



## 実践統計解析【第6回】

# 外れ値にご用心

川瀬 雅也<sup>1\*</sup>・松田 史生<sup>2</sup>

Aさん：先輩，先日教えてもらいながらやった培養実験のデータを見てもらえますか。

B君：抗生物質生産菌株の生産量比較だね。もちろん、いいよー。

Aさん：3回実験をしたんですけど，3回目がおかしいかなと思って，土曜日に4回目をやりましたよね（表1）。

表1. 抗生物質生産量（mg/L）の測定結果

微生物	1回目	2回目	3回目	4回目
A株	10.2	9.8	10.6	10.8
B株	8.1	9.5	10.5	8.5

A株の方が高生産かなと思って検定してみると

```
> a <- c(10.2, 9.8, 10.6, 10.8)
```

```
> b <- c(8.1, 9.5, 10.5, 8.5)
```

```
> t.test(a, b)
```

Welch Two Sample t-test

```
t = 2.063, df = 3.9915, p-value = 0.1082
```

で有意差はなかったんですが，3回目がおかしいように思えて外してみると，

```
> a <- c(10.2, 9.8, 10.8)
```

```
> b <- c(8.1, 9.5, 8.5)
```

```
> t.test(a, b)
```

Welch Two Sample t-test

```
t = 3.0857, df = 3.5749, p-value = 0.04268
```

となって，有意差が出るんです。3回目のデータはどう取り扱えばいいんですか。

B君：困ったな。以前なら，外してしまおうと言うところだけど，外すとまずい気もするし。ところで，指導してもらっているC先生はなんて言っているの？

Aさん：「外してもいいんじゃない？」なんですけど…

B君：X教授に聞いてみよう。

このような状況は，皆さんの研究室でもよく起こっているのでは。その時，皆さんはどう対処されているだろうか。

### データの正確さ

X教授：今日はどうしたのかな。

Aさん：実は……という訳なんです。

B君：僕も，統計学を勉強してみても，簡単にデータを外せないような気がして。

X教授：B君も，随分成長したじゃないか。

B君：そうですね。

Aさん：でも，土曜の実験の後……。

B君：ダメ，ダメ。

X教授：このまま雑談でも，一向に構わないが，話を元に戻そうか。つまり，3回目のデータを外れ値と扱っていいかどうかだね。もし，3回目のB株のデータだけがおかしいとしたら，どうなるかな。

```
> a <- c(10.2, 9.8, 10.6, 10.8)
```

```
> b <- c(8.1, 9.5, 8.5)
```

```
> t.test(a, b)
```

Welch Two Sample t-test

```
t = 3.498, df = 3.1278, p-value = 0.03702
```

で，やっぱり，有意差は出るね。

Aさん：B株の3回目だけ外せばいいんですね。

X教授：そう簡単にはいかないんだ。少しデータについてみてみよう。そもそも，測定データには誤差が付きものだというの聞いたことがあるね。

Aさん・B君：はい

X教授：「誤差とは何か」知っているかね。

B君：正しい値との差だと習いました。

X教授：その通り。でも，正しい値は，ほとんどの場合知ることはできないね。そこで，いくつもデータをとって，その平均を正しい値と考えたんだ。誤差が正規分布に従うというのは中心極限定理で示されているから成り立つんだが，ここまでの話で，一つ不正確な個所があるんだが分かるかな。

