ひかり統計塾

統計検定２級
公式問題集(CBT対応版)の解説
Part.2 カテゴリー1～10
Part.3 模擬テスト
お試しサイト版

統計検定2級　CBT問題集　目次

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: 1変数記述統計の分野
問1,2(p24-27)　@お試しサイト

(p24.1)[C1]問1.相対度数の計算

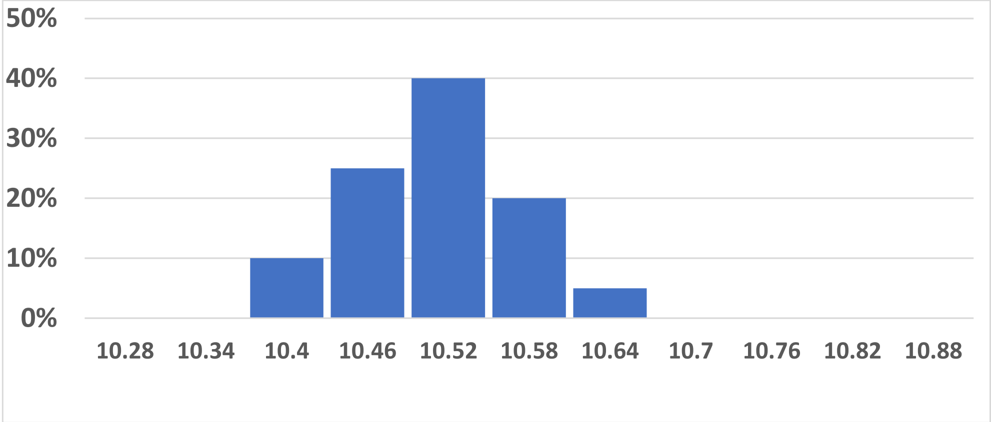
(Aランク)

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)

