

## センター試験への心得（データの分析）

本章は、統計という考え方の中で、必要となる数学があまり高度でないものだけを扱っている。したがって、学習のポイントは、

平均値、中央値、四分位数、箱ひげ図という最も初歩的な考え

分散、標準偏差、相関係数という少しだけ進んだ数学的な概念

とを一旦は区別して（つまりそれぞれに対し、別々の覚悟をもって）学習することが大切である。前者からの出題は、短い時間内に手計算で解決できる、という制約条件から、中央値や平均値を巡るクイズ的な問題か、四分位数や箱ひげ図の統計的意味の直観的理解を問うものに限られるとあってよい。

他方、後者からの出題も、本格的な数のデータの処理となると、計算があまりに繁雑になるため、ごく少数のデータについて、概念の定義と、概念間の関係についての知識があれば解決が容易なものに限られるとあってよいだろう。必要とされる知識を具体的に述べるならば、 $n$  個のデータ  $x_1, x_2, \dots, x_n$  についての平均値を  $m$ 、分散を  $s^2$  とすると

$$\begin{cases} m = \frac{1}{n}(x_1 + x_2 + \dots + x_n) \\ s^2 = \frac{1}{n}\{(x_1 - m)^2 + (x_2 - m)^2 + \dots + (x_n - m)^2\} \end{cases}$$

という定義と、これから導かれる

$$\begin{cases} x_1 + x_2 + \dots + x_n = nm \\ s^2 = \frac{1}{n}(x_1^2 + x_2^2 + \dots + x_n^2) - m^2 \end{cases}$$

といった関係式である。実用的には、仮平均の考え方が以外に有力である。統計という数学的手法は、コンピュータの発達と普及に伴って、今後ますます大きな重要性をもつであろう。