第35回　第５章　データの分析

データと度数分布表

監修・執筆
湯浅弘一

今回学ぶこと

クラスであるテストを行ったとき、このテストが難しいかどうかは受験者の点数から見ればわかります。つまり、低得点ばかりならそのテストは、難しかったとわかるし、高得点ばかりならそのテストは、易しかったとわかります。今回はその点数からテストの難易度を分析してみましょう。

学習のポイント

①データから度数分布表を作る
②データからヒストグラムを作る
③データの相対度数を求める

ポイント1　データから度数分布表を作る

まず、「データ」とは何でしょう？　よく使われるデータという言葉。データ＝資料、と辞書に書いてあったりします。ここでいうデータとは、数的資料ということになります。

例えば、10人のグループに英単語のテストをしたとします。問題は50問です。

以下はその結果です。

生徒番号	1	2	3	4	5	6	7	8	9	10
得点	16	40	38	36	40	22	44	38	32	28

上の表をザッと見ただけでは、このテストの難しさはわかりません。

そこでデータを10点刻みで振り分けたのが、次ページの度数分布表です。

度数分布表

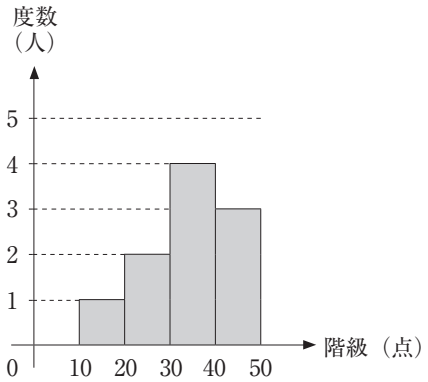
得点の階級（点）	階級値（点）	度数（人）
0 以上～ 10 未満	5	0
10 ～ 20	15	1
20 ～ 30	25	2
30 ～ 40	35	4
40 ～ 50	45	3
計	―――	10

この各区間を **階級** 。階級の真ん中の値を **階級値** 。その階級に当てはまるデータの個数を **度数** といいます。

この度数分布表から、このテストはどうやら高得点にデータが集中しているようです。

ポイント2　データからヒストグラムを作る

ポイント１の度数分布表をヒストグラムに表してみましょう。



これで高得点に人が集まっている様子がわかります。

ポイント3　データの相対度数を求める

相対度数とは、$\dfrac{\text{各階級の度数}}{\text{合計}}$ のことで、全体に占める割合がわかります。

相対度数分布表

得点の階級（点）	階級値（点）	度数（人）	相対度数	
0 以上～ 10 未満	5	0	0	← 0 ÷ 10
10 ～ 20	15	1	0.1	← 1 ÷ 10
20 ～ 30	25	2	0.2	← 2 ÷ 10
30 ～ 40	35	4	0.4	← 4 ÷ 10
40 ～ 50	45	3	0.3	← 3 ÷ 10
計	―――	10	1	

パーセントで表せば

得点の階級（点）	相対度数
0 以上～ 10 未満	0%
10 ～ 20	10%
20 ～ 30	20%
30 ～ 40	40%
40 ～ 50	30%

上位 2 階級に全体の 70%がいます !!

これらのデータから、この英単語のテストは易しい！　と分析できます。