平均して背が高いかどうかについては、一般的に言うことはできない。男性と女性を比較する場合や、成長ホルモンを投与されたラットと投与されなかったラットを比較する場合などにおいても、同じようなことが言える。それはなぜだろうか？

答えはもちろん、すべてが同一の素粒子とは対照的に、ヒト（他の生物、器官、細胞も同様であるが）はすべてそれぞれ異なるからである。換言すれば、生物は**多様性（変動性）**を示すので、唯一の人、細胞、実験用に処置された生物が全体を代表するものではない。生物学者は数多くの仕事を試み、統計を使用することによって、多様性の問題を克服できるのである。

まずは、可能な限りの観察をすべてやりつくし、ある**標本***を**反復***して**観測**しなければならない。そうすることにより、**母集団***全体の状況について、適切な**推定値***を算定できる。例えば、金髪と黒髪の女性について1人ずつ測定するかわりに、それぞれほぼ20人ずつの女性について測定し、これらのデータから金髪と黒髪の2つのグループの平均身長を推定しなければならない。そうすれば、グループのうちの一方が平均してより高かったことが確実にわかるだろう（例えば、金髪の女性のほうがより高かったと）。金髪と黒髪の女性の間には本当に身長の差があるからそうなるのである。しかしながら、特定の人々を選んだことによってこの差が偶然に得られただけである可能性もある。こうした可能性を減らす

※標本（サンプル）

母集団から無作為抽出される部分集合で、そこから測定値が得られる。これらの測定値は母集団のパラメータ（母数）を推定するのに用いられる。

※反復データ点

個々のデータ点。

※母集団

この集団から測定値が得られるような、無限に大きな可能性をもった集団。母集団を特徴づけるパラメータ（母数）は、通常、標本のデータ点から推定されねばならない。

※推定値

母集団のパラメータ（母数）に対し、これを推定する値を標本データから算定した値。

ために、次の段階のプロセスを実行しなければならない。それは**統計的検定**（ここでの例の場合、2標本 t 検定）を実行することであり、導かれた結果が偶然出現した確率を計算する。この確率が十分低い場合、偶然起こる可能性がほとんどないと判断できて、結果が**有意**であると結論できる。ここでの例の場合、金髪の女性が黒髪の女性より有意に背が高いと結論できるのである。

1.3 統計学の論理と仮説検定を理解しよう

問題はすべて、統計学の論理が直観と逆転しているということに由来する。あなたが科学のテーマについて研究しているとき、通常は、興味深い何かが起こっているという仮説を設定する。例えば、金髪の女性は黒髪の女性より背が高いと仮定してから、この仮定を検証する。統計学の検定では、逆の論理を使う。つまり、そうした興味深いことは何も起こっていないという仮説、**帰無仮説**[※]を設定する。例えば、金髪と黒髪の女性は同じ平均身長であるとする仮説を設定してから、この帰無仮説が真であるかどうかについて検定を行う。統計的検定には、4つの主要なステップがある。

※帰無仮説

統計的仮説検定において、データが差を示していないとか、関連を示していないとかを前もって仮定すること。そのとき、統計的仮説検定では、そのような帰無仮説に適合するデータを偶然にとれる確率を算定する。

● ステップ 1：帰無仮説を設定する

あなたが設定しなければならない帰無仮説として、普通の科学的な仮説の正反対のものをとる。つまり、差がない、または関連がないと設定する（例えば、金髪と黒髪の女性の場合の帰無仮説として、どちらも同じ身長であるという仮説を設定する）。

● ステップ 2：検定統計量を算定する

算定する**検定統計量**は、標本の中に含まれる変動性（ばらつき）に対し、何らかの有効な結果（通常、グループ間の差や測定値の間の関連）の度合を吟味する。普通（例外もある）、この度合が大きくなればなるほど、検定統計量の値もますます大きくなる。

● ステップ 3：有意確率を算定する

検定統計量と標本のサイズがわかれば、帰無仮説が真実であるとした場合に、そうした測定値の結果が偶然得られる確率を算定できる。これは**有意確率**[※]として知られている。一般に、検定統計量と標本のサイズが大きくなればなるほど、有意確率はますます小さくなる。

