

デザインはなぜ大切か

- 本章ではまず、良い科学研究をデザインする方法を身につけるとどういった利益があるのか（1.1 節）、身につけないとどういった害悪があるのか（1.2 節）を説明する。
- つぎに 1.3 節で、研究のデザインととったデータの解析法との関係について述べる。
- 実験をデザインするとき、最も重要なポイントとなるのは、ランダムなばらつきと交絡因子をどう処理するかである。1.4 節ではこれらの概念について導入的な話をする。
- どんな研究でも、母集団について何かを知りたいときは、母集団からサンプルをとって、サンプルだけを測定するのが普通である。1.5 節では、本書で実験デザインの専門用語がどのように使われるかを知ってもらうための第一歩として、サンプルを構成する「もの」の呼び名について説明する。

1.1 実験デザインはなぜ必要か

全員がそうだとはいえないが、生命学者は「実験デザイン」という言葉を見ると、安心して熟睡してしまうか、悲鳴をあげて逃げ出すかのどちらかである。実験デザインは、彼らの多くにとって、数学や統計学の授業の嫌な記憶を

呼び起こすもので、たいていは、何だか難しいもの、統計学者にまかせるべきもの、と考えられているのだ。これは間違った考えである！単純でも効果的な実験をデザインするのに、難しい数学は必要ない。それより実験デザインに関係するのは常識、生物学的な洞察力、そして念入りな計画である。ただ、ここで必要とされる常識は確かに特別な種類のものだし、いくつか基本的なルールもある。本書では、読者が効果的な実験デザインを考えられるようになることを目指しつつ、あまり苦痛を感じないですむ道を案内していきたいと思っている。

では、多くの生命科学者はなぜこうもデザインを考えるのを嫌がるのだろうか。その理由の一部はおそらく、実験のデザインに費やす時間があるなら、実験をするのに使ったほうが良い、と考えるのが楽だからだろう。とにかくわたしたちは生物学者なのだから、生物学に集中して、デザインや解析のことは統計学者に心配してもらおうじゃないか、というわけだ。このような態度から、2つのとんでもない思い込みが生まれてきた。これらは研究者の卵からベテランの大先生まで、あらゆる年齢層の生命科学者の口から聞くことができる。

思い込み1：どのようにデータをとるかは重要ではない。統計的な「応急処置」は必ずあるので、どのようにとったデータでも解析はできる。

もしこれが本当なら、どんなに素晴らしいことだろう。でも、残念ながらそうではない。

世の中には、かぞえきれないほど多くの統計検定が存在する。このため、どんな状況でもそれに合った検定が1つはあるに違いない、という誤った印象を抱きがちだ。しかし、統計検定と名のつくものには、必ずデータについて何らかの前提条件が設けられているので、データがまずその条件を満たさなければ、検定を適用する意味がない。それらの条件の中には、ごく少数の検定に特化されたきわめて特殊なものもある。もちろん、とったデータがそれらの条件を満たさなくても、データがもっている別の特徴を前提とする別の検定は見つかるかもしれない。だが、そちらの検定を使っても、最初に立てた問いとは別の問いへの答が出てくるだけだ。そして、その別の問いはほぼ確実に元の問いより

つまらないだろう（でなければ、なぜ最初にこちらの問いを立てなかったのかということになる）。それに、非常によく使われる統計検定の多くに共通する基本的な前提がいくつかあって、それらを見無視するというなら危険を覚悟でやるしかない。たとえば、たいていの統計検定は、データが**独立データ点 independent data point**と呼ばれるものから成っていることを前提としている（詳しいことは第5章で述べる）。とったデータがこの前提を満たしていなければ、使える統計検定は大幅に減ってしまうのだ。

このような結末は、最初に注意深くデザインしておくことで避けられる。それに、たいていの場合、注意深くデザインされた実験は、粗雑なデザインの実験にくらべて、より単純な統計解析法しか必要としない。だから、時間をかけて注意深くデザインすれば、あとになって必要以上に複雑な統計法を習うために膨大な時間を無駄にしないで済む。

実験で使う被験体のグループ（サンプルと呼ばれる）は、わたしたちが知りたいと思っているもっと大きな個体の集団（母集団と呼ばれる）を代表していなければならない。そのために大事なことは、サンプルを選ぶとき、サンプル内の個体どうしが何らかのつながりをもっていることがないように、つまり個体どうしが互いに独立であるように気をつけることだ。たとえば、人間の食の好みを調べているとき、同じ家族の5人から食の好みのデータをとっても、5つの独立データ点は得られないだろう。家族はたいてい長年食事を共にしてきているので、同じ家族の構成員どうしは、全くつながりのない二者どうしよりも、食の好みが似ていると考えられるからだ。同様に、同一人物から時間を置いて5回データををとっても、5つの独立データ点は得られない。ある人のある時点での好みは、もう少し時間が経ったときのその人の好みを知るための、有力な手がかりになる可能性が高

