

統計学基礎 練習問題

第19回 2変量データの記述(8)

2012年12月3日

問1 次のデータは、中四国地方9県の商業従業者数のデータである。このデータについて、以下の問に答えよ。ただし割り切れない数値がある場合は、四捨五入して小数点以下第2位まで求めること。

鳥取	島根	岡山	広島	山口	徳島	香川	愛媛	高知
5	6	16	26	12	6	9	12	7

データ出典: 経済産業省『商業統計調査』(単位: 万人) (2007年6月1日現在)

1. 算術平均 \bar{x} を求めよ。

2. メディアンを求めよ。

3. 分散 s^2 を求めよ。

4. 標準偏差 s を求めよ。

5. レンジを求めよ。

6. 次のデータは各県の漁業従業者数のデータである。この2変量の相関係数はいくつになるであろうか。次ページの選択肢の中から選べ。(商業従業者数を横軸に、漁業従業者数を縦軸に取った散布図を描き、参考にする事。)

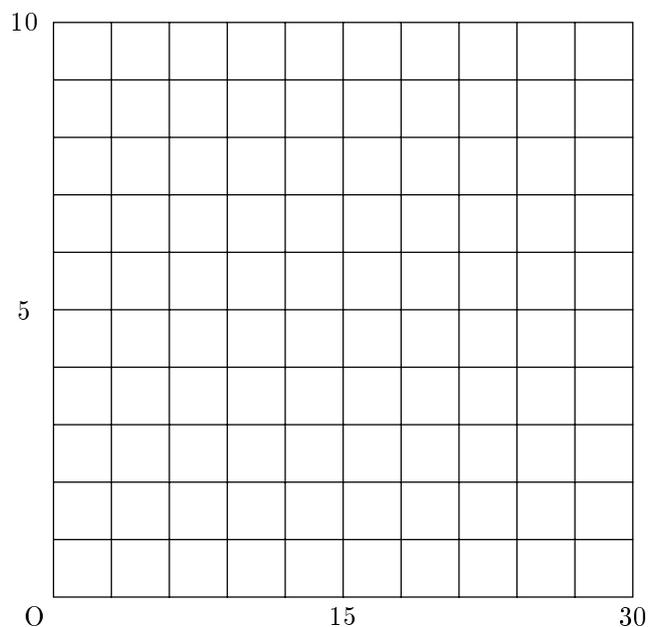
鳥取	島根	岡山	広島	山口	徳島	香川	愛媛	高知
2	4	2	5	7	3	3	9	5

データ出典: 農林水産省『漁業センサス』(単位: 千人) (2008年11月1日現在)

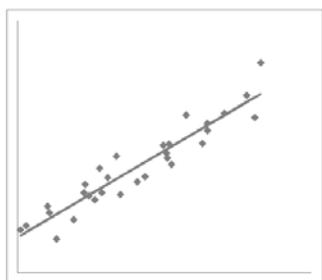
(1) 0.223

(2) -0.828

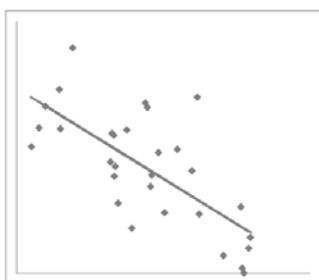
(3) 0.935



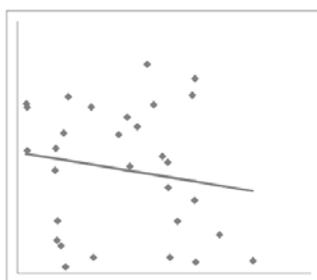
問 2 下の散布図を (a) ~ (e) について、 $y = a + bx$ という式をあてはめる。推定された回帰式と決定係数を (1) ~ (5) の中から選び、選択肢で答えよ。



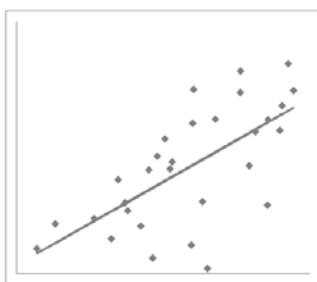
(a)



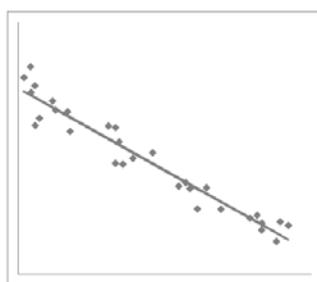
(b)



(c)



(d)



(e)

(a)	
(b)	
(c)	
(d)	
(e)	

