

3 級

第
1
章

第 1 章

手法編

1. データの取り方・まとめ方 (1)

問 1

一般的に母集団は，分布の中心の位置とばらつきの程度で，その状態を知ることができる．分布の中心の位置を表す統計量としてもっとも適切なものをひとつ選べ．

- ア．平均値
- イ．偏差
- ウ．最大値
- エ．相関係数

2

ばらつきの程度を表す統計量としてもっとも適切なものをひとつ選べ．

- ア．中央値
- イ．算術平均
- ウ．標準偏差
- エ．寄与率

問 3

データの最大値と最小値の差を何というか．これは，簡単に求められるばらつきの尺度であり，データの数が 10 以下の場合によく用いられる．もっとも適切なものをひとつ選べ．

- ア. 中央値
- イ. 範囲
- ウ. 不偏分散
- エ. 標準偏差

解説

この問題は、母集団から抽出したサンプルによって得られたデータから計算で得られる統計量について問うものである。

本問では、大問2と関連して、分布の中心位置を表す統計量、ばらつきの大きさを表す統計量について理解しているかどうかポイントである。

解答

問1 ア 問2 ウ 問3 イ

問1

分布の中心の位置を表す代表的な統計量として、平均値 $\bar{x}$ 、メディアン $\tilde{x}$ がある。よって、正解はアである。

中心の位置を表す統計量には、ほかに、ミッドレンジ M （最大値と最小値の平均）や出現頻度がもっとも多いデータを表す最頻値（モード）がある。

問2

ばらつきの程度（大きさ）を表す統計量として、不偏分散 V 、標準偏差 s 、範囲 R がある。よって、正解はウである。

問 3

取られたサンプルのデータの中の最大値と最小値を用いて範囲 R は,

$$R = \text{最大値} - \text{最小値}$$

で求めることができる. これもばらつきを表す統計量である. よって, 正解はイである.

範囲 R は, 標準偏差 s を求める場合に必要な 2 乗の計算や $\sqrt{\quad}$ (平方根) の計算の必要もなく, 計算が楽であるが, データのうち, 二つのデータしか使わないので, データ数が大きくなると効率が悪くなるため, 10 以下の場合に使われることが多い. 不偏分散 V , 標準偏差 s は, データ数が $n = 2$ からいくらでも大きなデータ数でも対応できる.