独立データ点：independent data point

独立データ点は、互いに関連のない個体からとられたデータである。

ある個体の測定値がそのとりうる値の範囲内でどのあたりになるかについて、別の個体の測定値が何の手がかりもあたえないとき、それら2つの測定値は独立である。

【問 1.1】 刑務所に収容されている人々の中に、宗教の教えを実践している人と左利きの人がそれぞれどのくらいいるかを知りたいとき、どのような母集団からサンプルをとるか。

【問 1.2】 前問で、とったサンプルに属する2人が、過去12ヵ月間、刑務所の同じ監房に入れられていたことがわかったとする。2人からとったデータは独立か。

いからだ。

思い込み2：とにかくデータをたくさんとりさえすれば、何かしらおもしろい結果が出てくるし、非常に微妙な効果でさえも検出できる。

データでいっぱいノートがあると安心だ。少なくとも、指導教授にそれを見せれば、あなたが頑張っていることを認めてもらえる。しかし、データの量は、データの質の代わりにはならない。少量でも注意深くとったデータは、強力な統計法で容易に解析できるし、そのようなデータからは生物学的に興味深い効果が検出できる可能性も高い。逆に、どんなにたくさんデータをとっても、質が悪ければ何の発見も得られないだろう。もっとつらいのは、それだけ多くのデータをとるためには、良質のより少ないデータをとるよりもはるかに多くの時間と資金を使わなければならないだろう、ということだ。

➡ 効果的な実験のデザインに必要なのは、数学の計算ではなく、生物学的な考察である。最初に注意深く実験をデザインしておけば、データを解析するときになって大量の汗をかいったり涙を流したりしなくてすむ。

1.2 貧弱なデザインの害悪

1.2.1 時間と金の無駄遣い

効果的にデザインされていない実験からは、注ぎ込んだ資金と努力に見合わない結果が得られるのがせいぜいで、最悪の場合は何の成果も得られない。せっかくデータをとっても、解析する方法が見つからなかったり、あなたが立てた問いに答えるものでなかったりすれば、貴重な時間と実験材料を無駄にしまったことは明らかだ。だが、これほど重大な失敗でなくても、貧弱にデザインされた実験には、思うように効果があがらない原因が他にもいろいろある。よくある誤解は、実験の規模は大きければ大きいほど良い、というものだ。

しかし、問いに効果的に答えるために本当に必要とするより多くのデータをとるのは、時間と金の無駄遣いである。逆に、ひどく時間のかかることが予想されたり、高価な消耗品が必要だったりすると、実験の規模をできるだけ小さくしたくなる。しかし、規模が小さすぎて、肝心の効果を検出できる可能性がほとんどないようでは、時間も金も節約したことにはならない。おそらくもう1回、今度はきちんと、実験をやり直さなければならなくなるからだ。このような問題は少し注意深く考えるだけで回避できる。そのやり方は後の章（とくに第6章）で見ることにしよう。

同様にありがちな間違いは、なぜそうするのかをよく考えもせずに、サンプルからできる限りさまざまな種類の測定値をとることだ。これは、ひいき目に見ても、多くの時間を費やして使い道のないデータをとっているとしか言えない。下手をすると、余計なデータにかまけて、自分の問いの解決にとって決定的に重要な情報を集めそこねているかもしれないし、本当に重要な情報の収集には十分な時間と集中力を使っていないかもしれない。欲張りは禁物だ。3つの問いに当て推量で答えるより、1つの問いに明確な答を出すことに集中しよう。

➡ 一刻も早く実験にとりかかりたい気持ちはわかるが、最初に時間をかけて実験の計画とデザインを練ることは、長い目で見れば時間と金の節約になる（それに言うまでもなく、あとで恥をかかずにすむ）。

1.2.2 倫理の問題

粗雑なデザインの実験は、努力と資金の無駄遣いである。それだけでも十分悪いが、生命科学の場合は、実験に動物（または動物からとった材料）を頻繁に使うので、問題がいっそう複雑になる。実験は動物にとって大きなストレスになると考えられる。単に実験室に入れておく、野外で観察する、といったことだけでも十分ストレスになるだろう。したがって、実験デザインに細心の注意を払い、動物に引き起こすストレスと苦しみを絶対最小限に抑えるよう努力し

なければならない。そこでよくやるのは、使う動物の数をできるだけ減らすことだが、ここでもやはり、意味のある結果が出てくることが見込まれるくらいの規模は確保しなければならない（第6章参照）。

何かの実験をしたいとき、方法はいくつもあることが多い。よく問題になるのは、複数の異なる処理を同一の個体に施すか、それとも異なる処理は別々の個体に施すか、ということだ。前者の場合、実験に使う動物の数は少なくすむだろうが、個々の動物はより長いあいだ実験室に拘束され、より頻繁に人にさわられることになる。後者の場合、より多くの動物を使うことになるが、拘束時間と人にさわられる回数は少なくすむ。どちらが良いのか、実験を始める前にじっくり比較検討し、動物の苦しみを最小限にする方法をとらなければならない。この種の問題については第10章でもっと詳しく考える。

