ひかり統計塾

統計検定２級
公式問題集(CBT対応版)の解説
カテゴリー1（p24-41）
1変数記述統計の分野

統計検定2級　CBT問題集　PART.2 目次

[C1]
[CATEGORY.1]
1変数記述統計の分野

(p24.1)[C1]問1.相対度数の計算

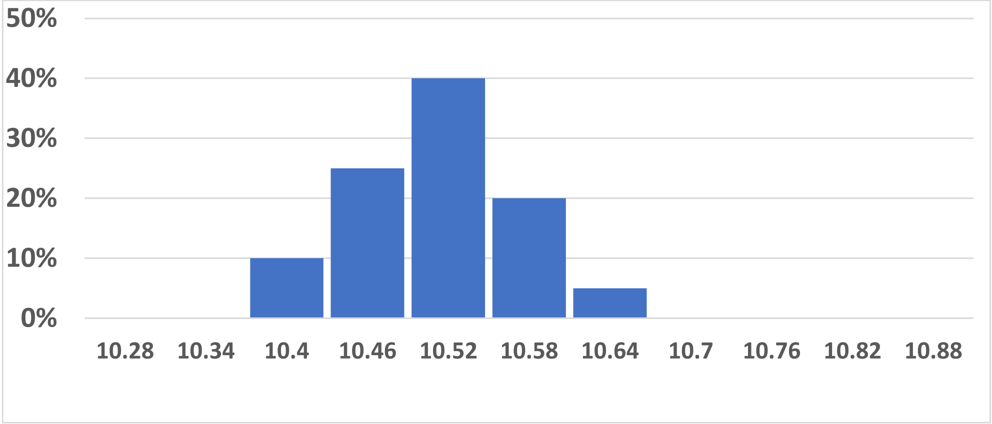
(Aランク)

類題

(相対)度数分布表

	以上		未満	合計		
階級No	区間			中心値	度数	相対度数(%)
1	10.250	～	10.310	10.280	0	0
2	10.310	～	10.370	10.340	0	0
3	10.370	～	10.430	10.400	20	10
4	10.430	～	10.490	10.460	50	25
5	10.490	～	10.550	10.520	80	40
6	10.550	～	10.610	10.580	40	20
7	10.610	～	10.670	10.640	10	5
8	10.670	～	10.730	10.700	0	0
				合計	200	100

相対度数分布



(ポイント)
相対度数の総和は100(%)

$$10 + 25 + ? + 20 + 5 = 100$$
$$35 + ? + 25 = 100$$

$$? = 100 - 60 = \mathbf{40}$$

CBT問題集:
p24,問1 (ア)、(イ)の値は何でしょう？

・1952年・・・97.8+2.2=100となっている

(ア)=100-(85.1+2.1)=100-87.2=12.8
(イ)=100-(76.6+17.0+2.1)=100-95.7=4.3
⇒答:⑤

済

(p26.1)[C1]問2.中央値を含む階級

(Aランク)

類題 **度数分布表**

Q. 中央値が含まれている階級番号は？

(公式)
中央値が含まれている階級
=「累積相対度数=50%」を
含む階級

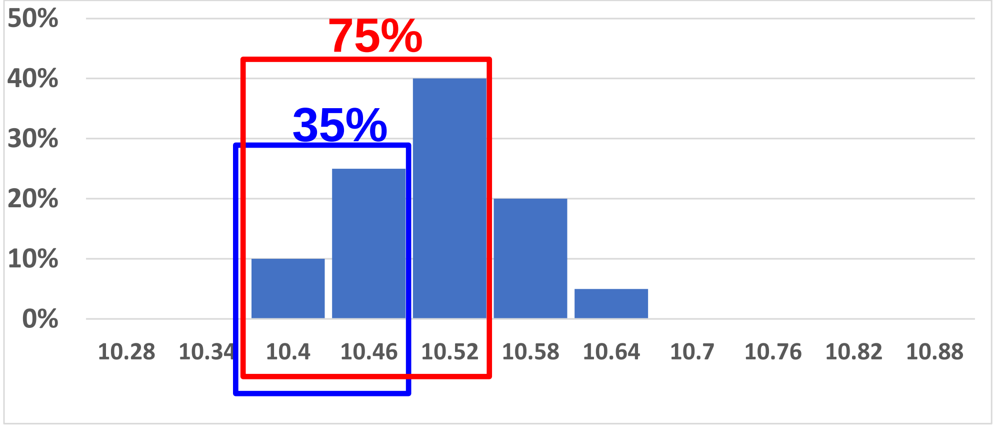
	以上		未満	合計			
階級No	区間			中心値	度数	相対度数(%)	累積相対度数(%)
1	10.250	～	10.310	10.280	0	0	0
2	10.310	～	10.370	10.340	0	0	0
3	10.370	～	10.430	10.400	20	10	10
4	10.430	～	10.490	10.460	50	25	35
5	10.490	～	10.550	10.520	80	40	75
6	10.550	～	10.610	10.580	40	20	95
7	10.610	～	10.670	10.640	10	5	100
8	10.670	～	10.730	10.700	0	0	100
				合計	200	100	

←50%を超えていない

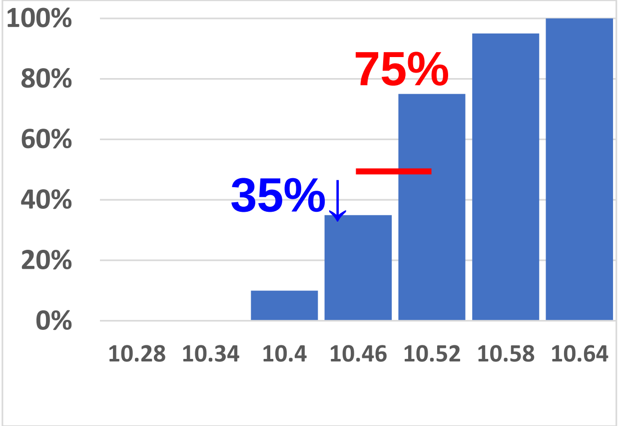
←50%を超えている
（ここに中央値が含まれている)

(答)階級No.5

相対度数分布



累積相対度数分布



(p26.2)[C1]問2.中央値を含む階級

(Aランク)

13.2+7.2=20.4

	相対度数(%)	累積相対度数(%)	
	13.2	13.2	
	7.2	20.4	
	7.0	27.4	
	6.1	33.5	
(E)	5.6	39.1	
(F)	5.5	44.6	
(G)	4.5	49.1	←(G) 50%を越えていない
(H)	4.2	53.3	←(H) 50%を越えている（ここに中央値が含まれている)
(I)	3.3	・・・	
(J)	3.2		
(K)	6.0		
・・・	・・・		

階級(H)　⇒(答) ①

(ポイント)
中央値が含まれている階級
=「累積相対度数=50%」を含む階級

済

(p28.1)[C1]問3. 箱ひげ図と度数分布

(Aランク)

箱ひげ図の復習

①データを大小の順で並べる

120
82
68
60
56
52
52
50
46
45
38
30
28
25
-20

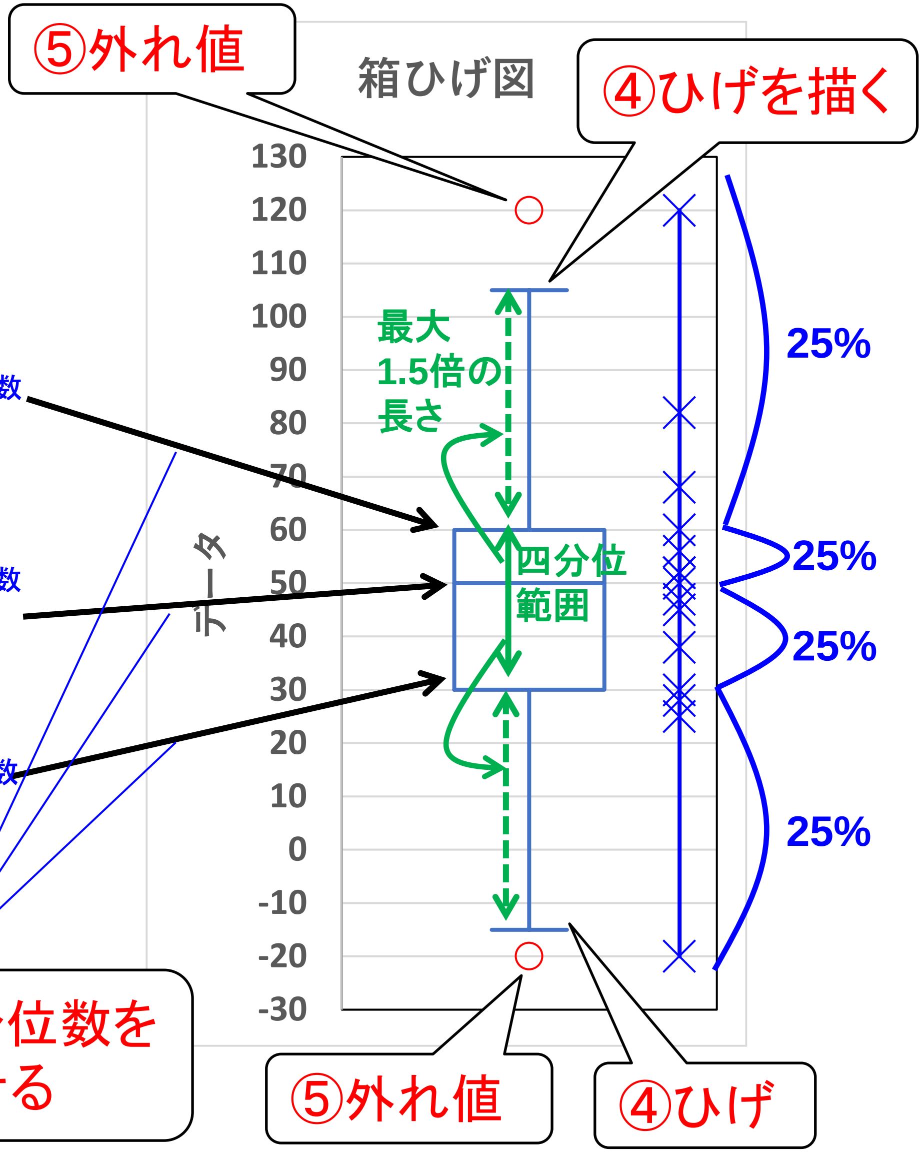
②4つに分ける

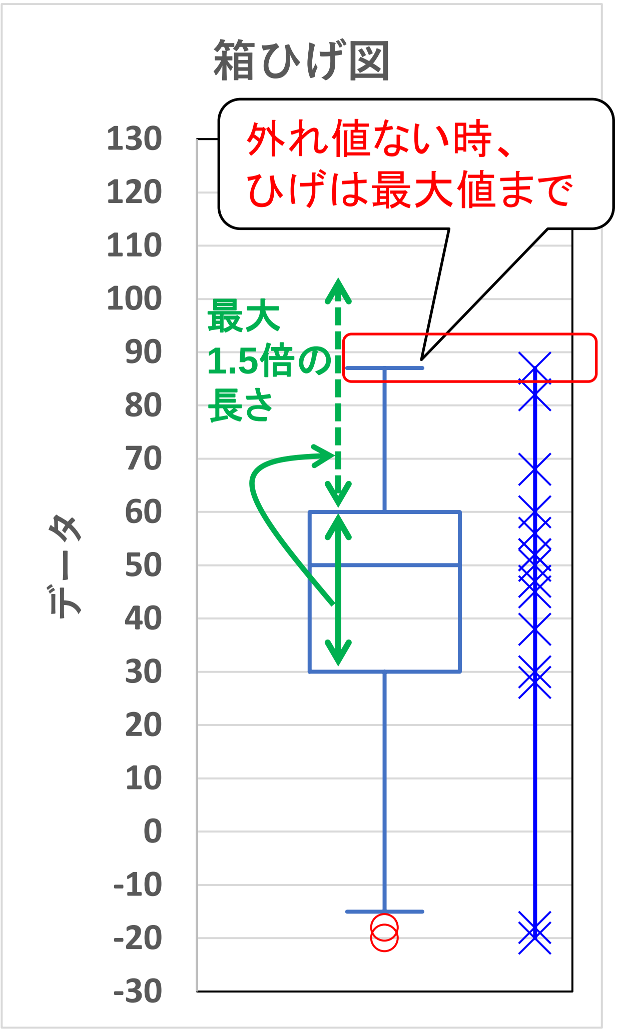
25%
25%
25%
25%

5数要約

番号	データ	
15	120	(e)最大値
14	82	
13	68	
12	60	(d)第３四分位数
11	56	
10	52	
9	52	
8	50	(c)第２四分位数（中央値）
7	46	
6	45	
5	38	
4	30	(b)第1四分位数
3	28	
2	25	
1	-20	(a)最小値

③3つの四分位数を箱に対応させる





(p28.2)[C1]問3. 箱ひげ図と度数分布

(Aランク)

問題: p29の度数分布表で、東京の度数はどれですか？

ヒント:この様な問題では、まず、「端部」(最大値、最小値、外れ値)に着目してください

p28. 東京の箱ひげ図より

・2℃以上4℃未満に最小値がある ⇒ 満たすのは(A),(B),(C)

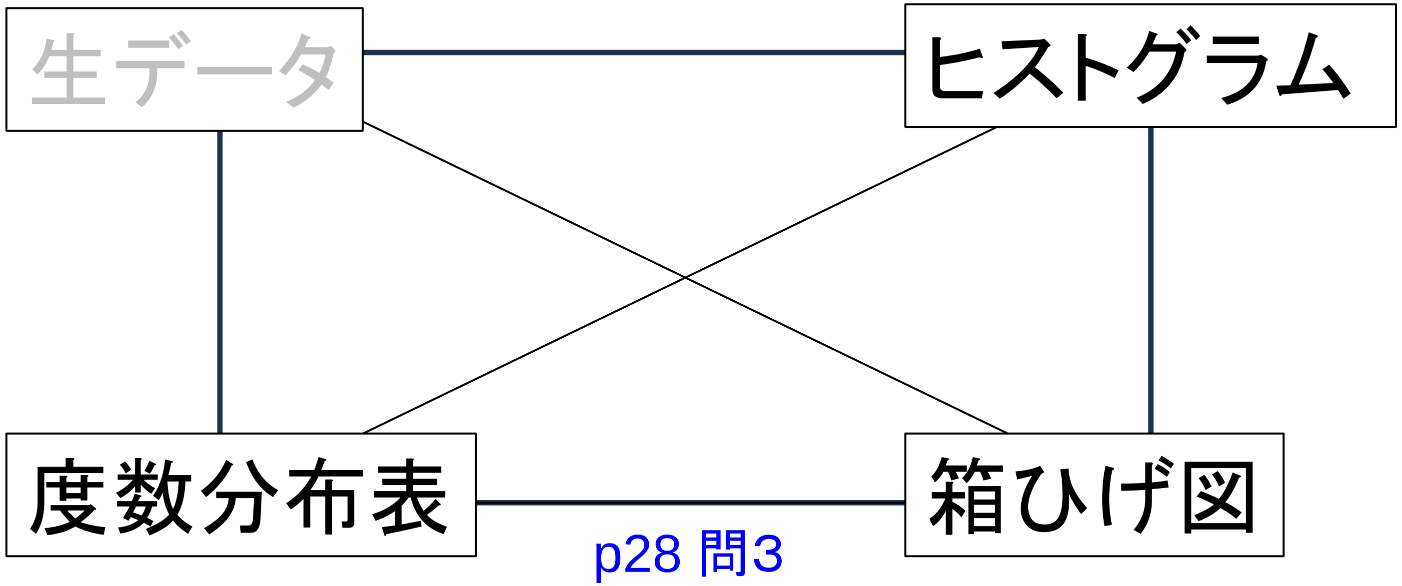
・16℃以上18℃未満に外れ値が2個以上ある (⇒最大値) ⇒ (A),(B),(C)のうち、満たすのは(A)