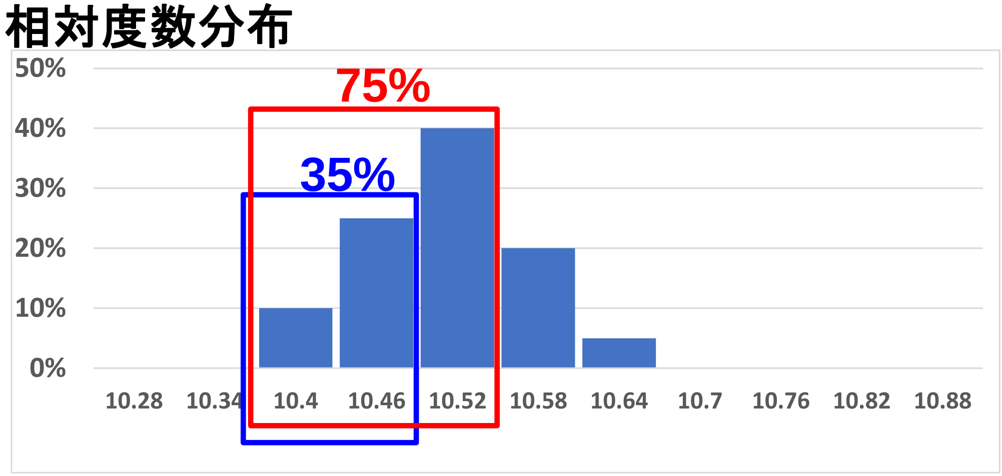
類題 **度数分布表**

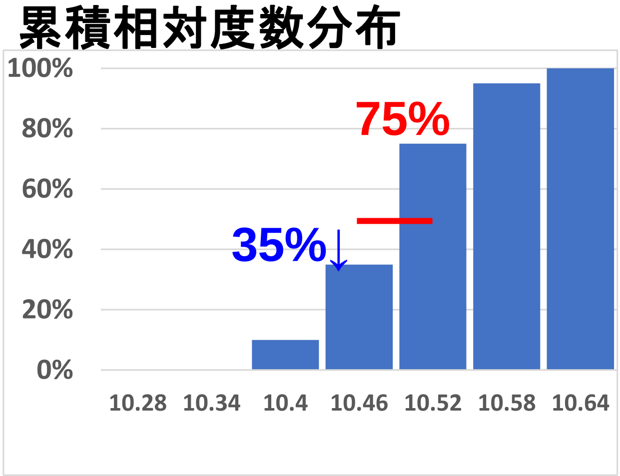
Q. 中央値が含まれている階級番号は？

(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 ←50%を超えていない
5	10.490	～	10.550	10.520	80	40	75 ←50%を超えている（ここに中央値が含まれている)
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	

(答)階級No.5





(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

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: 1変数記述統計の分野
問9 (p38-39)コレログラムの基礎 @お試しサイト

(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

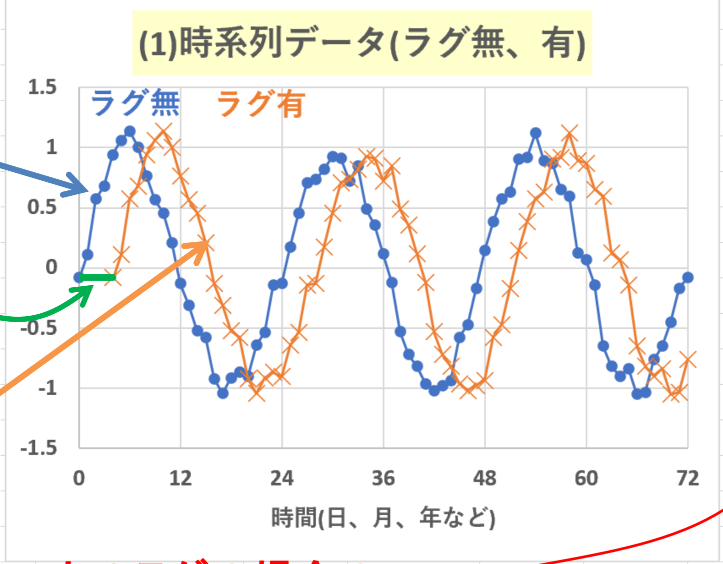
（1 周期＝24）

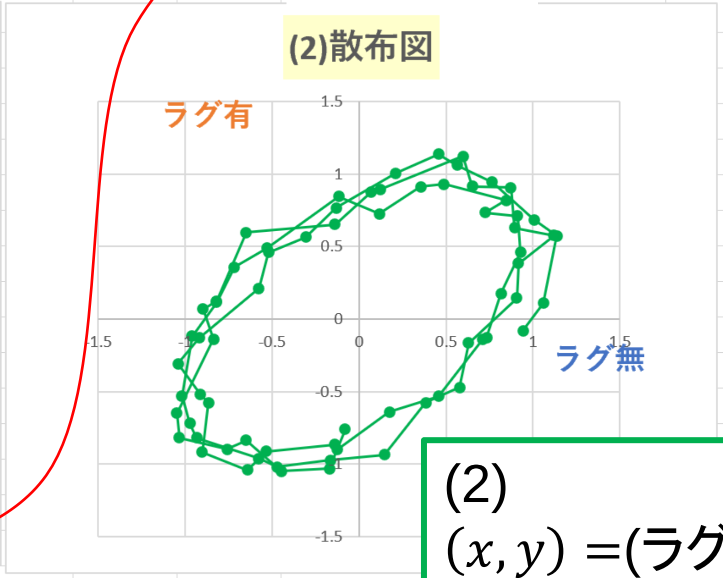
自己相関係数= 0.5037

(1)元の時系列データ(ラグ無)