倫理の問題は、室内で動物実験をする科学者だけに関係しているわけではない。野外実験も、実験者が環境に侵入することで、そこに棲む生き物たちに有害な影響を及ぼす可能性がある。そうは言ってもせいぜい研究対象の生物が影響を受けるだけだ、と思うかもしれない。しかしそれは根拠のない期待である。たとえば、山頂で地衣類のサンプルを集めれば、鳥の巣作りの邪魔になるかもしれない。それに、病原菌が採集用具に付着して、あちこちに運ばれてしまうかもしれない。このようなわけで、倫理は本書を通じて何度もくり返し現れるテーマとなるだろう。

➡ **劣悪な実験デザインで時とエネルギーを無駄遣いするのが愚かであることは言うまでもない。それに加えて、絶対に必要とされる以上に人や動物を苦しめたり、生態系を乱したりするのが言語道断であることは、どんなに強調してもしすぎることはない。**

1.3 実験デザインと統計解析法の関係

本書はデータの統計解析については詳しく述べない、と言ったら、読者はび

っくりするかもしれない。著者がそう言うからには、統計検定は実験デザインにとって重要ではなく、それについては考えなくてもよいのだろうか。そんなことはない！ 実験デザインと統計はじつは密接につながっていて、データをとる前から、データ解析のために使う統計法について考えておくことは不可欠だ。すでに述べたとおり、どの統計法も、それが解析できるデータの種類や、検証できる仮説の種類について、他の検定とは少し異なる前提条件を定めている。だから、統計検定をおこなうときは、あなたのとったデータを解析でき、あなたの検証したい仮説を調べられるとわかっている検定法を使うことが非常に重要である。これを確実にするための唯一の方法は、とり終わったデータをどのように解析するか、前もって決めておくことだ。本書では、統計の細部には立ち入らない代わりに、デザインする上で統計について考えることが決定的に重要なときは、そのことを強調して注意を促すことにしたい。

本書が統計に重点を置かない理由は2つある。そのひとつは単に、すぐれた統計学の本がすでに何冊かあって、どの検定にするか決める段になったら、それらを参照すればいいからだ（参考のために、著者たちが使った本を巻末に何冊か挙げてある）。しかし、もうひとつの理由のほうが重要で、著者たちの固い信念によれば、実験デザインというものは、あなたが使う統計検定をはるかに超えるものだからだ。このことは、細かく込み入った統計の議論の中で、とかく見失われがちな点である。実験をデザインするとは、とったデータの解析に使われる統計法の仕組みを知ることだけではない。それ以上に、科学的なものの考え方を学ぶことである。自分のデータに自信をもつことである。自分が測っていると思っているものを本当に測っていることと知っていることである。また、特定のタイプの実験から何が結論できて、何が結論できないかを知っていることである。

統計解析の考え方には哲学的立場から2つの学派があり、一方は帰無仮説の検証に使われる p 値を重視し、もう一方（ベイズ統計）はモデルを当てはめることを重視している。

本書は、統計に関しては帰無仮説の検証という文脈で話をしているが、ベイズ流のアプローチをするときも、実験デザインへの取り組み方はまったく変わ

らないはずだ。

➡ 実験をデザインするとは、研究対象であるシステムの生物学的実態について考えることだ。したがって、生物学実験を計画するのに一番適しているのは生物学者、つまりあなた自身である。

1.4 良い実験デザインはなぜ特に生命科学者にとって重要なのか

実験について考えるとき（したがって本書全体を通じて）、しょっちゅうぶつかる2つの重要な概念は、交絡因子とランダムなばらつきである。実際、実験デザインの二大目標は、交絡因子を考慮することと、ランダムなばらつきを最小限に抑えることだと言ってよい。これらのテーマはそれぞれ第3章と第4章で詳しく扱うが、ここでこれらを取り上げるのは、実験デザインがなぜ特に生命科学者にとって重要なのかをわかってもらうためである。

1.4.1 ランダムなばらつき

ランダムなばらつき random variation は、個体間のばらつき、個体間差異、処理内差異、ノイズなどとも呼ばれている。これは単に、サンプル内の個体（たとえば動物、植物、人、実験区画、組織標本など）が、わたしたちが関心をもっている理由以外の理由で、互いにどのくらい異なっているかを表すものである。このようなばらつきが研究に及ぼす影響については、第4章で詳しく述べる。

ランダムなばらつき：random variation
関心の的ではない因子が原因で起こるばらつきをランダムなばらつき（またはノイズ）という。実験をデザインするとは、ランダムなばらつきを取り除いたり制御したりすることによって、関心の的である因子の影響がよく見えるようにすることである。

たとえば、10歳の男の子は全員が同じ身長ではない。今、わたしたちの研究の目的が、10歳男児の身長、国による差異を明らかにすることだとしよ