（重なっている可能性がある）

(答)①

済

(p28.3)[C1]問3. 箱ひげ図と度数分布



対応付けの問題:
よく出ます

ポイント: まずは「端部」に
注目してください

(p30.1)[C1]問4.幹葉図(みきはず)の読み取り

(Aランク)

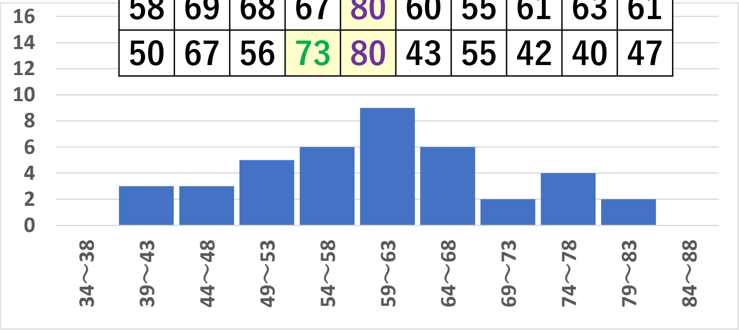
幹葉図(みきはず) (かんようず) の説明:

例: A塾でのテストの点数

A塾

51	54	54	65	47	74	59	63	74	74
49	75	49	52	68	61	64	48	62	62
58	69	68	67	80	60	55	61	63	61
50	67	56	73	80	43	55	42	40	47

人数



点数ごとの
人数を整理

点数	人数
40	1
42	1
43	1
47	2
48	1
49	2
50	1
51	1
52	1
54	2
55	2
56	1
58	1
59	1
60	1
61	3
62	2
63	2
64	1
65	1
67	2
68	2
69	1
73	1
74	3
75	1
80	2
合計	40

幹 十の位	葉 一の位
4	02377899
5	0124455689
6	011122334577889
7	34445
8	00

73点が1人、74点が3人、75点が1人
80点が2人いるという意味です。

(p30.2)[C1]問4.幹葉図の読み取り

(Aランク)

①最高得点:90点。92点ではないので、×

②最低得点は40点。56点ではないので、×

③60点未満の人は、7人なので、〇

④学生数=25名。20%=5名。

成績は、上から順に、90,82,80,78,78, 76,76,...

上位20%=5名の最低点は、78点であり、76点ではないので、×

⑤ 70点は一人だけ。58点は5人、60点は3人

70点が最頻値となることはない⇒×

(答) ③

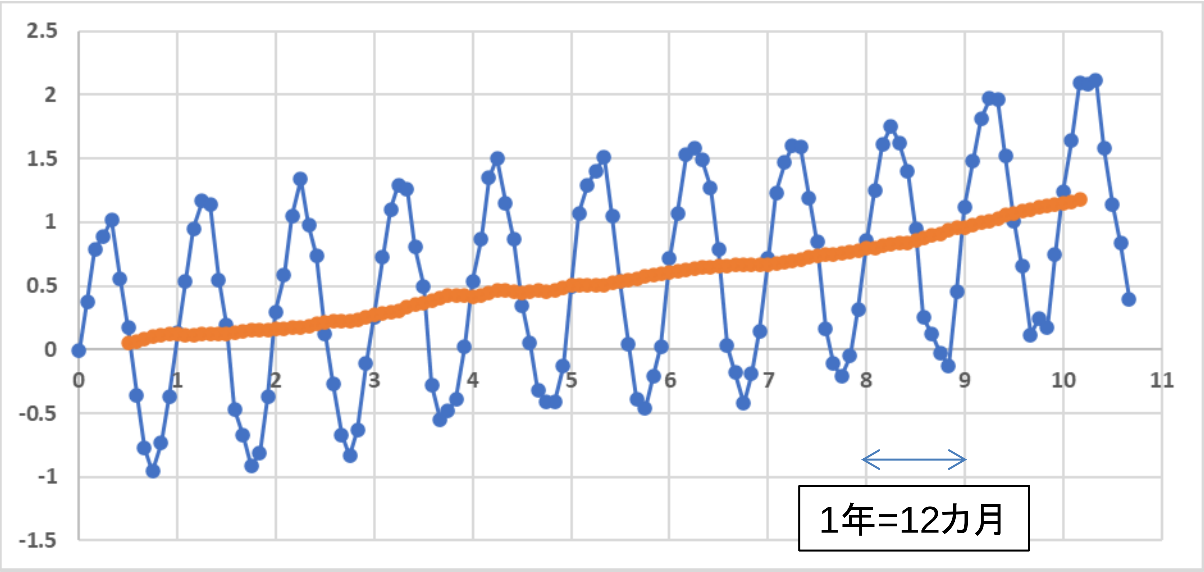
済

(p31.1)[C1]問5.時系列の変動の性質

時系列データの変動分解

- ・傾向変動
- ・季節変動　(1年周期)
- ・不規則変動

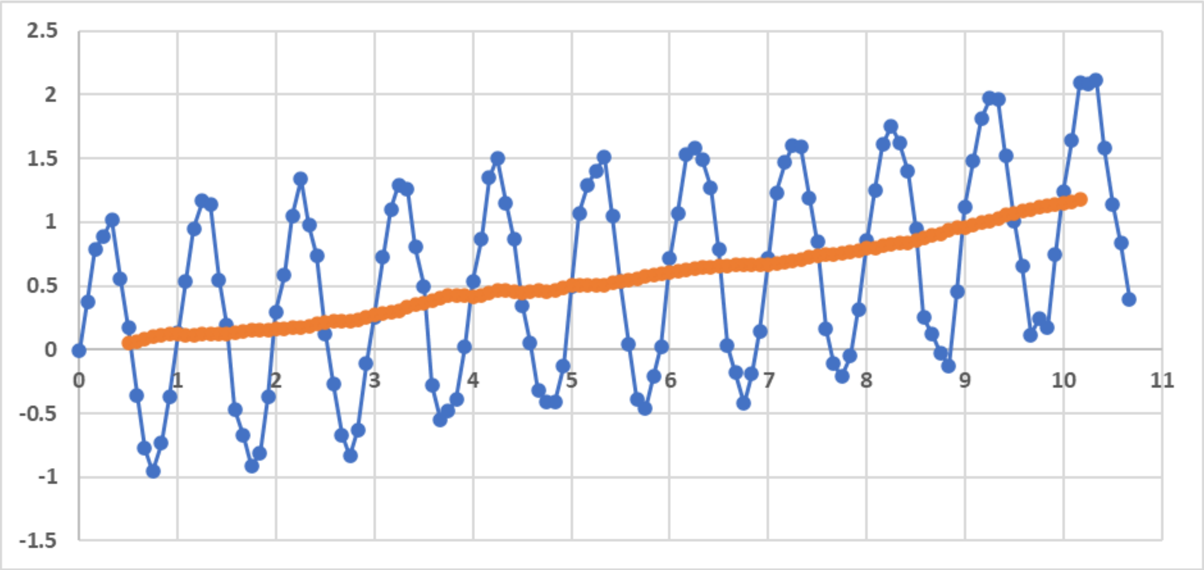
ある農作物の収穫量（仮想的なデータ）
（基準時点での値: 0 と設定）



(p31.2)[C1]問5.時系列の変動の性質

(Aランク)

I: 傾向変動は直線でない場合もある、×



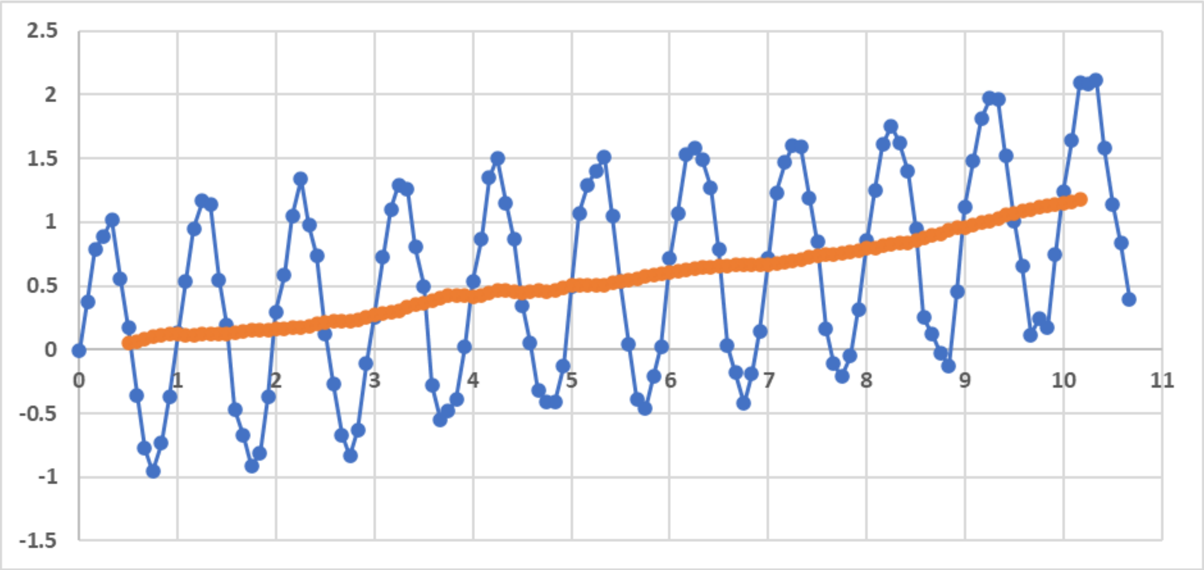
II:　正しい　○

III: 不規則変動には、予測が困難な偶然変動も含まれるので、　×

IIのみが正しい⇒(答)②

済

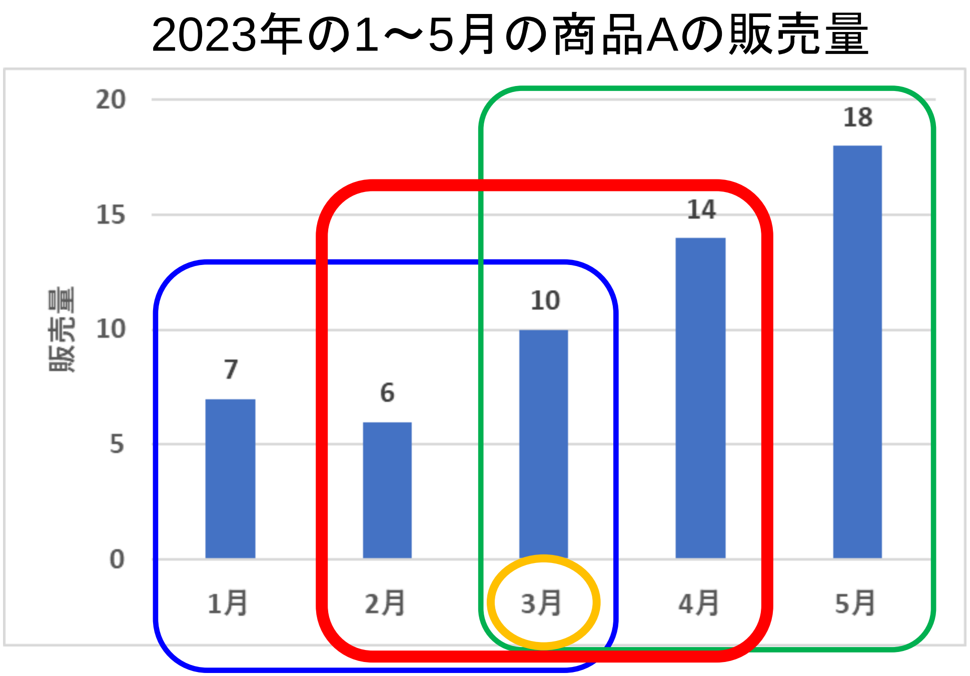
(p31.3a)[C1](参考)移動平均



- ・傾向変動　を抽出したい時、
- ・季節変動　（1年周期）を除去するといい　⇒周期(1年=12カ月)で平均化すればいい
- ・不規則変動

(p31.3b)[C1](参考)移動平均

3項移動平均：
2018年11月問3に
出題例があります



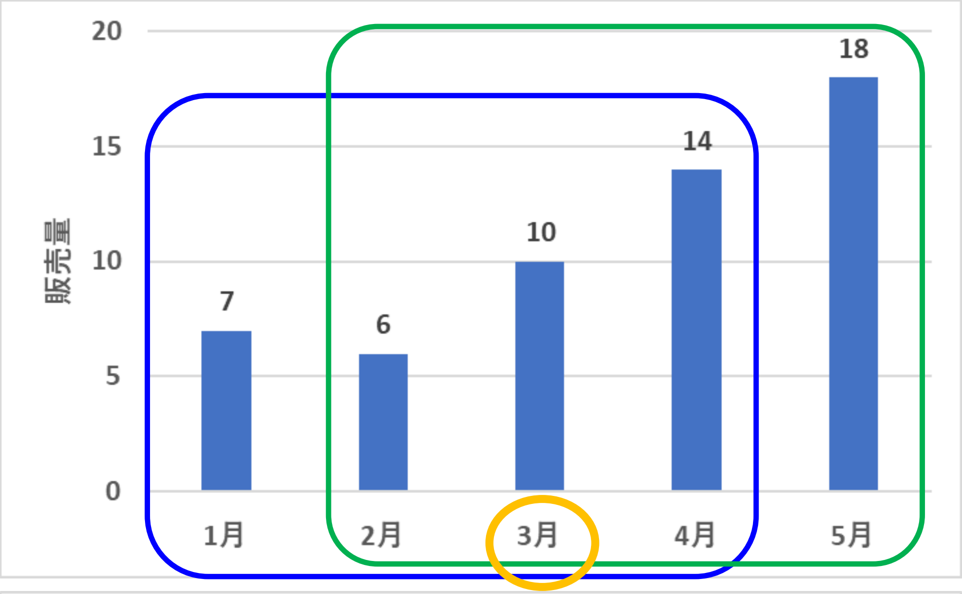
Q1: 2023年3月の
3項移動平均（＝3か月間での平均）は
どう計算すべき？