ずらした量：ラグ

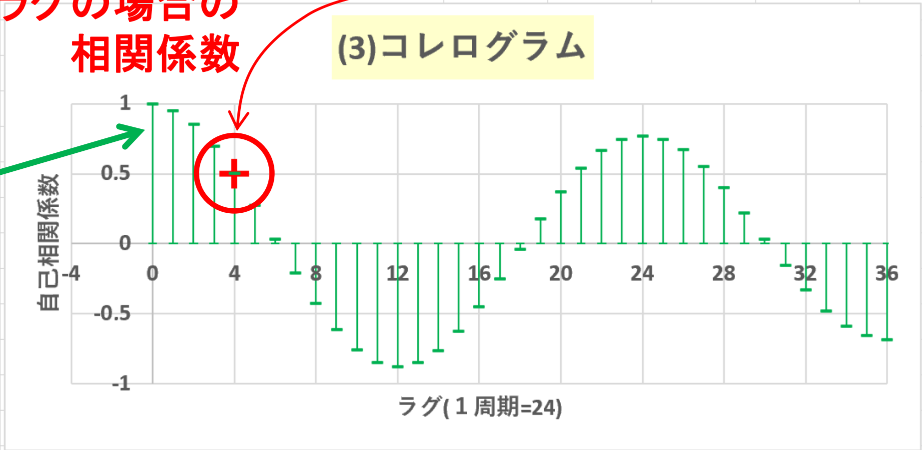
元の時系列データを右にずらしたもの(ラグ有)