Aさん：分かりません。

B君：僕も。

X教授：少し教科書的な説明になるが，この図（図1）を見てもらおうか。

著者紹介 <sup>1\*</sup>長浜バイオ大学（教授） E-mail: m\_kawase@nagahama-i-bio.ac.jp

<sup>2</sup>大阪大学大学院情報科学研究科（准教授）

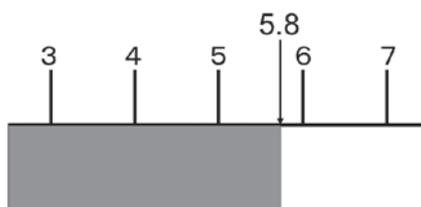


図1. 物差しを使い、ものの長さを測定  
最小目盛りの1桁下まで読み取る

よく実験で、測定器具の最小目盛りの1桁下まで読み取りなさいと言われなかったかい？

B君：いつも、言われています。

X教授：たとえば、この図では5.8と読んでいるが、2人はどう読む？

Aさん：私も5.8くらいかなと思います。

B君：僕は5.7の方がいいかなと思う。

X教授：そうだろう。人によって、最後のケタの読みが違ってくる。この他にも、サンプリングや試薬調製、測定機器の操作など人が関わる場面はいくつもあると思う。

B君：確かに。

X教授：このように、人が関わることで生じてくる誤差を偶然誤差と言うんだ。この偶然誤差が正規分布に従うんだ。では他に、誤差を生む要因はないかな？

Aさん：測定機器の調子が悪いときです。

X教授：そうだね。測定機器の調整不足や、測定方法でも出てくる。たとえば、この方法は少し高めの数値が出るなんて聞いたことがないかね？

B君：あります。

X教授：測定機器や測定法により生じる誤差は系統誤差と言うんだ。測定機器の場合は十分調整してやれば、誤差は無視できる程度にできるし、測定法の場合は誤差の出る方向（プラスかマイナスか）が決まっているので補正が可能だ。しかし、偶然誤差は生じる方向は決まっていないうし、大きさもまちまちなので補正できない。だから、測定誤差は避けられないんだ。

Aさん：誤差に2種類あるなんて知りませんでした。

B君：指導してもらっているC先生は、Aさんの使っている方法は精度の高い方法だから正確な値が出ると言っていたよね。でも、誤差は生じるんですよ。

X教授：本当にC先生はそう言ったのかね？とんでもない話だ、まったく……。

Aさん：落ち着いてください。

X教授：「精度が高い」と「正確な値が出る」ということ

はまったく関係がないんだ。よく間違えるので、説明しておこう。大事な三つの言葉があるんだ。

**正確さ (accuracy)**：測定値と真の値がどの程度一致しているかの程度。つまり、両者の差が小さいほど、正確だということになる。

**精度 (precision)**：同じ量を繰り返し測定した場合、得られるデータの一致の程度。つまり、ある量を3回測定した場合、3回の測定データの間の差が小さいほど精度が高いとなる。

**感度 (sensitivity)**：二つの異なった量を区別できる程度のことであり、言い換えれば、二つの量がどの程度違っていれば区別可能かを表す量である。

X教授：正確さと精度、感度はまったく違うことを表しているんだ。

B君：まったく知りませんでした。

X教授：有効数字という言葉を知ったことがあるね？

Aさん：はい。でも、なんかよく分からないんです。

X教授：測定値には、必ず誤差や測定の限界があるね。

だから測定で得られるデータには信頼性の限界があると言っていい。つまり、“何桁目までが信頼できるか”ということが、測定値の取り扱いのうえで重要だ。この信頼できる桁数（測定の精度によって保障される桁数）を有効数字と言っているんだ。有効数字には、

○有効数字の最後の桁には少なくとも $\pm 1$ 程度の不確かさがある。

○有効数字の桁数は小数点の位置とは無関係である。

○演算（加減乗除）において、得られる結果の有効数字は演算に用いた中でもっとも少ない桁数に一致させる。

という約束がある。たとえば、1.23と0.123は同じ3桁だね。1.23  $\times 10^0$ と1.23  $\times 10^{-1}$ とすればよく分かると思う。ただし、三つ目のルールは要注意だ。

たとえば、メスシリンダーやピペットを使って容量を測定した場合、最小目盛りの1桁下までを通常は読み取り、ここまでの有効数字となる。今、最小目盛りが0.1 mLであったとする。読み取った値が、1.11 mL、12.56 mLというケースが当然生じてくる。このとき、有効数字の桁は3桁と4桁となり、同じ測定器具を使っているのに、桁を揃えることができなくなる。このような場合は、有効数字は小数点以下2桁目までというように決めることになる。演算の時も同じ測定器具だ