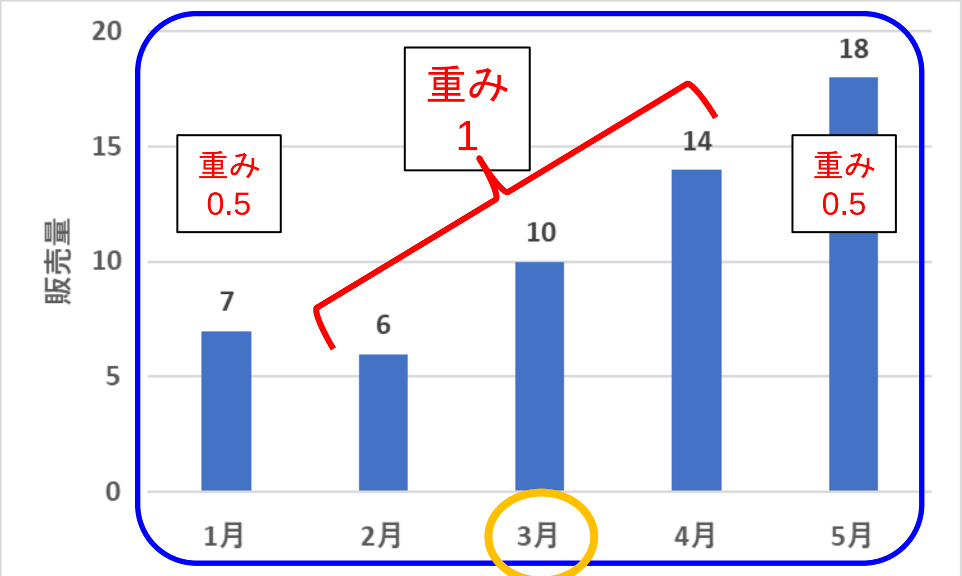
A1: 3月＋前後の2月、4月の販売量を使い、
2月～4月の3か月間で平均をとるといいですね。

$$\frac{1}{3}(6+10+14)=10$$

(p31.3c)[C1](参考)移動平均

2023年の1～5月の商品Aの販売量





Q2: 2023年3月の
4項移動平均（＝4か月間での平均）は？

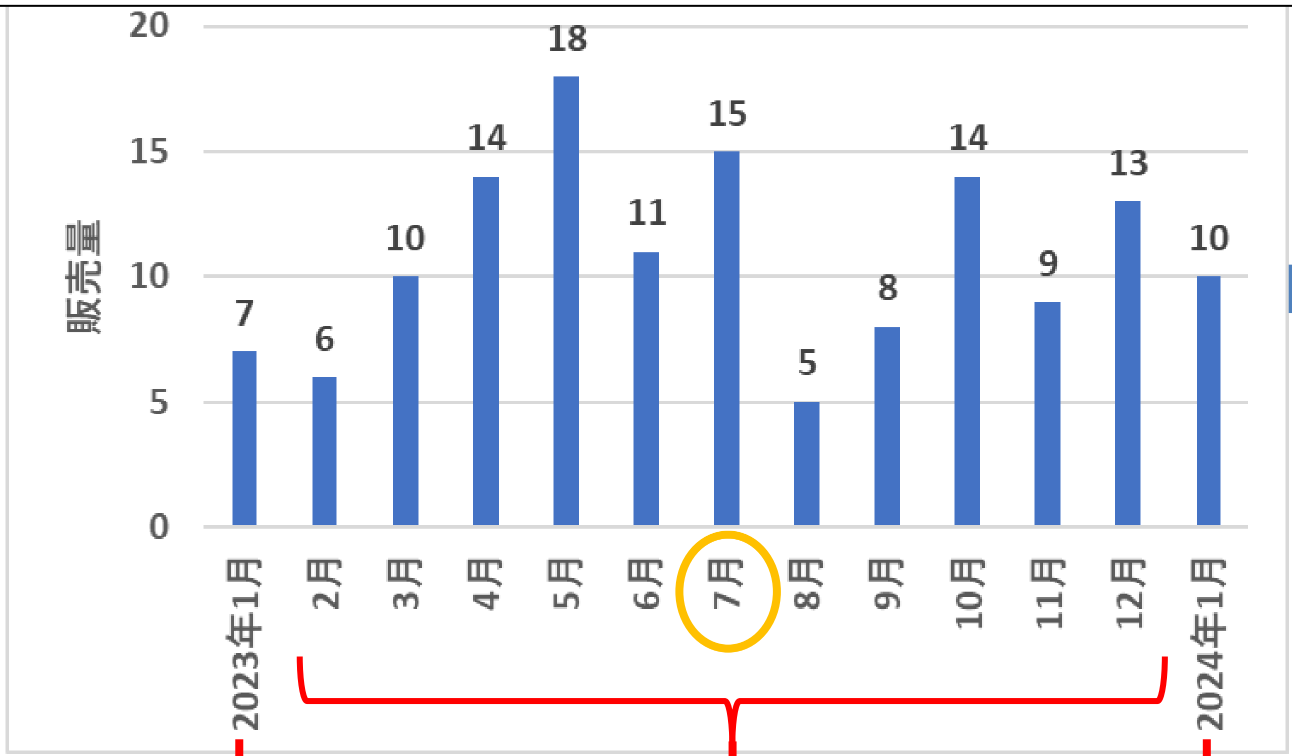
1月～4月の4か月間で平均をとるといい？
2月～5月の4か月間で平均をとるといい？

1月～5月の5か月間の値を使い、
2～4月分は、重み1で、
両端の1月,5月分は、重み0.5にして、計算するといい

$$\frac{1}{4}\left(\frac{7}{2}+6+10+14+\frac{18}{2}\right)=10.625$$

(p31.3d)[C1](参考)移動平均

7月における、12項移動平均:12カ月(1年)での移動平均

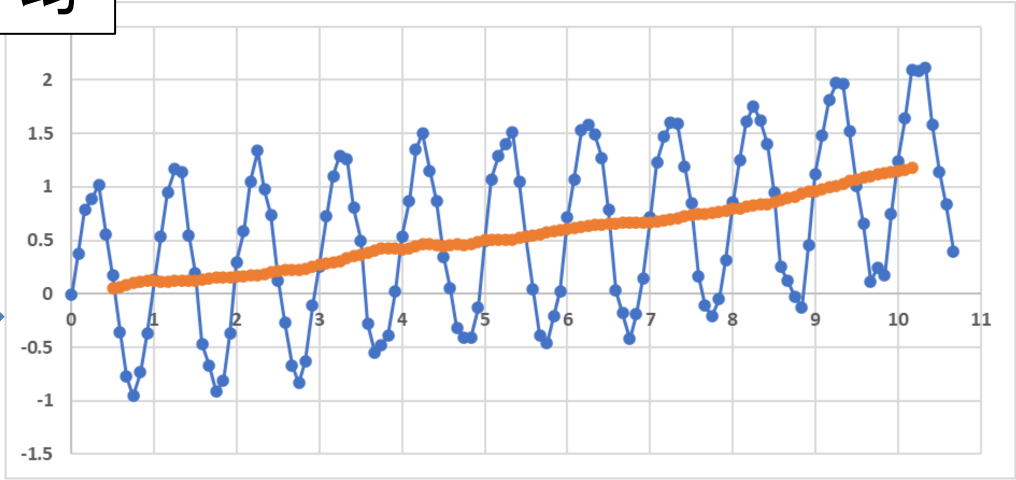


重み: 0.5

重み: 1
(11か月分)

重み: 0.5

$$\frac{1}{12}\left(\frac{7}{2}+6+10+\cdots+9+13+\frac{10}{2}\right)=10.958$$



まとめ

<u>奇数項の移動平均</u>

$$\frac{1}{3}(6+10+14)=10$$

<u>偶数項の移動平均</u>

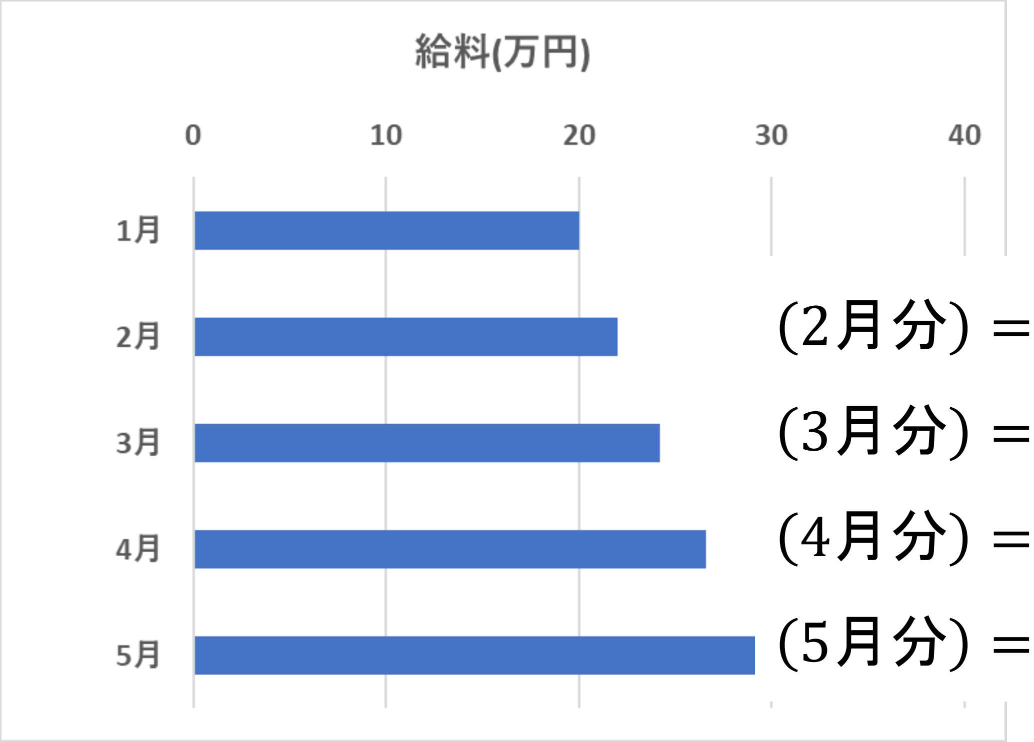
$$\frac{1}{4}\left(\frac{7}{2}+6+10+14+\frac{18}{2}\right)=10.625$$

(p32.1)[C1]問6.平均変化率の計算式

変化率

あるAI系の仕事の給料は、
1月は20万円、2月は1月の1.1倍、3月は2月の1.1倍、4月も3月の1.1倍・・・とします。

	給料(万円)	
1月	20	
		×1.1
2月	22	
		×1.1
3月	24.2	
		×1.1
4月	26.62	
		×1.1
5月	29.282	



$(2\text{月分}) = (1\text{月分}) \times 1.1 = (1\text{月分}) \times 1.1$

$(3\text{月分}) = (2\text{月分}) \times 1.1 = (1\text{月分}) \times 1.1^2$

$(4\text{月分}) = (3\text{月分}) \times 1.1 = (1\text{月分}) \times 1.1^3$

$(5\text{月分}) = (4\text{月分}) \times 1.1 = (1\text{月分}) \times 1.1^4$

$1.1\text{倍} = (1 + 0.1)\text{倍} = \left(1 + \frac{10}{100}\right)\text{倍}$ ⇒「変化率は10%」と表現します

(p32.2)[C1]問6.平均変化率の計算式

(ABランク)

同じ比で変化したと仮定します

		比			比	
1月	102.6					
	↓	1.0127	倍	⇒	$\boldsymbol{1+r'}$	倍
2月	103.9					
	↓	1.0029	倍	⇒	$\boldsymbol{1+r'}$	倍
3月	104.2					
	↓	1.0134	倍	⇒	$\boldsymbol{1+r'}$	倍
4月	105.6					

$\times(\boldsymbol{1+r'})^{3}$

但し、

平均変化率:$\boldsymbol{r}$(%)

$$\boldsymbol{r=100r'}$$

$$\boldsymbol{r'=\frac{r}{100}}$$

$$105.6=102.6\times(\boldsymbol{1+r'})^{3}$$

$$(\boldsymbol{1+r'})^{3}=\frac{105.6}{102.6}$$

$$1+r'=\left(\frac{105.6}{102.6}\right)^{\frac{1}{3}}$$

$$r'=\left(\frac{105.6}{102.6}\right)^{\frac{1}{3}}-1$$

(公式) ある月の●●値が, x_0で、
nか月(n年)後の●●値がx_nの時、
平均変化率をr(%)とすると、

$$x_n=x_0\times\left(\boldsymbol{1+\frac{r}{100}}\right)^{\boldsymbol{n}} \Rightarrow \boldsymbol{r=100\left\{\left(\frac{x_n}{x_0}\right)^{\frac{1}{n}}-1\right\}}$$

%で表現すると、$r=100r'$

$$r=100r'=100\left\{\left(\frac{105.6}{102.6}\right)^{\frac{1}{3}}-1\right\}$$

(答)④

済

(p32.3)[C1]問6.平均変化率の計算式(類題)

同じ比で変化したと仮定します

		比			比	
1月	102.6					
				⇒	$\boldsymbol{1+r'}$	倍
2月						
				⇒	$\boldsymbol{1+r'}$	倍
3月						
				⇒	$\boldsymbol{1+r'}$	倍
4月						
				⇒	$\boldsymbol{1+r'}$	倍
5月	105.6					