(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



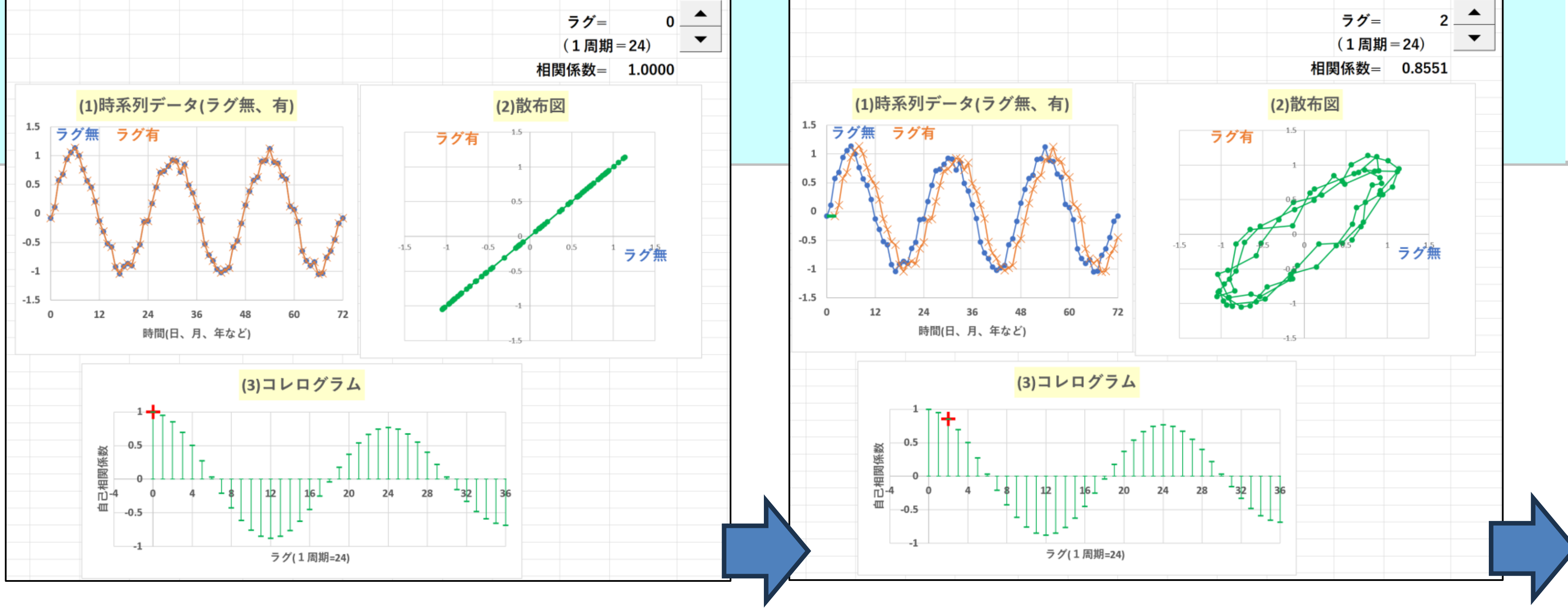
(3)コレログラム：ラグを変えた時の(ラグ無、ラグ有)の相関係数

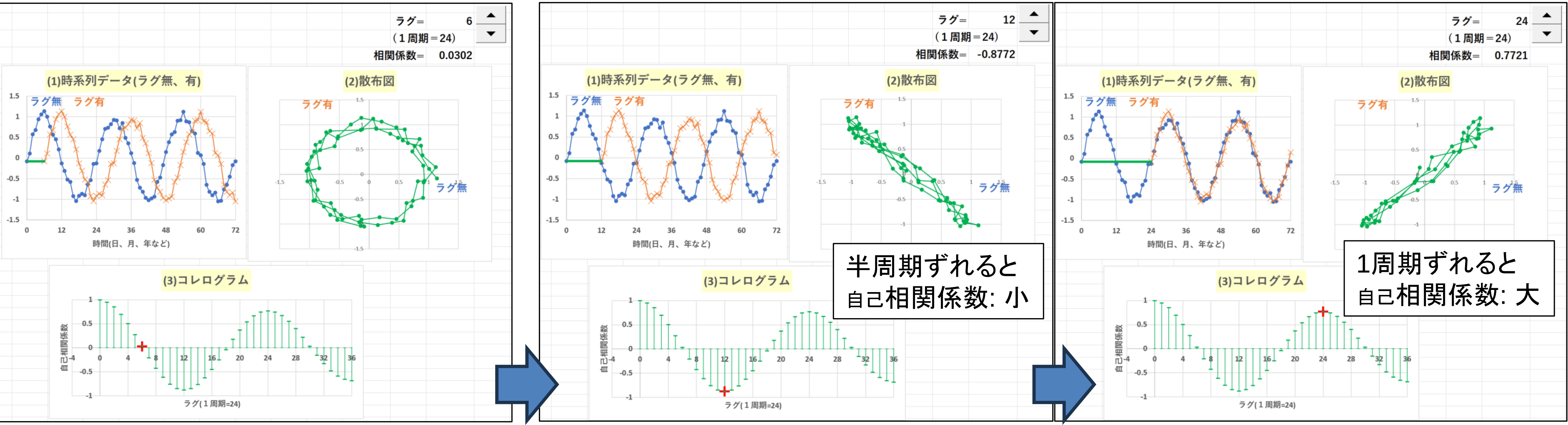
正確には自己相関係数

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)