● ステップ 4：帰無仮説を棄却するかどうかを決定する

最後のステップは、帰無仮説を棄却すべきかどうかを決定することである。慣例に従い、有意確率が $1/20$ （5%の確率、または 0.05）以下の場合に帰無仮説を棄却できると決められている（5%カットオフ）。有意確率が5%を超える場合は、帰無仮説を棄却する証拠がないことになる。しかし、このことは帰無仮説を支持する証拠があることを意味するわけではない。

※有意確率

帰無仮説が真であるとしたときに、ある1組の結果が得られる確率。

5%カットオフは、生物学者が実際に起こっていることについて見落とす可能性を減らすための、実は妥協の産物である。例えば、たとえ真の結果でなかったとしても、一見有意な結果を見出す可能性があって、その確率が1/20である。カットオフ点を、例えば1/100すなわち1%に下げれば、この類の見落としの過失をする（これは**第1種の誤差***として知られている）可能性を減らすことができる。他方、カットオフ点を下げることにより、真の結果を見出すのに失敗する（これは**第2種の誤差***として知られている）可能性は高くなる。

※第1種の誤差

本当は母集団の間に差や関連がないにもかかわらず、有意な差や関連が検出されること。

※第2種の誤差

本当は母集団の間に差や関連があるにもかかわらず、有意な差や関連を検出するのに失敗すること。

このような確率的な性質の結果のため、統計的検定を実行しても、実は何も決定的に証明することはできないのである。検定結果が有意であっても、そうではないかもしれない可能性がまだ残っている。同様に、ある選択した検定で有意でなかったとしても、その結果が真実であったかもしれない可能性がなおある。

1.4 なぜ、たくさんの統計的検定があるのか

統計的検定が生物学で必要であることを受け入れ、さらに統計学の逆転した論理に対処できたとしよう。それでも、たった1つのある統計的検定法を使うことによって、すべてのデータを解析できると期待するのは理にかなったことである。しかしながら、統計学の本には、しばしば多くのいろいろな検定法が掲載されている。なぜ、それほど多いのか？

※データ

観測や実験を通じて得た結果としての観測値や測定値。これらの値は世の中の事象を考察するために使用される。

それには、主に2つの理由がある。第1の理由は、物事を定量化するのに非常に多種多様な方法があって、その結果いろいろなタイプのデータが収集されるからである。そして、これらのデータは、いろいろ違った形で変動するからである。第2の理由は、収集したデータについて問いかけたいことが非常に多様だからである。

※測定値

ある数値で有意に表される1つの特性状態。

1.4.1 データ*のタイプ

● 測定値

生物の特性を定量化するために用いられる最もありふれた方法は、**測定値***（例えば、身長、質量またはpHのようなもの）を得て、**区間データ**を作成することである。こうしたデータとして、例えば体重（21.23 kgや34.651 kg）のように連続的に変化する場合と、例えばショウジョウバエの毛の数（12本または18本）のように不連続的に変化する場合とがある。2章で解説しているように、これらの測定値の多くは**正規分布***に従っている。いわゆる**パラメトリック検定***と呼ばれる一連の検定がある（3～5章を参照）が、パラメトリック検定では正規分布に従う測定値を取り扱う。他方、多くの測定値は正規分布に従わない。この場合、データが正規分布に従うようになるまで**変換**しなければならないか（7章）、またはそれが可能でないときは、データが正規分布に従わないことを仮定する別の一連の検定法、**ノンパラメトリック検定***（8章）を使用して解析しなければならない。

※正規分布

多くの因子から影響を受けた測定値が従う、左右対称でベル形をした分布形態。

※パラメトリック検定

データが正規分布しているとしたときの統計的検定。

※ノンパラメトリック検定

データが正規分布していると仮定しないで、そのかわりに観測値の順位を利用する統計的検定。

●

※順位

データ点に付与した数の順序。

● 順位*

他の場合として、現実の意味のある実測値がなくて、測定値に順位をつけることだけが可能な場合がある。この順位づけられたデータ、すなわち順序を示すデータには、雌鶏がつつく順番（例えば第1番目、第12番目）や、感染症の程度（例えば、なし、軽い、中ぐらい、重い）、またはアンケート・データの結果（例えば、1 = 悪いから、5 = 非常に良いまで）などが含まれる。この種のデータは、8章で述べるノンパラメトリック検定を使用して解析しなければならない。