$\times(\boldsymbol{1+r'})^{\boldsymbol{4}}$

さきほどは、「1月⇒4月」で平均変化率を求めました。
もしも、「1月⇒5月」の場合には、どうなりますか？
（5月での値は、105.6と仮定します。）

$$105.6 = 102.6\times(\boldsymbol{1+r'})^{\boldsymbol{4}}$$

$$(\boldsymbol{1+r'})^{\boldsymbol{4}} = \frac{105.6}{102.6}$$

$$1+r' = \left(\frac{105.6}{102.6}\right)^{\frac{1}{4}}$$

$$r' = \left(\frac{105.6}{102.6}\right)^{\frac{1}{4}} - 1$$

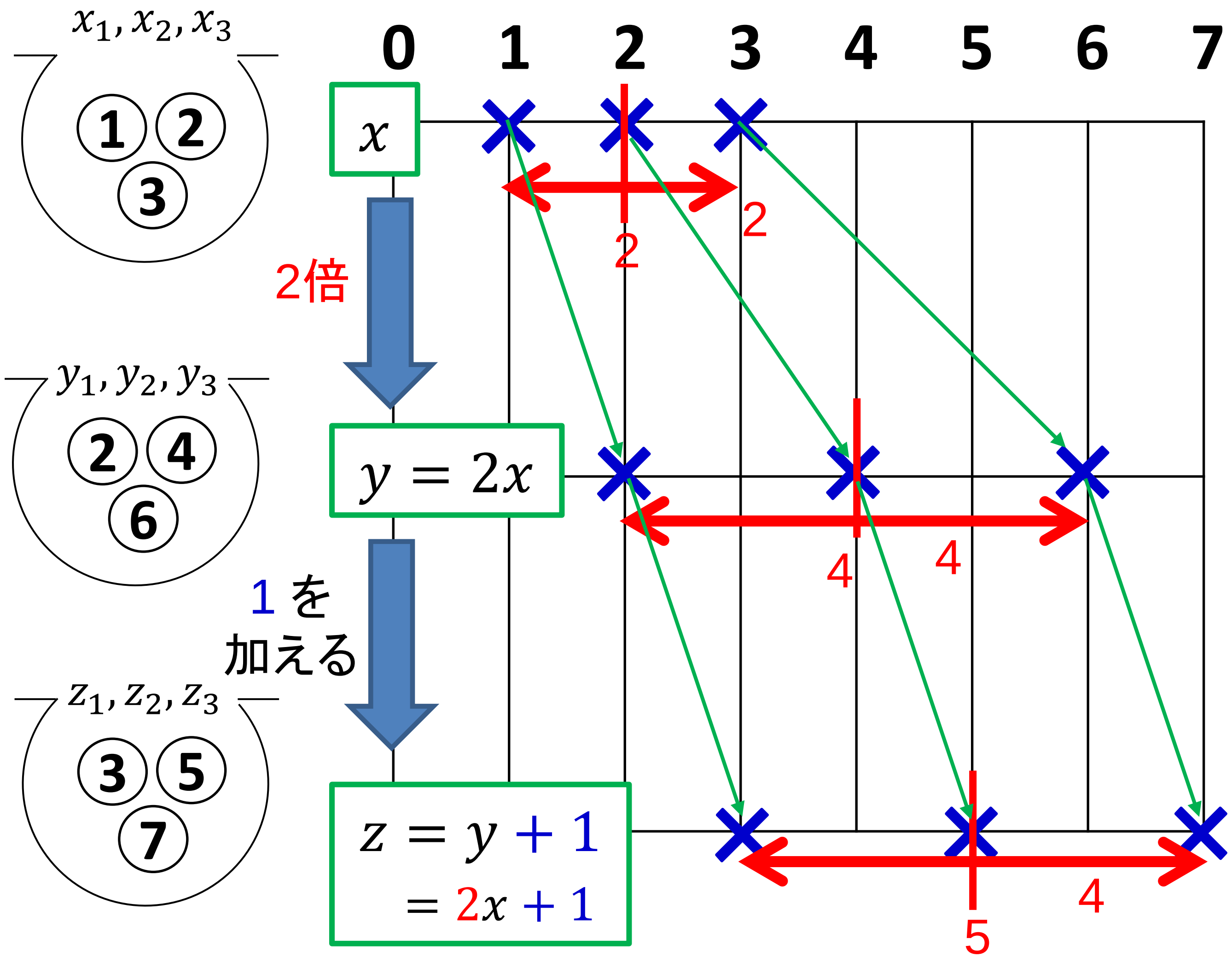
と求まります。これは電卓で実際に計算できますか？

$$\left(\frac{105.6}{102.6}\right)^{\frac{1}{4}} = \sqrt{\sqrt{\frac{105.6}{102.6}}} = \sqrt{\sqrt{1.02923}} = \sqrt{1.0145} = 1.00723$$

$r' = 0.00723 \;\Rightarrow r = 0.723\%$　実計算可能です

$$(x)^{\frac{1}{2}} = \sqrt{x},\qquad (x)^{\frac{1}{4}} = \sqrt{\sqrt{x}},\qquad (x)^{\frac{1}{8}} = \sqrt{\sqrt{\sqrt{x}}}, \ldots$$

(p34.1)[C1]問7.線形変換による平均・標準偏差



目安=最大-最小

中心 (平均)	広がり (標準偏差)	広がりの2乗 (分散)
$\bar{x} = 2$	$s_x = \sqrt{2/3}$	$s_x^2 = 2/3$
2倍になる	2倍になる	2^2倍になる
$\bar{y} = 2\bar{x}$	$s_y = 2s_x$	$s_y^2 = 2^2\, s_x^2$
1 だけ大きくなる	変わらない	変わらない
$\bar{z} = \bar{y} + 1$ $= 2\bar{x} + 1$	$s_z = s_y$ $= 2s_x$	$s_z^2 = s_y^2$ $= 2^2\, s_x^2$

(p34.2)[C1]問7.線形変換による平均・標準偏差 (ABランク)

<公式> 超重要！

n個のデータ: $x_1, x_2, \dots, x_n$

(平均: $\bar{x}$) (分散:s_x^2, 標準偏差: s_x)

に対して、

$y_i = ax_i + b$ (a, b: 定数)

($i = 1,2,\dots,n$)の

変換を行った時、

その平均は $\bar{y} = a\bar{x} + b$

その分散は $s_y^2 = a^2 s_x^2$

標準偏差は $s_y = |a| s_x$

となります。

・摂氏での平均: $\bar{C} = 2.4$, 標準偏差: $s_C = 7.0$

・$F_i = 1.8C_i + 32$ の変換をした時、

Q. 華氏での平均: $\bar{F} =$ ？, 標準偏差: $s_F =$ ？

$x \Rightarrow C$(摂氏)、$y \Rightarrow F$(華氏)と置き換えて考えます。

$$F_i = aC_i + b \ (a = 1.8, b = 32)$$

$\bar{y} = a\bar{x} + b$

$\Rightarrow$ 平均：$\bar{F} = a\bar{C} + b = 1.8 \times 2.4 + 32 = 36.32$

$s_y = |a| s_x$

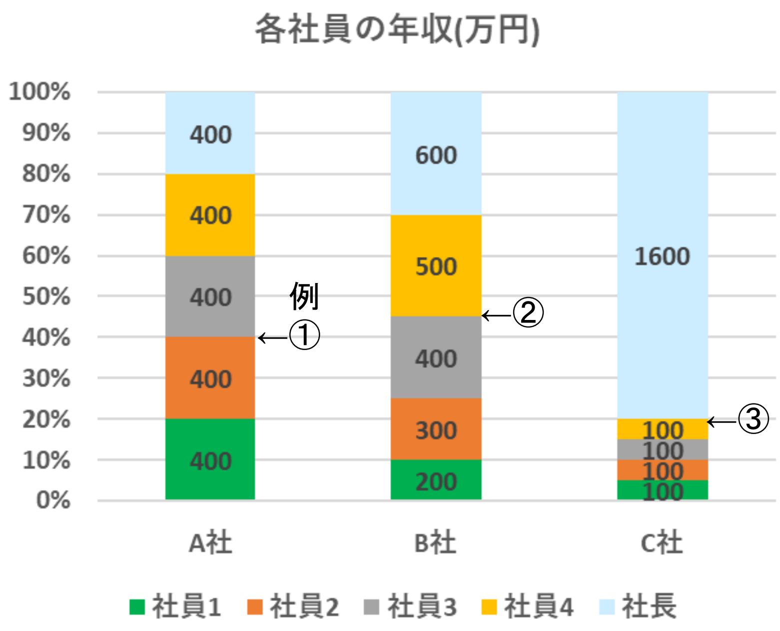
$\Rightarrow$ 標準偏差：$s_F = |a| s_C = 1.8 \times 7.0 = 12.6$

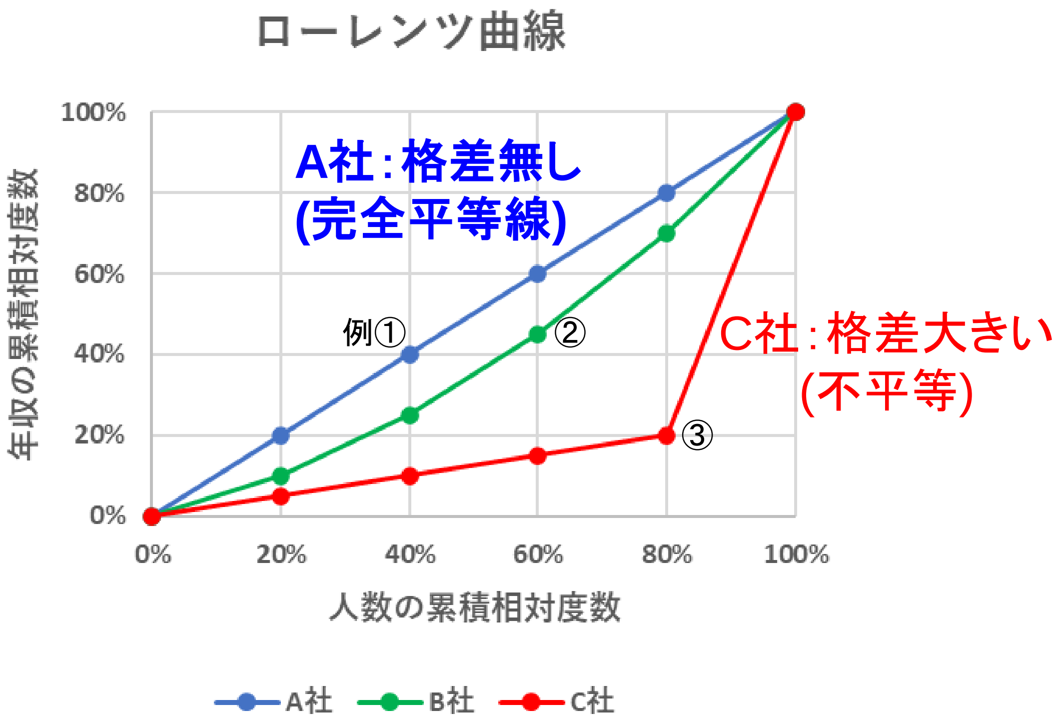
(答)④

済

(p36.1)[C1]問8.ローレンツ曲線・ジニ係数の説明

あなたは、新卒で就職活動中だとします。
就職先の候補として、3社があり、それぞれ5名(社員4名と社長)が働いています。
どの会社に入りたいですか？

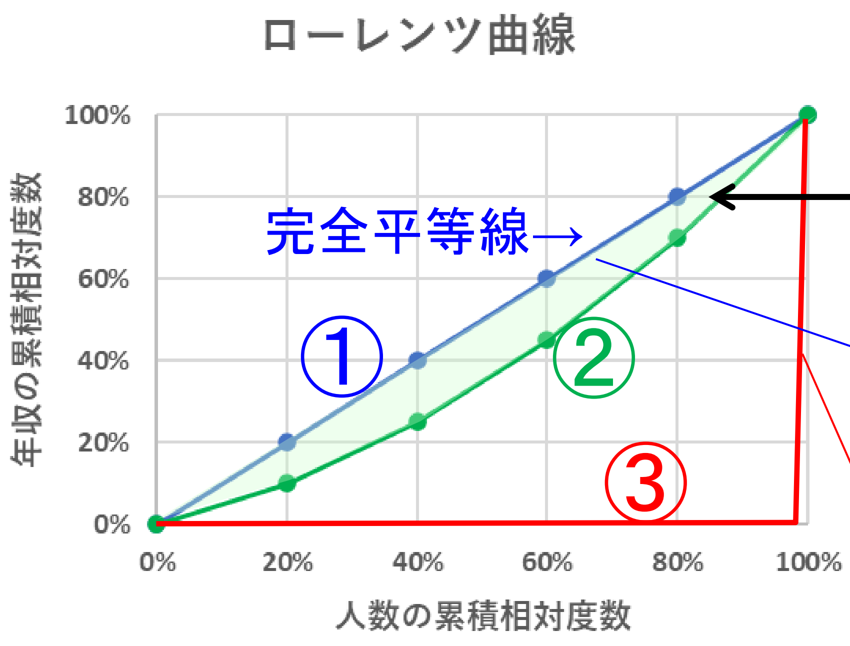




ローレンツ曲線は、格差の度合い（平等・不平等の度合い）を表現します。

総人数が(例)500人の場合、年収の低い順に100人、100人、...の様に、(例)5グループに分けて
総年収に対する各グループの寄与を調べ、図示します。

(p36.2)[C1]問8.ローレンツ曲線・ジニ係数の説明



この領域の面積の2倍がジニ係数

ローレンツ曲線が完全平等線①に一致の時、
ジニ係数≒0となる。平等。(例:全員の年収は同じ)

ジニ係数が大きくなると、
不平等の度合いが大きくなる

ローレンツ曲線が③の場合：
ジニ係数≒1となる。不平等。
(例:社員が多い会社で、社長以外はほぼ無給、
社長のみ年収100億円)

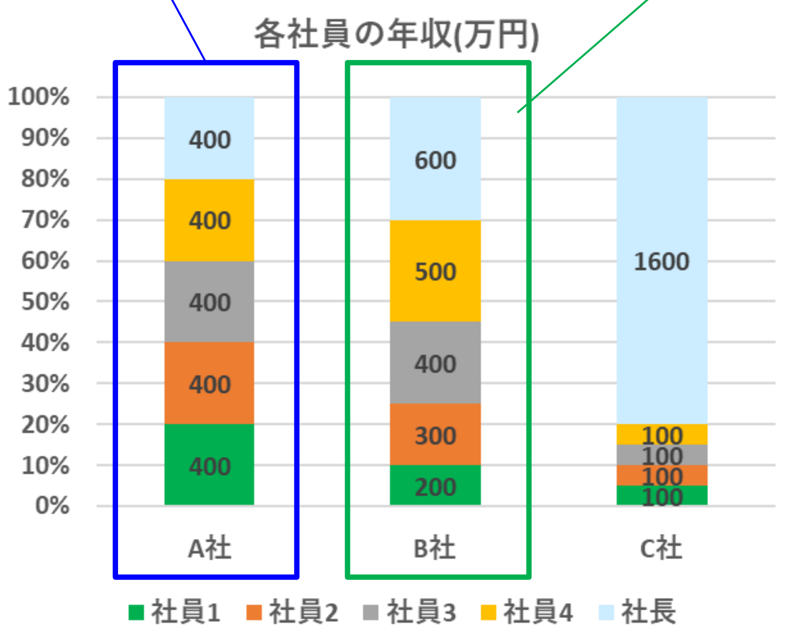
過去問
2018年6月回　問3
図からジニ係数を求める問題が
出ています

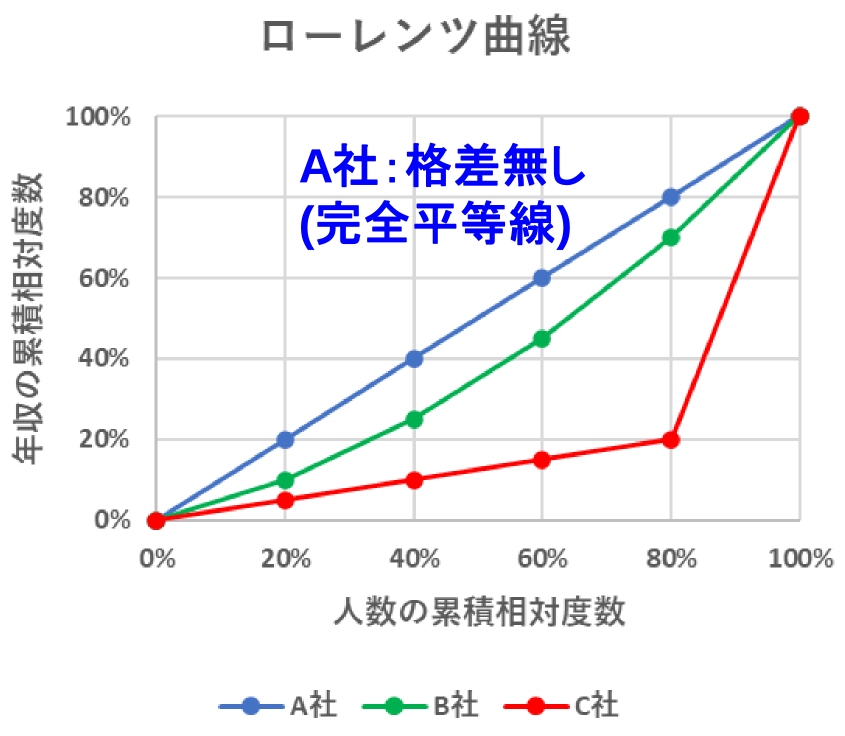
(p36.3)[C1]問8.ローレンツ曲線・ジニ係数の説明　(ABランク)

