

■ 慶應大学経済学部の第3問に、データの分析の問題が出題されている(2016年2月13日実施)。

[3] ある野生動物を10匹捕獲し、0から9の番号で区別して体長と体重を記録したところ以下の表のようになった。体長と体重の単位は省略する。

番号	0	1	2	3	4	5	6	7	8	9
体長	60	66	52	69	54	72	74	60	58	61
体重	5.5	5.7	5.9	5.9	6.0	6.2	6.2	6.4	6.5	6.7

- (1) この10匹の体長の最小値は(34)(35)、最大値は(36)(37)である。
- (2) この10匹は5匹ずつAとBの2種類に分類できる。1つの種類の中では体長と体重は正の相関を持つ。10匹の体長と体重の相関係数は0.05以下だが、種類Aの5匹に限れば0.95以上であり、種類Bの5匹も0.95以上である。また、番号2の個体は種類Bである。このとき、種類Aの5匹の番号は小さいほうから順に(38)、(39)、(40)、(41)、(42)であり、その5匹の体長の平均値は(43)(44)(45)となる。
- (3) 10匹のうち体長の大きいほうから5匹の体長の平均値は(46)(47)(48)である。(2)で求めた平均値と異なるのは、体長の大きい5匹のうち番号(49)の個体が種類Bだからである。
- (4) (2)で求めた種類Aの5匹の体重の偏差と体長の偏差の積の和は6.6、体重の偏差の2乗の和の平方根は小数第3位を四捨五入すると0.62、体長の偏差の2乗の和の平方根は小数第1位を四捨五入すると(50)(51)である。

■ (2)のA, Bのグループ分けがポイントで、これができなければ以降は壊滅する。相関係数だけから2グループに分けよと言うことだが、計算して求めよということではないはずだ。答だけ出せばよいので「こうだろう」という予想で突き進むしかないということだろうが、それはいかなるものか。

これを書いている現段階では、東進予備校が「解答」をWeb上にupしていて、該当箇所は右の通り。

解答の苦労の跡が忍ばれる。

(2) 最も体長・体重AとBに分類する

体長	52	54	58	60	60	61	66	69	72	74
番号	2	4	8	0	7	9	1	3	5	6

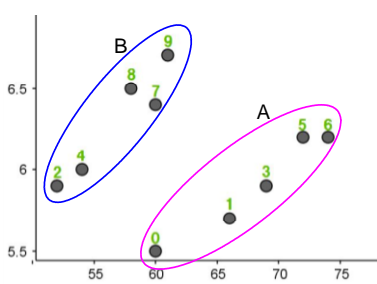
体重	5.5	5.7	5.9	5.9	6.0	6.2	6.2	6.4	6.5	6.7
番号	0	1	2	3	4	5	6	7	8	9

相関係数 $+0.45$ 以上となったのは
 「体長が大きいほど、体重も大きくなる」
 というように分類するが、A
 「2, 4, 8, 7, 9」と「0, 1, 3, 5, 6」
 が妥当かと思われ、
 (正確には、計算は正確な計算機、ここでは省略)
 条件の番号は B から
 A 「0, 1, 3, 5, 6」

■ さて、これはどう解くのがよいのか? 多分、bestな方法は「散布図」である。

これを見ると2本の「直線」が見えてくる。

番号2の個体を含む上側の集団がB, 下側の集団がAである。なお、「直線」とは各集団の回帰直線のことである。



実際、A, B 2集団の相関係数はそれぞれ、0.9741, 0.9517 で条件を満たしているとはいえ、解答時間内に計算して確認できることではない。

■ この試験科目はIA+IIBだが、数学IAのみを課するような大学などでは、「データの分析」の問題を出題せざるを得ない事情もあろう。

しかし、きちんとした解答の検証が受験中に困難なこの問題が入試問題として適切かどうか、疑問の余地がある。