と呼ばれる新分野を開拓したのです。フィッシャーは来日したこともあります。彼は農事試験場において、生物進化や遺伝を研究し、多くの業績を残しています。その研究の中で、彼は様々なデータ分析の手法を編み出しましたが、その1つがこれから話す分散分析なのです。

試験で得たデータを分析するためには、推計学が出てきたのは必然だったんですね。

そうじゃ。それでは、さっそく例としてトマトのハウス栽培の収益を考えてみよう。

ケース1

トマトのハウス栽培

栽培のための肥料は、4種（A、B、C、D）を使い、それ以外の条件はできるだけ均一化することにして、同一面積の区画からいくらの収益があったかを調べたところ、次のような結果になりました。データは、各肥料について8区画から得ているとします。

表2-1　8区画からの収益　（単位：万円）

肥料A	肥料B	肥料C	肥料D
8.40	6.79	5.79	7.30
4.44	5.74	8.65	9.20
7.71	5.02	10.38	7.71
7.23	6.71	6.25	8.14
3.57	8.57	8.22	7.14
3.53	8.49	7.15	11.35
3.77	7.80	9.32	6.77
7.31	5.29	5.62	7.97

さて、肥料の種類と収益はどんな関係があるのでしょうか。

今注目している収益に関連すると思われる項目はなんじゃ？

肥料？

そうじゃ。まあ、肥料以外にも本当は天気とかいろいろ絡んでくるかもしれんが、今回は、他の要因は均一に設定しているという前提で、肥料1つに絞って考えるんじゃ。

つまり、1つの要因だけみればいいのですね、うん、なんだかできそう。

このように、一因子（ここでは肥料）の効果を調べる方法を「**一元配置の分散分析**」といいます。一般に、統計資料に影響を及ぼす要因はさまざまありますが、今回の例のように、ある1つの要因を除いて、他の条件はできるだけ均一に設定したとします。理想的には、このときその要因だけがデータに影響を及ぼすはずということです。

ポイント

・肥料に相当するもの…「因子」
・肥料A、B、Cのような種類…「水準」
一要因にだけ着目して、資料への影響を分析する方法が「一元配置の分析」

まず、分散分析では、偏差に注目したのを覚えているかね？

えっと、ヘンサ、へんさ……。あ、偏差ですね。

一元配置の分散分析でも偏差は絶対に使うので、しっかり覚えておくのじゃぞ！
ちなみに、前の表から平均は何かな？

あ、平均はわかります。全部足してデータの個数32で割ればいいんですよね。7.10です。

「**偏差**」とは、データの値と平均値の差でしたね。肥料の効果は、その偏差の中に現れると考えます。実際、成長により肥料を与えられたトマトは、平均値よりも大きく成長するはずであり、結果として偏差は大きくなります。逆の場合もまたそう考えられます。

この例で、各データについて、実際にこの偏差を計算してみましょう。平均値は7.10なので、各偏差は次のようになります。この偏差の中に、肥料の効果が現れているというのがポイントです。

表2-2　偏差：各データ－全体の平均7.10

（単位：万円）

肥料A	肥料B	肥料C	肥料D
1.30	－0.31	－1.31	0.20
－2.66	－1.36	1.55	2.10
0.61	－2.08	3.28	0.61
0.13	－0.39	－0.85	1.04
－3.53	1.47	1.12	0.04
－3.57	1.39	0.05	4.25
－3.33	0.70	2.22	－0.33
0.21	－1.81	－1.48	0.87

肥料の効果が現れているってどういうことですか？