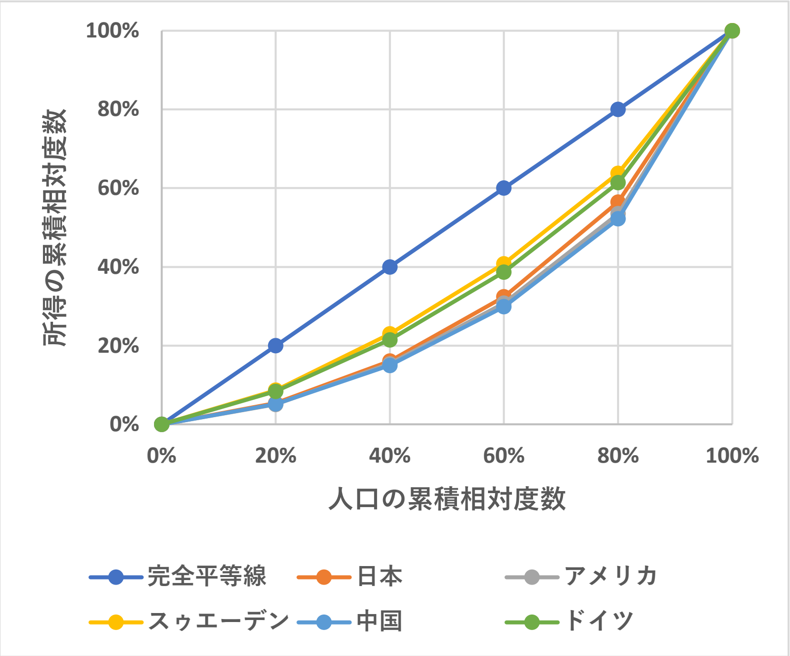
I:　正しいですか？

I:　各国の人口を5等分して、所得を調べている。完全平等線ならば、各階級の寄与が20%ずつになっているはず。表を見ると、そうはなっていない。その結果、ローレンツ曲線は完全平等線の下に弧を描く。 ⇒ Iは正しい

完全平等線　　問8の問題例に近い



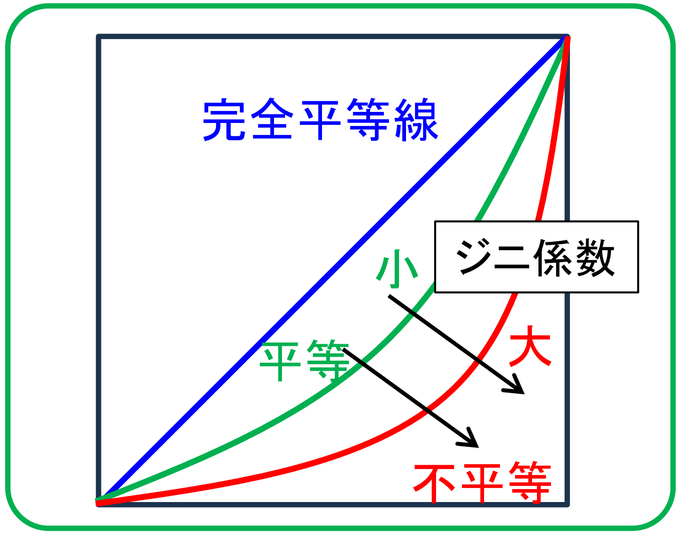




(p36.4)[C1]問8.ローレンツ曲線・ジニ係数の説明

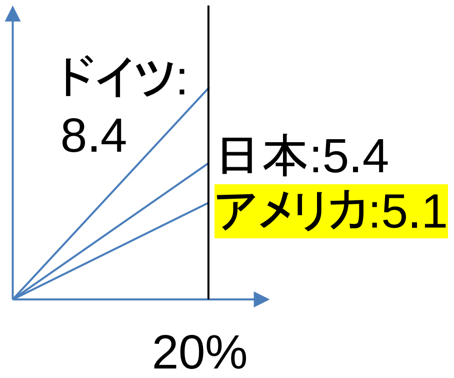
(ABランク)

II:　正しいですか？



II:　もし、「ジニ係数：アメリカが最も小さい」が正しければ、
「アメリカが最も平等」になるはず。
「アメリカが最も不平等」は矛盾している。

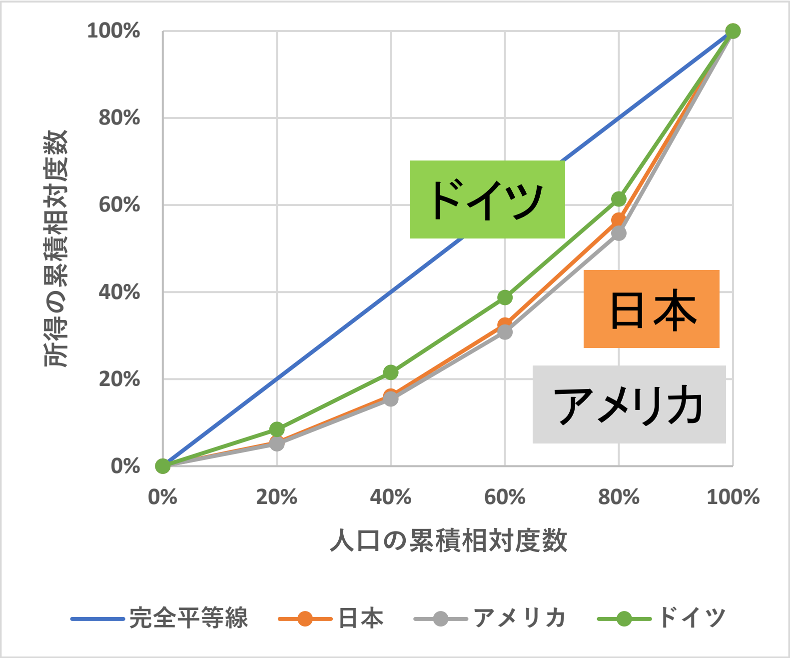
⇒　IIは正しくない



	第１五分位階級	第２五分位階級	第３五分位階級	第４五分位階級	第５五分位階級
日本	5.4	10.7	16.3	24.1	43.5
アメリカ	5.1	10.3	15.4	22.7	46.4
ドイツ	8.4	13.1	17.2	22.7	38.6

黄色の数値：第1～4の階級で、最小の値

アメリカ：第１～第４の各階級で、最小の値：
⇒ローレンツ曲線：他より、下側に位置する
⇒ジニ係数：大⇔不平等

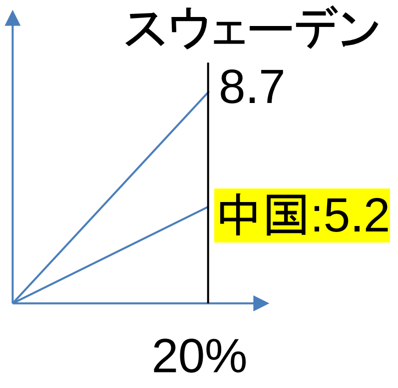


(p36.5)[C1]問8.ローレンツ曲線・ジニ係数の説明

(ABランク)

III:　正しいですか？

III:



	第１五分位階級	第２五分位階級	第３五分位階級	第４五分位階級	第５五分位階級
スウェーデン	8.7	14.3	17.8	23	36.2
中国	5.2	9.8	14.9	22.3	47.9

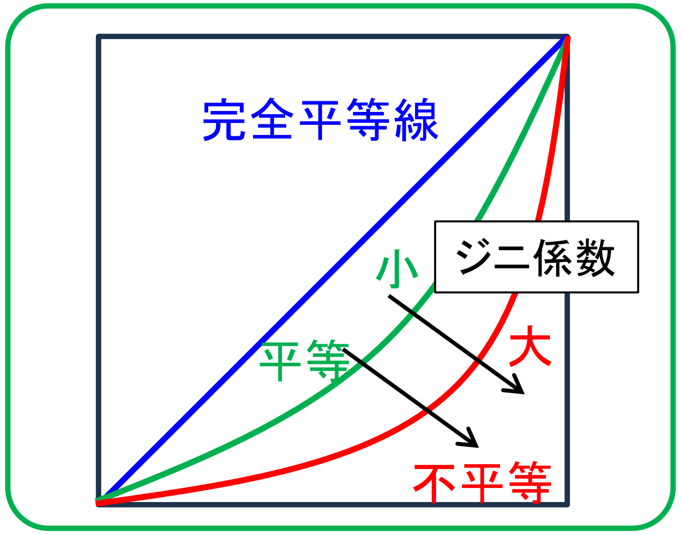
黄色の数値:第1～4の階級で、小さい値

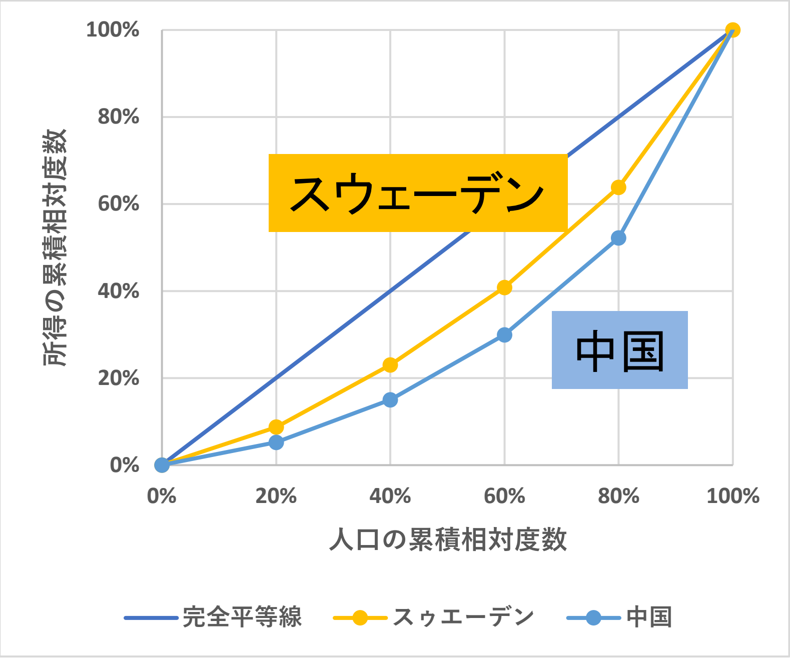
中国:第１～第４の階級で、小さい値：
⇒ローレンツ曲線：他より、下側に位置する
⇒ジニ係数：大⇔不平等　⇒ IIIは正しい

I　は正しい
II　は正しくない
III は正しい

(答)⑤

済





(p38.1a)[C1]問9.コレログラムの選択

コレログラムの説明

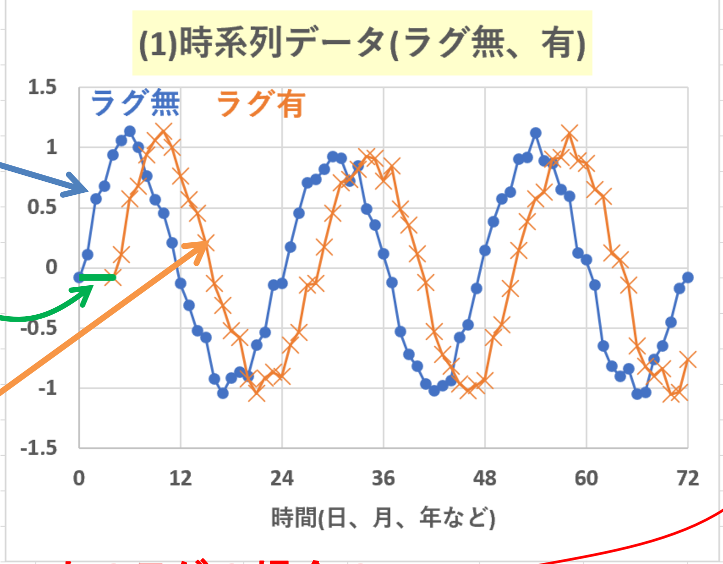
・・・時系列データの周期を調べる手法

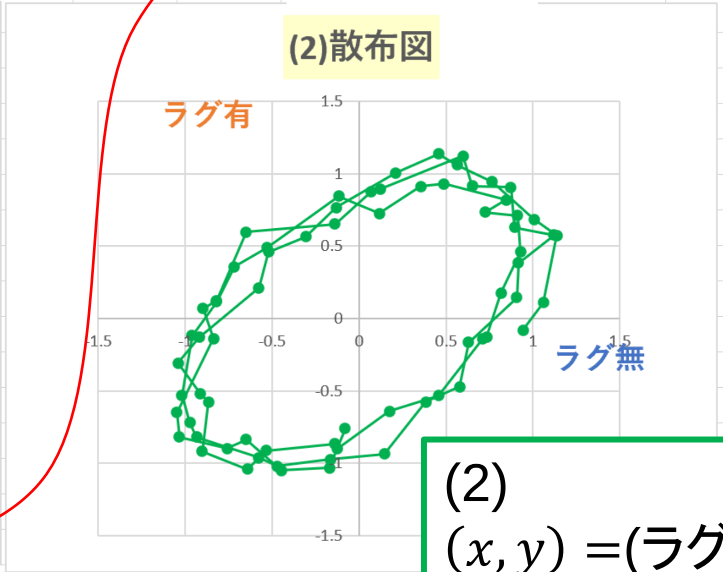
ラグ= 4
(1周期=24)
自己相関係数= 0.5037

(1)元の時系列データ(ラグ無)