※カテゴリー

1つのある数値で有意に表せない特性状態。

● カテゴリーデータ

生物の特性には、いかなる方法でも定量化するのが不可能な場合があって、これらを単に異なるカテゴリー*に分類することのみが可能な場合がある。例えば、鳥が異なる種、異なる色により分類される場合や、人が健康であるか、病気にかかっているかの場合、細胞が変異体であるか、非変異体であるかの場合、などである。この種のデータを定量化する唯一の方法は、それぞれのカテゴリーが生じる度数（頻度）*を数えることである。この種のデータは通常、 χ^2 （カイ2乗）検定（6章）を使用して解析される。

※度数（頻度）

ある特性状態が現れる回数。

1.4.2 問いかけのタイプ

統計的検定で調べたい問いかけには、主に、測定値の組の間に差があるだろうか？ 測定値の組の間に関連があるだろうか？の2つのタイプがある。

● 測定値の組の間の差に対する検定

2つのグループ間、または生物の種の間には差があるかどうか検定したいと考える状況はたくさんある。例えば、金髪と黒髪の女性の身長を比較する場合についてすでに述べた。より普通の状況として、実験を行う場合がある。実験用に処理された生物や細胞が、対照群と異なるかどうか知りたい場合である。または、1組のグループを対象にとられた2組の測定値を比較したい場合がある。例えば、人々がある治療を受ける前後を比較したい場合である。これらの問いかけに答えるための検定法は、3章と8章で説明している。いま一つは、いくつかの異なる種の生物（例えば、5つの異なる菌株）についてや、異なったタイプの処理を施されたもの（例えば、異なった濃度の窒素とリンで処理された小麦）について、互いに差があるかどうか知りたい場合がある。これらの問いかけに答える検定法は、4章と8章で述べている。最後に、異なるカテゴリーに分類される生物の度数が期待されるカテゴリーごとの度数と違うかどうかを決めたい場合がある。例えば、ラットが迷路において左の方へ向かうよりも右の方へより頻繁に向かう場合である。この種の問いかけに答える検定法は6章で解説している。

● 変数間の関連に対する検定

実行したいもう1つの検定は、生物や細胞の1つのグループを対象にして2つ

以上の測定値を得たとき、測定値がどのように**関連**しているかを調べることである。例えば、人の心拍数が血圧によってどう変化するか、体重が年齢とともにどう変化するか、または、ニューロン中のいろいろな陽イオンの濃度が相互にどう変化するかなどについて知りたい場合がある。この類の知識は、生物がどのように機能するのかを理解するのを手助けし、また、その生物に関することを予測できるようにもする。5章では、測定値の間の関係を定量化して、見かけ上の関連が本当かどうかを知るために、どのように統計的検定法が利用されるかについて記述している。6章では、カテゴリー化された形質、例えば、人の眼の色と毛髪の色とが関連しているかどうかをどのようにして決められるかについて述べる。

1.5 本書の使用方法

1.4節で述べた論理を発展させるとともに拡張し、**統計的検定を選ぶためのフローチャート**を作成した。これは、170ページの図9.1と、本書の表紙裏に掲載してある。このチャートの最上部からスタートして、次々にそれぞれの問いに答えながら下に降りてゆくと、必要とする統計的検定法にたどり着くことができる。完全に包括したものではないが、チャートには実質的に、大学生が必要とするすべての検定法が含まれている。そして、多くの生物学の研究者にとっては、研究生生活において使用するであろうすべての検定法が含まれている。このチャートを使用していただくことを前提として、本書を読んでいただくことで、以下の6つのことについて理解できるようになるだろう。

- ① いろいろな統計的検定法がそれぞれ答えることができる問いかけの種類。
- ② 本書では、例題を用意しながら、いつそれぞれの検定を使用すると有効かについて述べているので、ある特定の検定がどのような状況に対して適切であるかの範囲がわかる。
- ③ 基本的に各検定がどのように進められるかを述べているので、それぞれの検定の理論的根拠と数学的基礎がわかる。
- ④ 電卓、および/またはコンピュータベースの統計ソフト SPSS を使用することによって、それぞれの検定を実行する方法。
- ⑤ あなたが提起したい問いかけに答えることができるよう、より良く実験を立案するにはどうしたらよいか。
- ⑥ 統計的検定の結果を提示する方法。

しかしながら、まず変動性の原因を調べる必要があり、そして、この変動性がどのように取り扱われ、どのように定量化されるかを考察する必要がある。