けしか使わない時は、小数点以下の桁数で揃えるほうがいい。個々のケースで考えないでね。

Aさん・B君：なるほど。

X教授：精度や感度は、この有効数字の桁数と関係する量なんだ。

少し横道にそれたけど、外れ値の話に戻ろう。

### 外れ値の判定

X教授：Aさんの4回分のデータをもう一度、よく見てみよう。この時は以前、説明した箱ひげ図（図2～4）を書いてみるといい。覚えているかな？

Aさん：何となく。

B君：勉強しておきます。

X教授：何度も繰り返すことが統計学の理解には必要だから、いい機会と思って勉強するといい。

```
> a <- c(10.2, 9.8, 10.6, 10.8)
```

```
> b <- c(8.1, 9.5, 10.5, 8.5)
```

```
> boxplot(a,b)
```

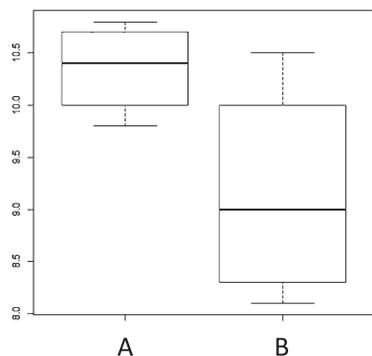


図2. 4回のデータの箱ひげ図

3回目の結果を外したらどうなるかやってみよう。

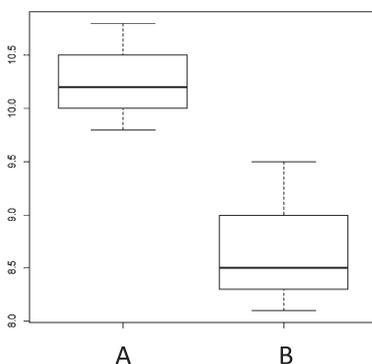


図3. 3回目のデータを除いた箱ひげ図

B株の3回目のデータだけを除いた場合は、

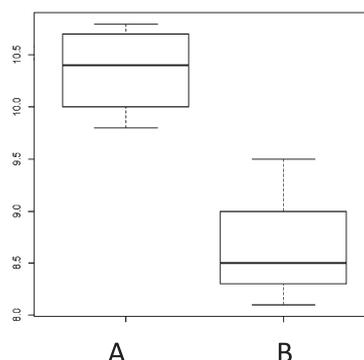


図4. B株の3回目のデータのみを除いた箱ひげ図

X教授：どうかな？

B君：やっぱり、B株の3回目のデータがないとまともってありますね。外れ値と言われるとそうだと思います。

Aさん：私も。

X教授：2株で差が出るという期待を持っているから、外れ値であってほしいと思えるのではないかな。

Aさん：確かにそうです。

X教授：2株に抗生物質の生産力の違いがあるとは、まだ分かっていないんだろう？分かっていないんだから、期待を持つのは悪いことではないが、期待の通りになる方向で判断するのはよくないな。まず、冷静に評価するという態度が必要だ。

B君：仰せの通りです。

X教授：客観的に判断しろと言われても困ると思うので、一つ、外れ値かどうかの判断を助けてくれる方法を教えよう。Dixon法(Qテストとのよばれることもある)という方法だ<sup>1,2)</sup>。

Aさん・B君：聞いたことがない方法です。

X教授：この方法は、データの母集団が正規分布であることを前提にしていることを忘れないでほしいんだ。測定誤差は正規分布に従うと言ったね。

Aさん：はい。

X教授：では、どんなデータにでも使えるかというところではない。たとえば、生物の行動や成長などはポアソン分布に従うケースも多々あると聞いている。このようなデータには使えないんだ。生物が作る物質の量を測定する場合でも、生産量の測定値の誤差は正規分布に従うけれど、生産量そのものが正規分布になって