ずらした量:ラグ

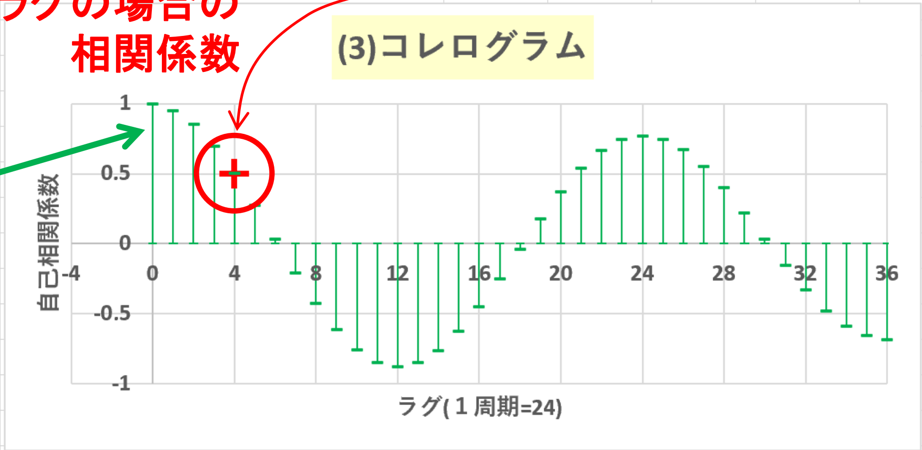
元の時系列データを右にずらしたもの(ラグ有)





(2)
(x, y) =(ラグ無、ラグ有)の散布図

上のラグの場合の相関係数

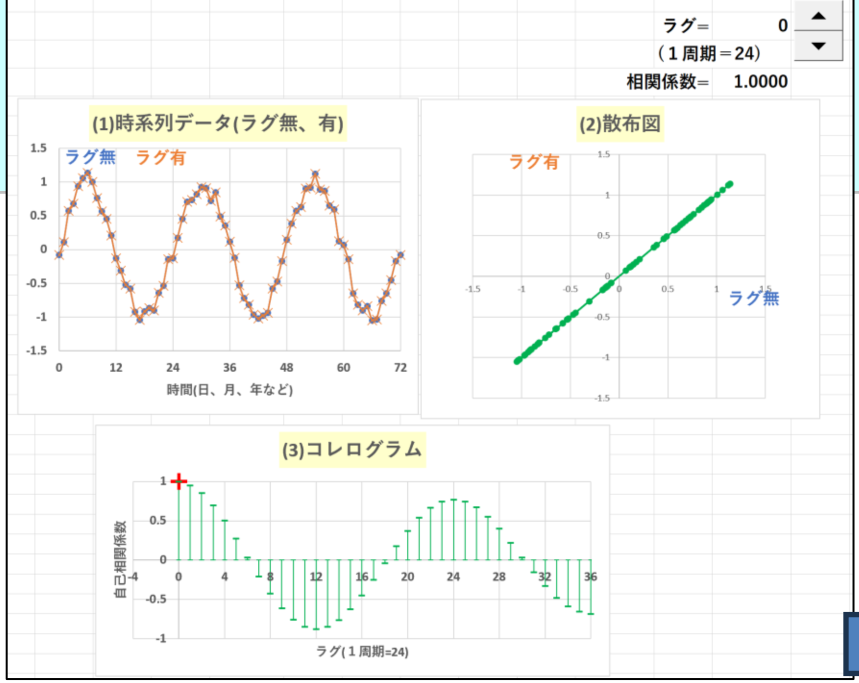


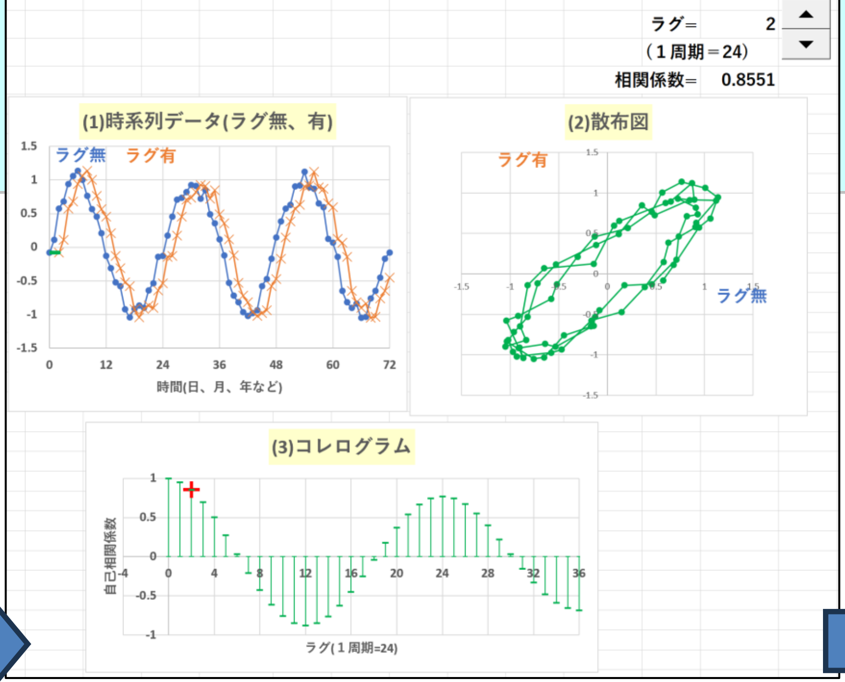
(3)コレログラム:
ラグを変えた時の(ラグ無、ラグ有)の相関係数
正確には自己相関係数

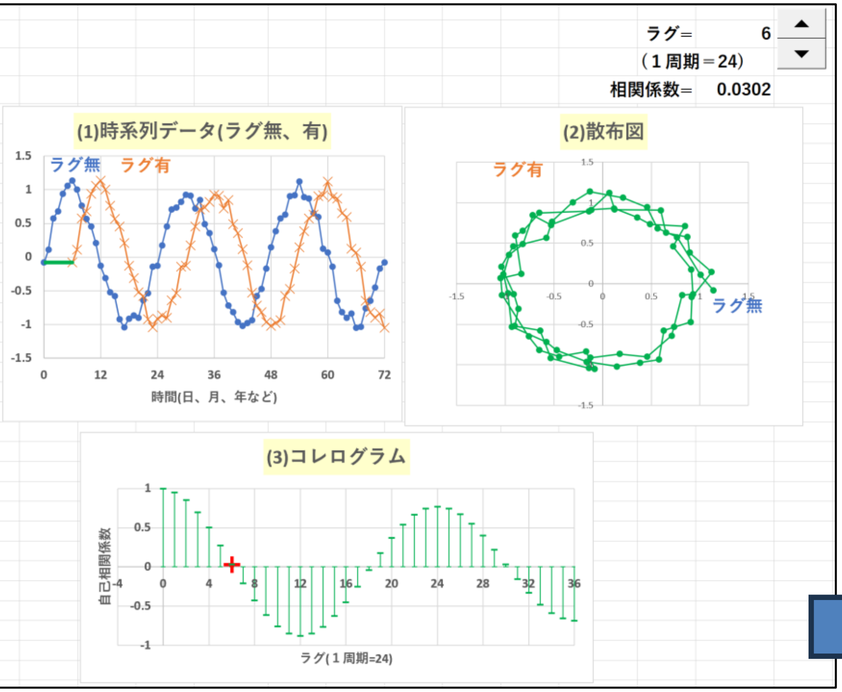
出題のされ方:
- 時系列データからコレログラムを選ぶ
- コレログラムから時系列データの特徴の読取

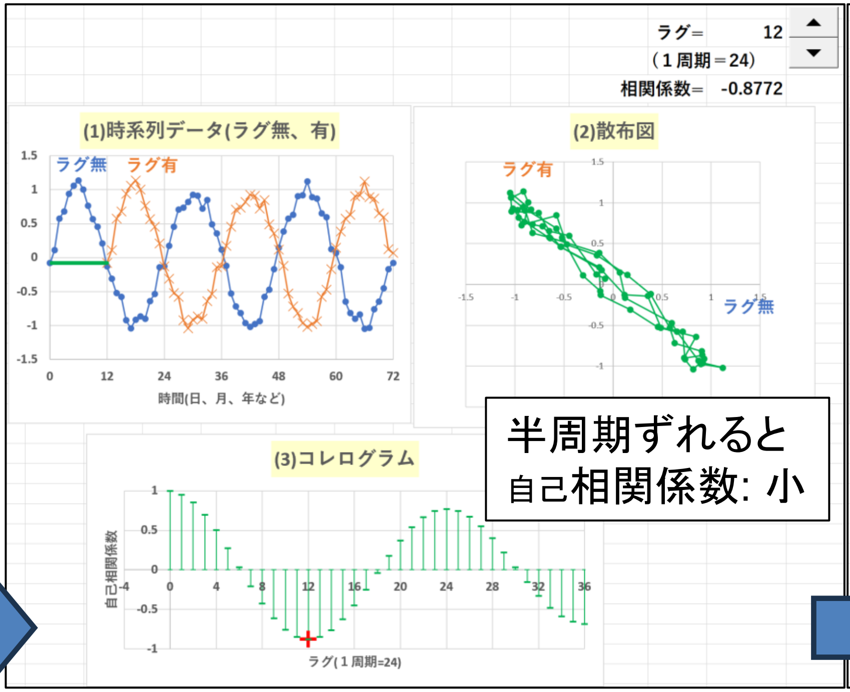
(p38.1b)

