

## お詫びと訂正

---

書籍『生命科学・生物工学のための間違いから学ぶ実践統計解析—R・Pythonによるデータ処理事始め』（近代科学社）のpp. 29–30に誤記がございました。

下記の通り、**太字網掛け部分**を訂正いたしますので、ご確認ください。

検出力は十分ある。では、このデータを100倍して、**次のようにCのデータを少し変えてみよう。差があるように見えるね。**

```
> C <- c(122, 138, 131)
> D <- c(141, 152, 137)
> t.test(C, D, var.equal = T)
```

**X教授**：**t = -2.0166, df = 4, p-value = 0.1139**となりt-検定の結果は有意差なしとなった。  
検出力を見てみると、

```
> mean(C) - mean(D)
[1] -13
> sqrt((2*var(C) + 2*var(D))/4)
[1] 7.895146
> power.t.test(n = 3, d = -13, sd = 7.895146)
power = 0.3403129
```

検出力不足になっている。つまりこの実験系では、データは3個では不足ということになるんだ。いくつかのデータが必要かと言うと、

```
> power.t.test(d = -13, sd = 7.895146, power = 0.8)
n = 6.896426
```

**7**個ずつデータが必要と言う結果になる。このデータはばらつきが大きいけど、もし、データの標準偏差（SD）がもう少し小さな場合では（SD = **4**の場合）、

```
> power.t.test(d = -13, sd = 4, power = 0.8)
n = 2.844429
```

データは**3**個ずつでいいわけだ。検定の話をしていないので、分からないところもあると思うが、「必要なデータ数は得られたデータの質によって違って来る」ということを、今は分かってくれればいい。どうかな。

---