いるのかどうか分からないね。つまり、Xという株があり、ある一つのフラスコ中の生産量を測定したとき、測定値の分布は正規分布になるが、いくつもフラスコがあり、すべてのフラスコの測定値を並べたときに、正規分布になるかどうかは分からない。生物に関するデータを扱う時は慎重になる必要があるという例だね。

B君：なるほど、この菌XはX教授みたいにひねくれものですね。

X教授：よく分かってきたね。今度、ゆっくりと話をしようじゃないか。

B君：冗談です。おゆるしを。

Aさん：早く先に進みましょう。

X教授：そうしよう。t-検定を行うんだから正規分布としてDixon法を使ってもいいことにしよう。

Dixon法での判定法は、まず、Q値という値を計算する。異常値と思われる数値をAb、Abにいちばん近い値をNn、データの最大値max、データの最小値minとして、Q値は

$$Q = \frac{|Ab - Nn|}{\max - \min}$$

で定義される。AさんのBの菌のデータについて計算すると、

$$Q = \frac{|10.5 - 9.5|}{10.5 - 8.1} = 0.417$$

となるね。これを臨界値というものと比べるんだ。

Aさん：臨界値ですか。

X教授：臨界値の表（表2）<sup>3)</sup>があるから安心して。

表2 Dixon法における臨界値

測定回数	信頼限界	
	90%	95%
3	0.941	0.970
4	0.765	0.829
5	0.642	0.710

（文献3）p. 84より抜粋

信頼限界とは、どの程度信頼できる区間を想定しているかということだ。以前説明した信頼区間の両端を信頼限界といったのを思い出してくれると分かると思う。

測定回数は4回だから、その値を見ると90%の信頼

限界が0.761、95%の信頼限界が0.829で、どちらの値もQ値より大きいね。つまり、外れ値を考慮することができないという結果なんだ。

Aさん：そうなんですか。何か、見た目の感じだけでいうと、随分、他のデータと比べて大きいから外してもよさそうに思えたんですが。

B君：C先生は、どうして外してもいいと言ったんだろう。

X教授：多分、今、Aさんが言ったように見た目以外れていそうだとということと、2株で差が出るという期待を持っているからじゃないかな。多くの研究者は、仮説を持って研究に当たるね。当然、必要なことなんだが、仮説の通りになればいいという意識が強くなると、データを無意識に合わせようとしてしまう。その結果、少し合わないデータを外れ値としてしまうんだ。こんなことがないように、十分注意しないといけないね。

ただ、Dixon法に頼りきるのはよくないことも知っておいてほしいんだ。もし、3回目の実験のとき、何か、気が付かない原因で測定装置の不具合があったとか、この時使った試薬に不備があったなどの可能性も0ではないと思う。

Aさん：うーん。あ、確かこのサンプルって、B先輩がかわりにサンプリングしてくれたやつですよ。

B君：そういえば…内部標準が見当たらずに自分の古いやつを使ったんだわ…

X教授：原因がはっきりしているときは、その回のデータは除外できるかな。とにかく、データの取り扱い方は、統計的な考察だけで決まるものではないことを十分理解しておいてほしいんだ。何も不都合なく測定したデータについては、統計的な処理が有効だと考えるといいと思うよ。

Aさん：分かりました。これは、内部標準を作りなおして土日にも再実験ですね。解決法は、これしかないと思いますから。先輩も付き合ってください！

X教授：それが一番いい解決法だね。

## 文 献

- 1) Dean, R. B. and Dixon, W. J.: *Anal. Chem.*, **23**, 636 (1951).
- 2) Rorabacher, D. B.: *Anal. Chem.*, **63**, 139 (1991).
- 3) 化学同人編集部編：実験データを正しく扱うために、化学同人 (2007).

（【第7回】は95巻4号に掲載予定です）