周期的に変動する
時系列データの場合

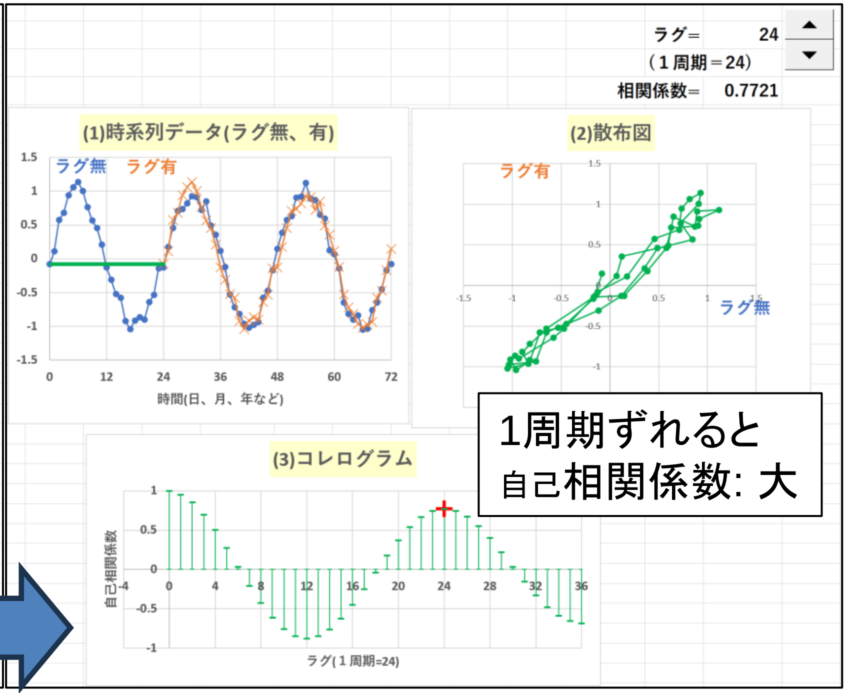








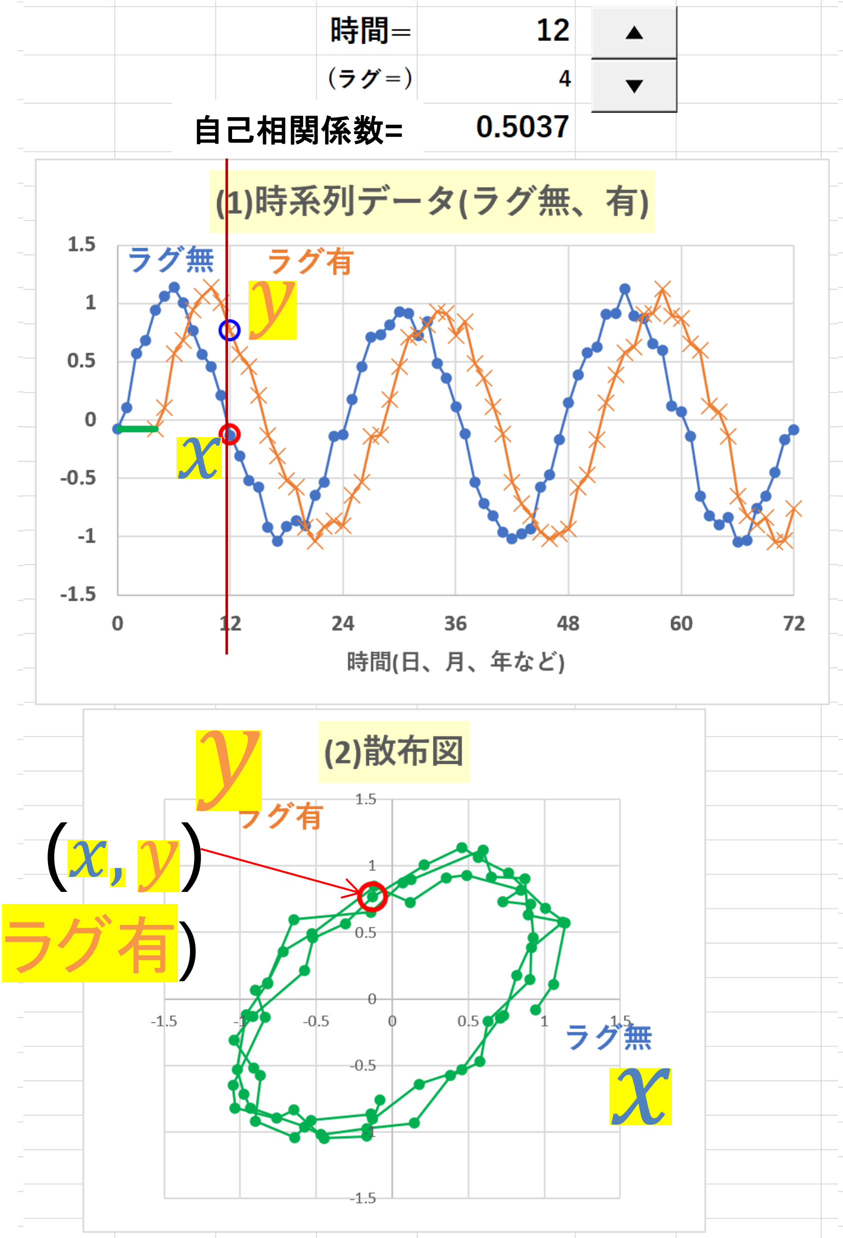
半周期ずれると
自己相関係数: 小

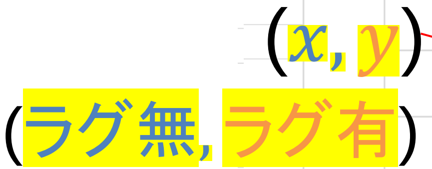


1周期ずれると
自己相関係数: 大

(p38.1c)[C1]問9.コレログラムの選択

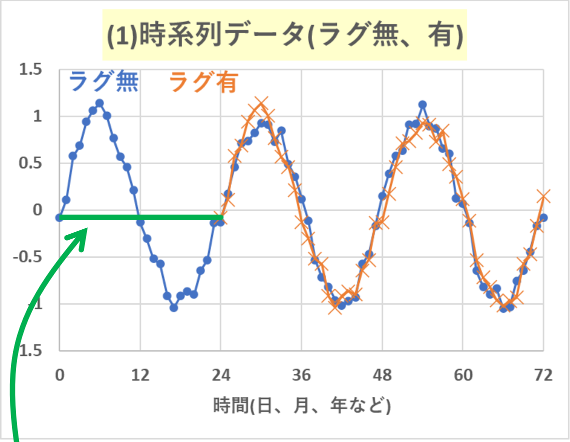
(1) 時系列データ　から
(2) 散布図へ





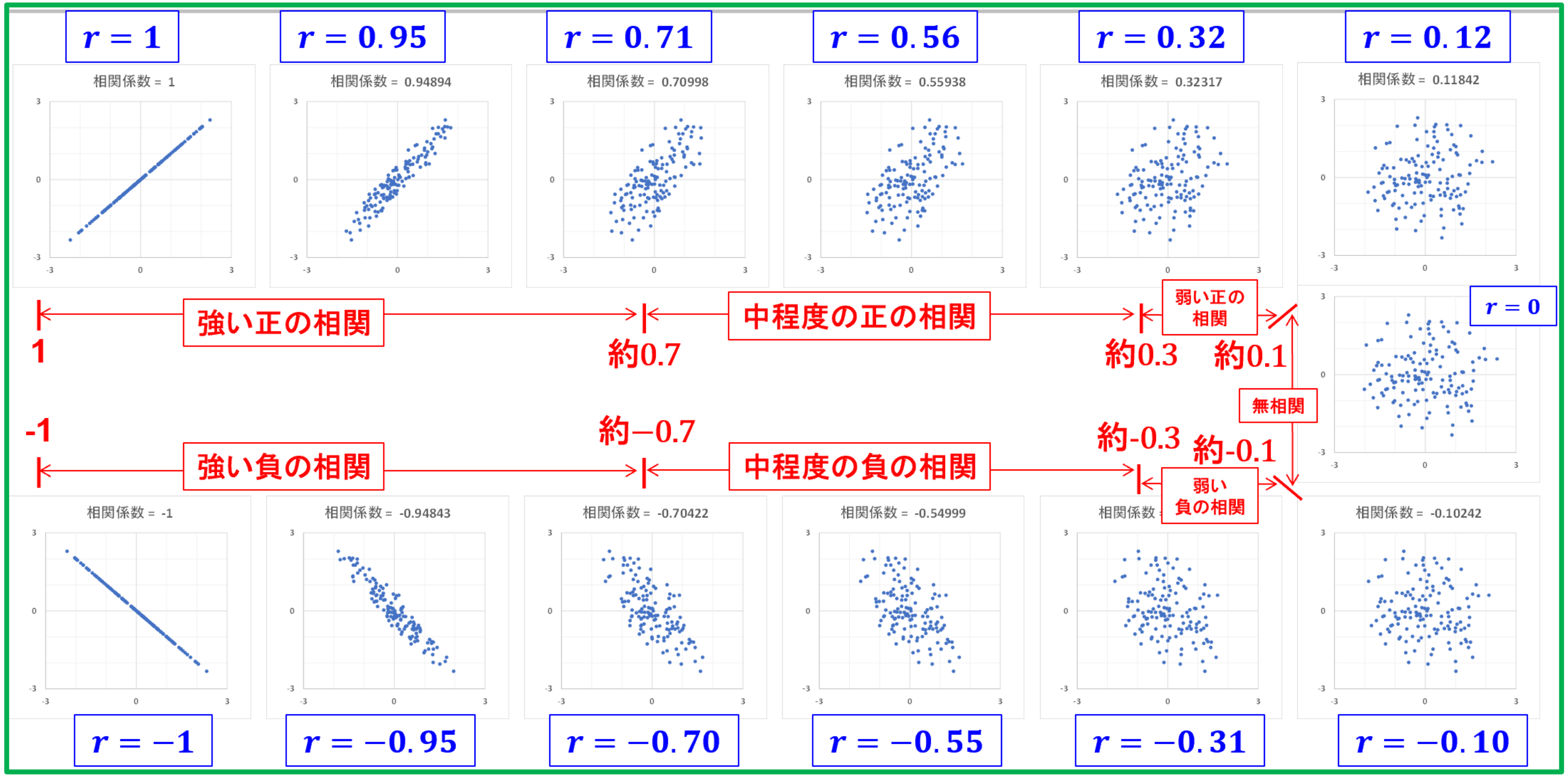
(p38.1d)[C1]問9.コレログラムの選択

(2)散布図　から
(3) (自己)相関係数へ
　（コレログラム）



ずれがある場合、
この部分の寄与は
計算しない⇒値が小さめ

散布図と相関係数の例



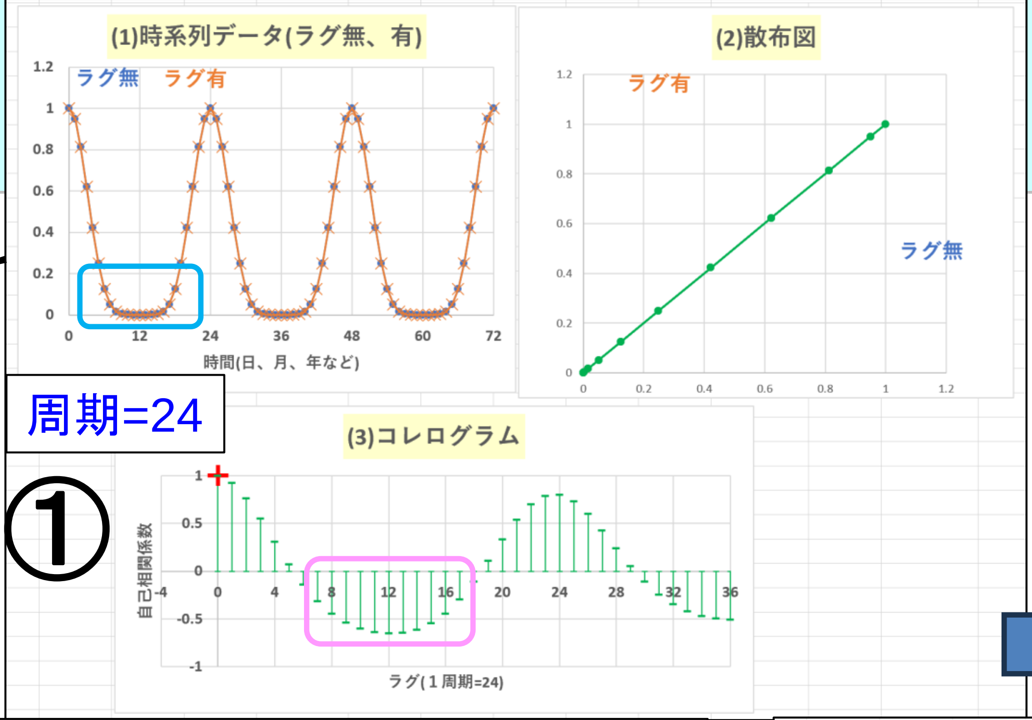
自己相関係数⇒最大値が1、最小値が-1にならない場合有

(p38.1e)

