

1 範囲と四分位範囲の違いと特徴について述べよ。

(解答)

範囲はデータの最大値と最小値の差であるのに対し、四分位範囲はデータの第3四分位数と第1四分位数の差である。

範囲をみることで、データ全体の大きさを確認することができる一方で、四分位範囲をみることで、外れ値などの影響を受けづらいうデータの大きさを確認することができる。

- ★ 2 右の表は、ある農家の昨年まで10年間の面積10^{アール}aあたりの米の収穫量をまとめたものである。このデータを用いて、今年の10aあたりの収穫量が例年より多かったといえるかを、仮説検定を用いて判断したい。

	10aあたりの 収穫量(kg)
①	500
②	520
③	490
④	530
⑤	480
⑥	490
⑦	460
⑧	500
⑨	510
⑩	520

- (1) この農家の昨年まで10年間の10aあたりの米の収穫量について、平均値と標準偏差を求めよ。
- (2) 検証したいこととは反対の事柄となる仮説を立て、判断に用いる棄却域を設定せよ。
- (3) この農家の今年の10aあたりの米の収穫量が525kgであった。今年の10aあたりの収穫量は、例年より多かったといえるか。仮説検定を用いて、判断せよ。

(解答)

- (1) 平均値は

$$\frac{1}{10}(500 + 520 + 490 + 530 + 480 + 490 + 460 + 500 + 510 + 520)$$

$$= 500 \text{ (kg)}$$

標準偏差は

$$\sqrt{\frac{1}{10}\{0^2 + 20^2 + (-10)^2 + 30^2 + (-20)^2 + (-10)^2 + (-40)^2 + 0^2 + 10^2 + 20^2\}}$$

$$= \sqrt{400} = 20 \text{ (kg)}$$

- (2) 検証したいことは

「今年の米の収穫量が例年より多い」かどうか

であるから

「今年の米の収穫量は例年と変わらない」

を仮説とする。

例年との比較は、過去10年間の収穫量の平均値で行うこととし、

「めったに起こらないこと」は

「米の収穫量が平均値から標準偏差の2倍以上離れた値となること」

とすると

$$500 + 20 \times 2 = 540$$

より、棄却域は **540kg 以上**

(3) 今年の 10a あたりの収穫量は 525kg であるから、

棄却域の「540kg 以上」に含まれない。

よって、立てた仮説は棄却されない。

したがって、このデータでは

今年の 10a あたりの収穫量が例年よりも多かったとはいえない。