(1) $y = 11.78 - 0.961x, R^2 = 0.502$

(2) $y = 1.97 + 0.803x, R^2 = 0.889$

(3) $y = 10.34 - 0.913x, R^2 = 0.937$

(4) $y = 5.77 - 0.192x, R^2 = 0.027$

(5) $y = 0.46 + 0.927x, R^2 = 0.441$

問3 次の2つのデータはそれぞれ大きさの順に並べてある。AとBで等しいものはどれか。下の①～④のうちから最も適切なものを一つ選べ。

A: 12, 14, 17, 23, 25, 34, 38, 39, 42, 52, 56, 58, 59, 64

B: 27, 29, 32, 38, 40, 49, 53, 54, 57, 67, 71, 73, 74, 79

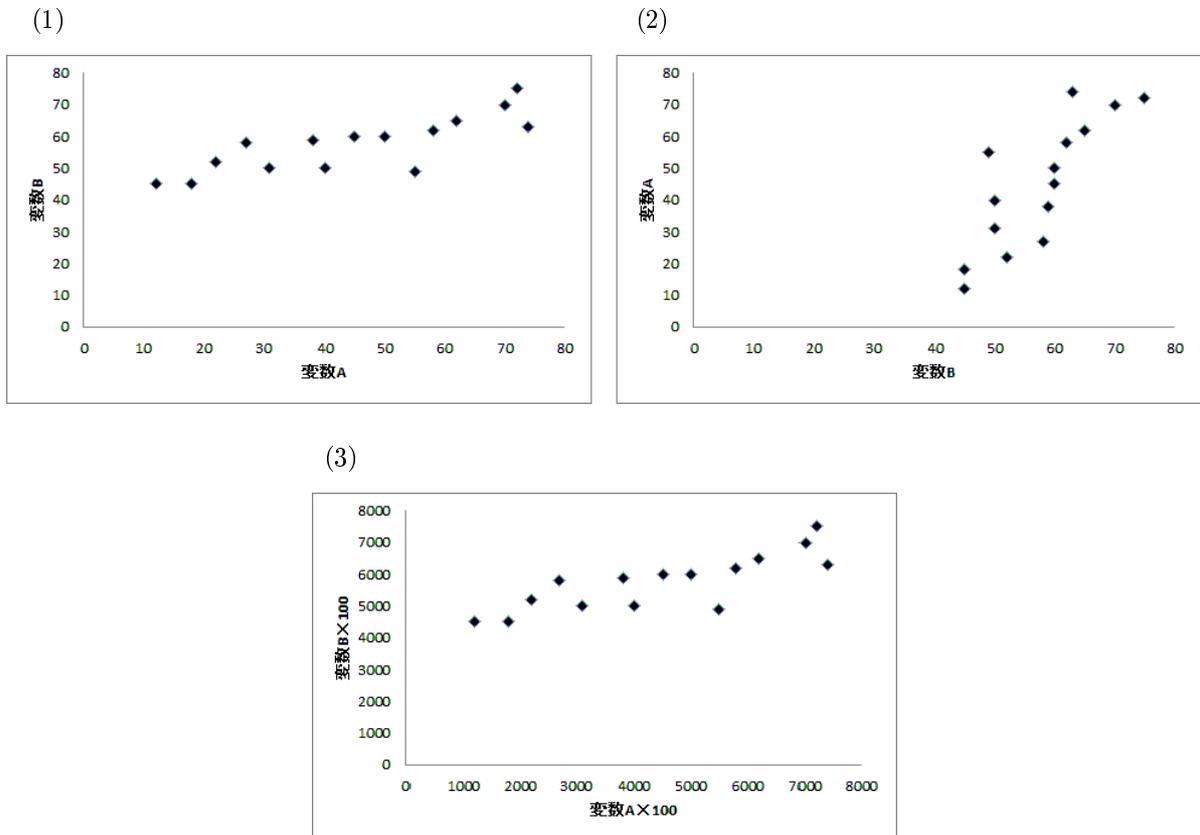
- ① 平均値 ② 中央値 ③ 分散 ④ すべて異なっている

問4 あるクラスで読んだ本の冊数を調査したところ、平均2冊、標準偏差1.2冊だった。その後、入力ミスが見つかり、各人が読んだ本の冊数は、本当はそれぞれ10倍の数値であることがわかった(例: 2冊と入力された人は、本当は20冊読んでいた)。このとき、本当の冊数での平均値と標準偏差の正しい組合せを、次の①～④のうちから一つ選べ。

- ① 平均値: 2(冊), 標準偏差: 1.2(冊)
 ② 平均値: 2(冊), 標準偏差: 12(冊)
 ③ 平均値: 20(冊), 標準偏差: 1.2(冊)
 ④ 平均値: 20(冊), 標準偏差: 12(冊)

問5 2つの変数A,Bについての観測値 $(a_1, b_1), \dots, (a_n, b_n)$ が求められたとき、以下の3つの散布図を次の手順で作成した。

- (1) は横軸に a , 縦軸に b を取った図
 (2) は縦軸に a , 横軸に b を取った図
 (3) は横軸に $100 \times a$, 縦軸に $100 \times b$ を取った図



このとき上の散布図の中で相関係数が最も大きいものはどれか。次の①～④のうちから最も適切なものを一つ選べ。

- ① (1) の散布図 ② (2) の散布図 ③ (3) の散布図 ④ (1),(2),(3) の相関係数は同じになる