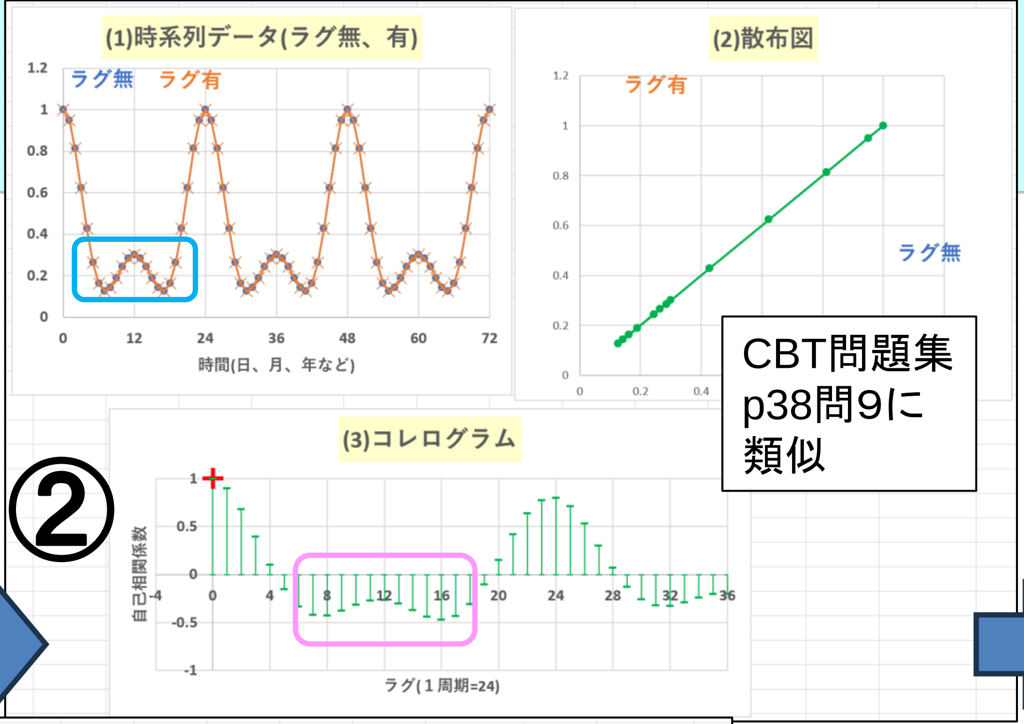
複雑な周期を持つ場合

周期=24

①

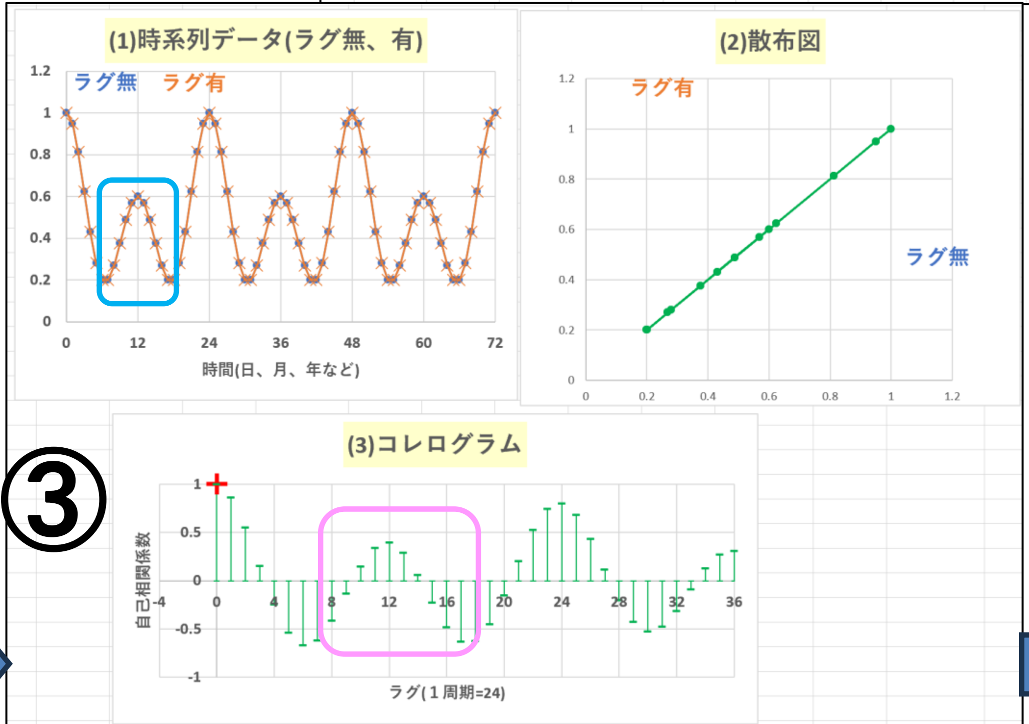


②



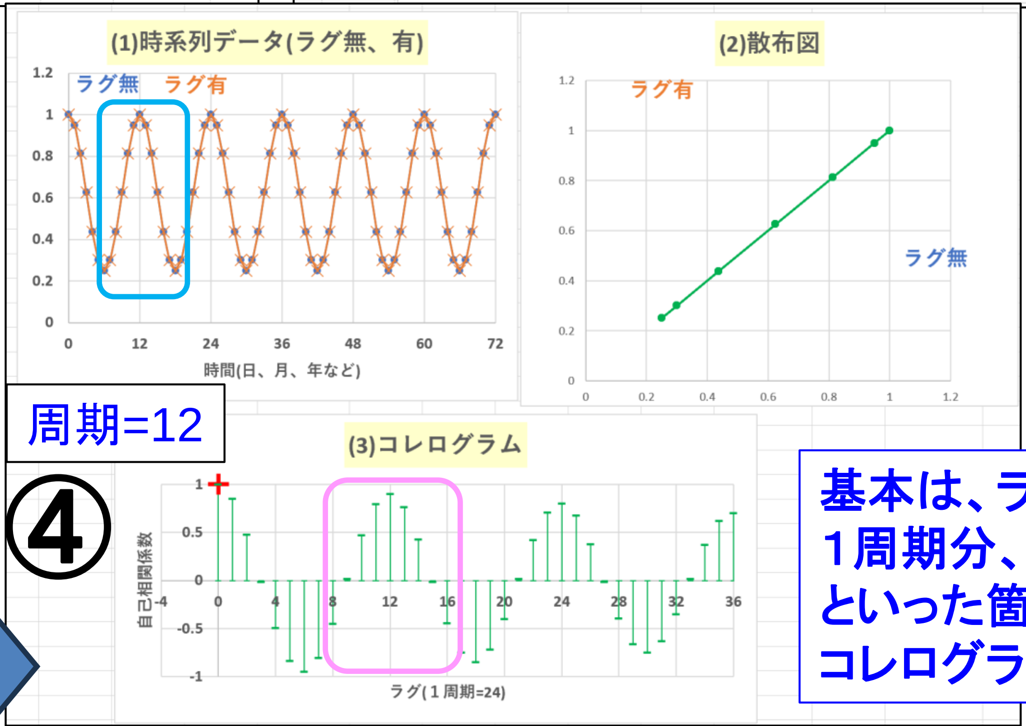
CBT問題集
p38問9に
類似

③



周期=12

④

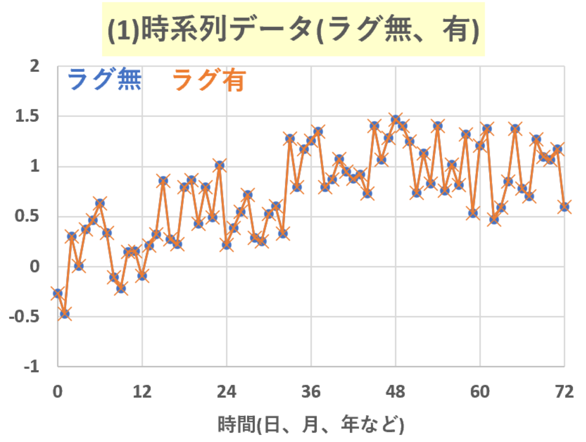


基本は、ラグの量が
1周期分、2周期分
といった箇所で、
コレログラムが大きくなる

(p38.1f)[C1]問9.コレログラムの選択

周期的に変動しない時系列データの場合

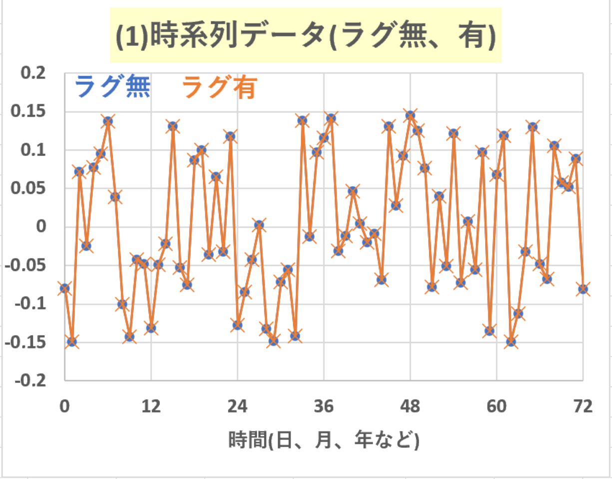
緩やかな変化＋ばらつき有

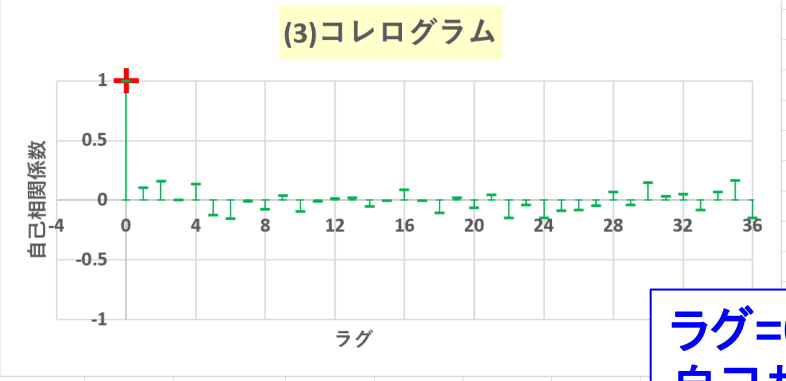


緩やかに変化する

類題：
統計検定2級
2017年6月
問3①　③

時系列の相関が無い場合



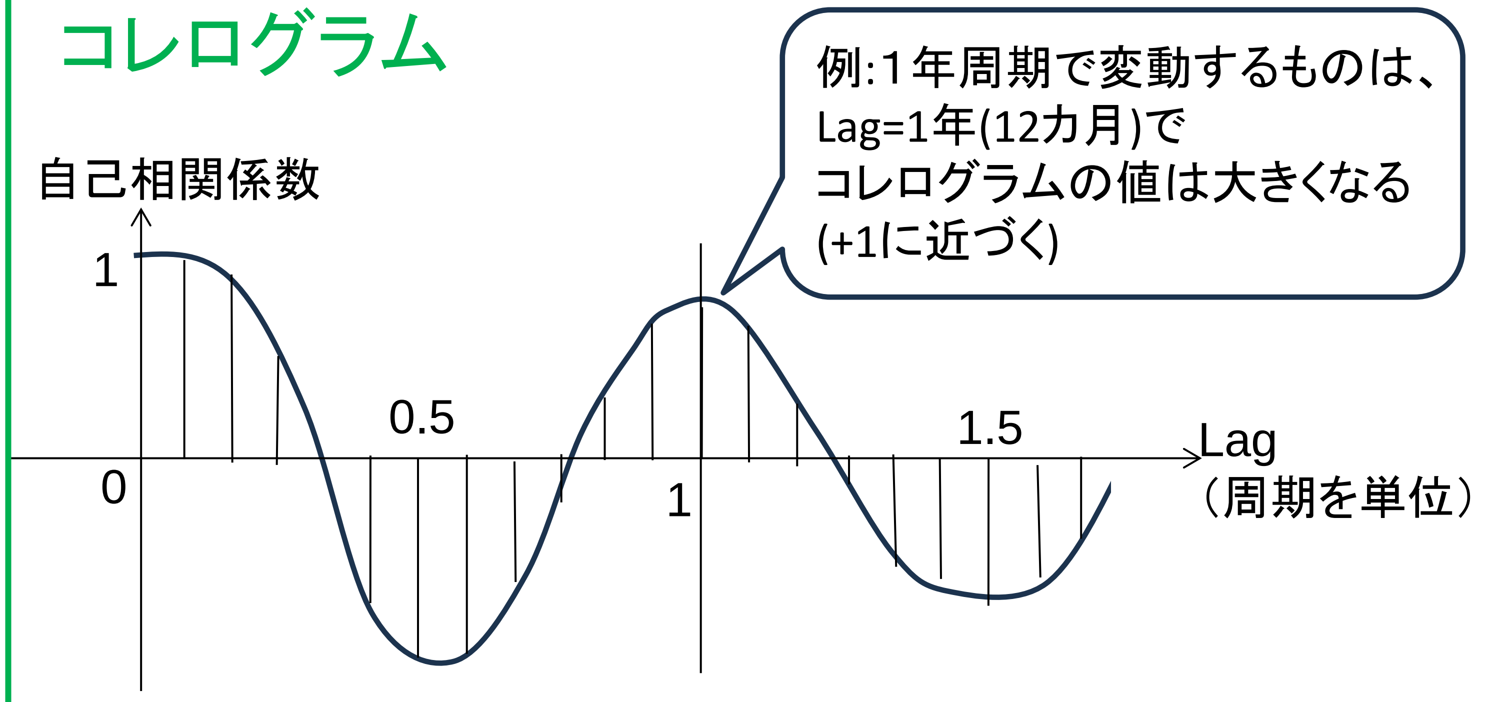


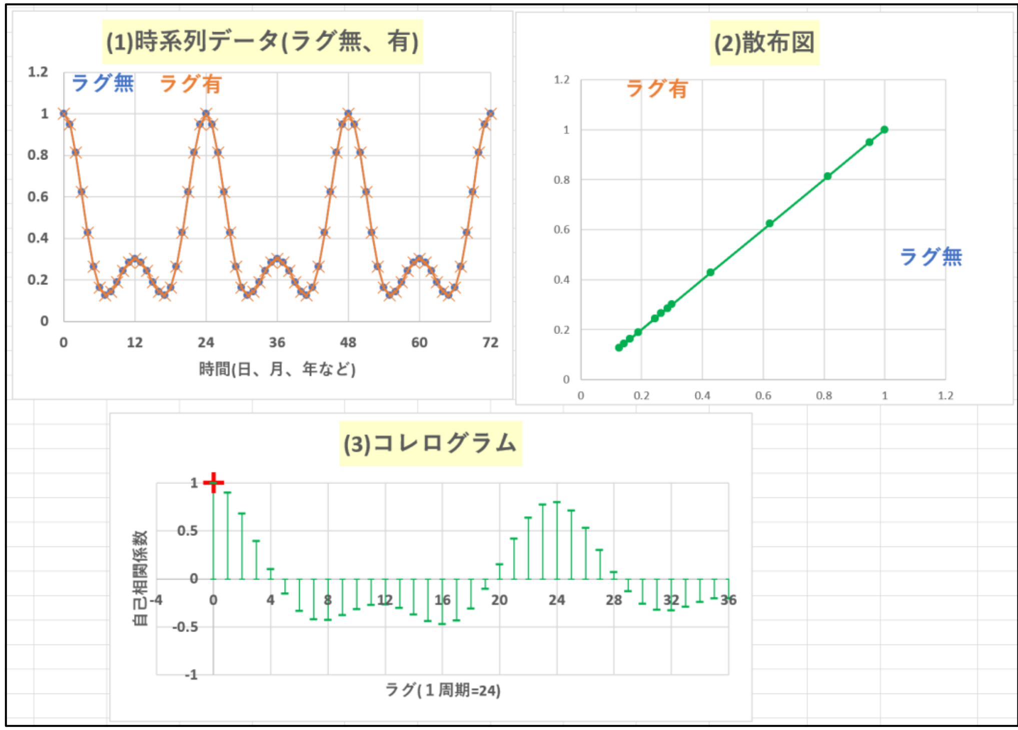
ラグ=0　以外は
自己相関係数≒0

(p38.2)[C1]問9.コレログラムの選択

(ABランク)

コレログラム





(答)②

済

(p40.1)[C1]問10.ラスパイレス指数の計算式

ラスパイレス指数とパーシェ指数

	商品A (例:梨)		商品B(例:ぶどう)	
2024年(比較年)	400 400 400 400 400	5個(円)	900 900	2個(円)
2023年(基準年)	300 300 300	3個(円)	600 600 600	3個(円)

Q: 全体として、
モノの値段は
何倍になりましたか？

売れた個数が違うので
単純比較できない

①基準年で売れた個数
に揃えて比を求める

消費者物価指数として使用

ラスパイレス指数=

$$\frac{400\ 400\ 400\ \text{3個} + 900\ 900\ 900\ \text{3個}}{300\ 300\ 300\ \text{3個} + 600\ 600\ 600\ \text{3個}} \times 100$$

頻繁に出題
(例)’18年11月問4

②比較年で売れた個数
に揃えて比を求める

パーシェ指数=

$$\frac{400\ 400\ 400\ 400\ 400\ \text{5個} + 900\ 900\ \text{2個}}{300\ 300\ 300\ 300\ 300\ \text{5個} + 600\ 600\ \text{2個}} \times 100$$

(例)’21年6月問3
出題

(p40.2)[C1]問10.ラスパイレス指数の計算式

ラスパイレス指数

$$= \frac{400 \times 3 + 900 \times 3}{300 \times 3 + 600 \times 3} \times 100$$

パーシェ指数

$$= \frac{400 \times 5 + 900 \times 2}{300 \times 5 + 600 \times 2} \times 100$$

公式

ラスパイレス指数

$$\frac{\text{比較年の値段} \times \text{基準年の個数}}{\text{基準年の値段} \times \text{基準年の個数}}$$

$$\frac{\text{比較年の値段} \times \text{比較年の個数}}{\text{基準年の値段} \times \text{比較年の個数}}$$

パーシェ指数

値段：調べやすい
個数：調べるのは大変
(特に比較年の「個数」を
調べるのは大変)
⇒基準年の(古い)個数を使う
ラスパイレス指数の方がより簡単
(CBT問題集p41,下から4行目
以降を参照ください)

「いつの個数を使うか」を
(基準年or比較年?)
おぼえてください！

(p40.3)[C1]問10.ラスパイレス指数の計算式

(ABランク)

CBT問題集の問題

	2016年		2017年	
	購入数量	平均価格	購入数量	平均価格
梨	3827	48.86	使わない	49.30
ぶどう	2422	107.09	使わない	115.36

$$\text{ラスパイレス指数} = \frac{49.30 \times 3827 + 115.36 \times 2422}{48.86 \times 3827 + 107.09 \times 2422} \times 100$$

梨 ← | → ぶどう

公式

$$\text{ラスパイレス指数} = \frac{\text{比較年の値段} \times \text{基準年の個数}}{\text{基準年の値段} \times \text{基準年の個数}}$$

$$\text{パーシェ指数} = \frac{\text{比較年の値段} \times \text{比較年の個数}}{\text{基準年の値段} \times \text{比較年の個数}}$$

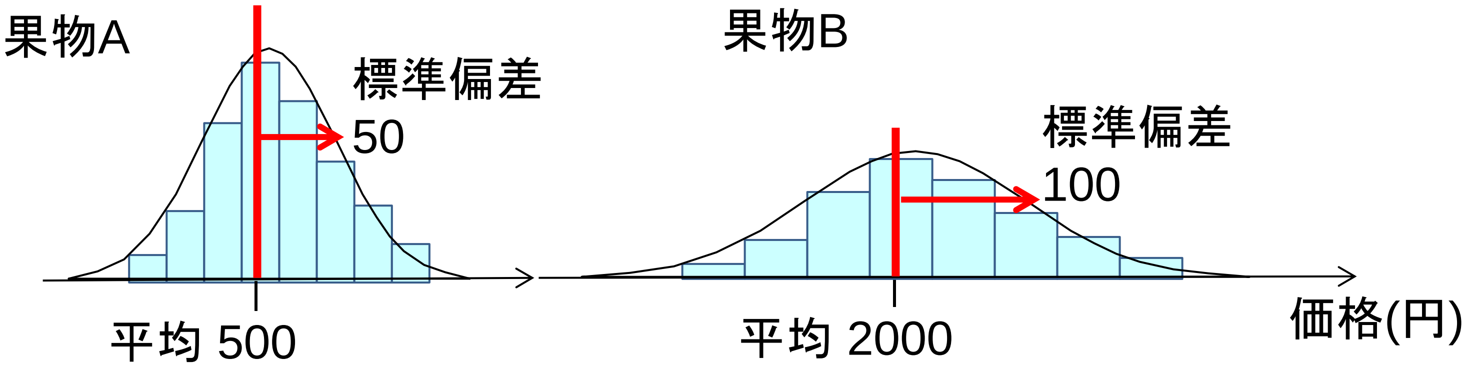
済

(答)②

(p41.1)[C1]追加事項:変動係数 (CBT本に記載無)

出題例: 2017年11月問3

例: ある期間の、果物Aと果物Bの価格の分布



Q: ばらつきは
どちらが大きいですか？

標準偏差は、
果物Bが大きいけど、
Bは平均も大きいし･･･

$$\text{変動係数} = \frac{50}{500} = 0.1$$

$$\text{変動係数} = \frac{100}{2000} = 0.05$$

変動係数は、果物Aの方が大きい

・$\text{変動係数} = \cancel{\dfrac{\text{平均}}{\text{標準偏差}}}$ にご注意ください

・$\text{平均} > 0$ の場合にのみ使用

平均が異なるものの
ばらつきの比較には

$$\text{変動係数} = \frac{\text{標準偏差}}{\text{平均}}$$

を使うといい

⇒果物A,Bでは、どうなりますか？