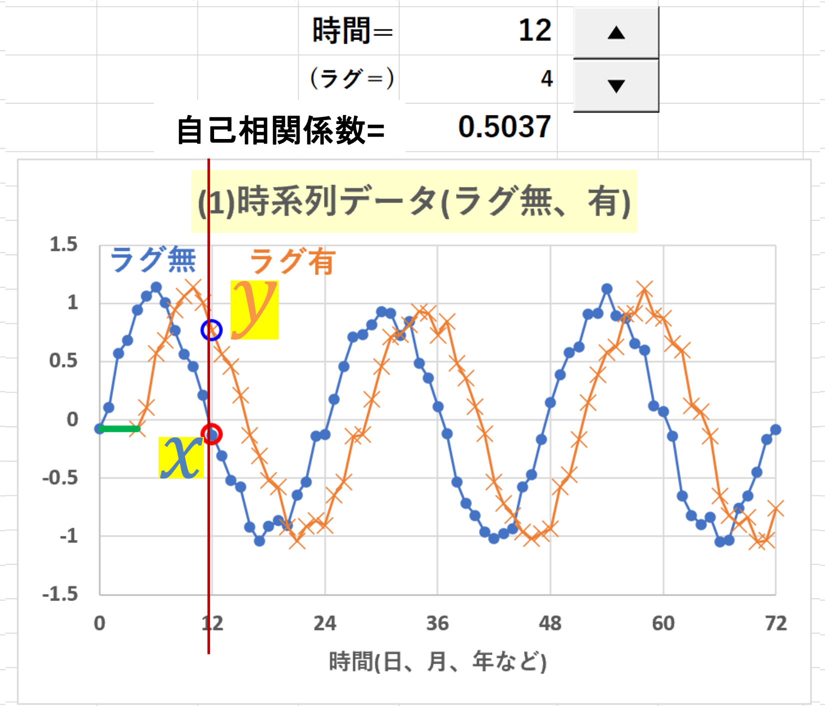
周期的に変動する
時系列データの場合

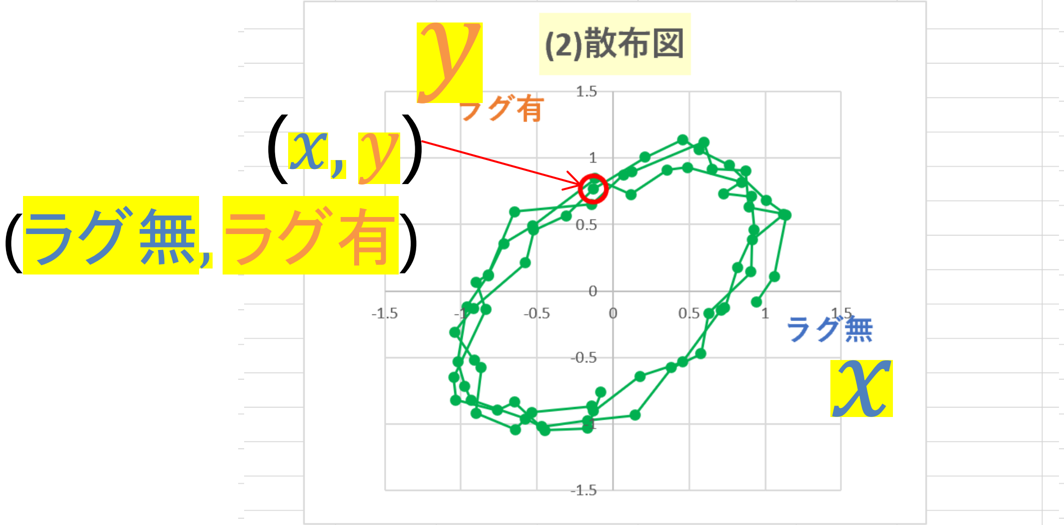




(p38.1c)[C1]問9.コレログラムの選択

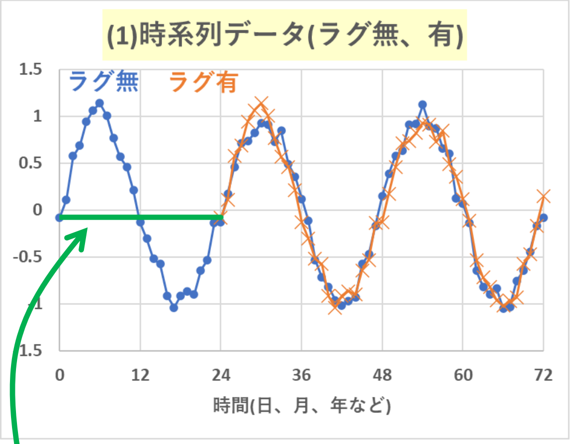
(1) 時系列データ から
(2) 散布図へ





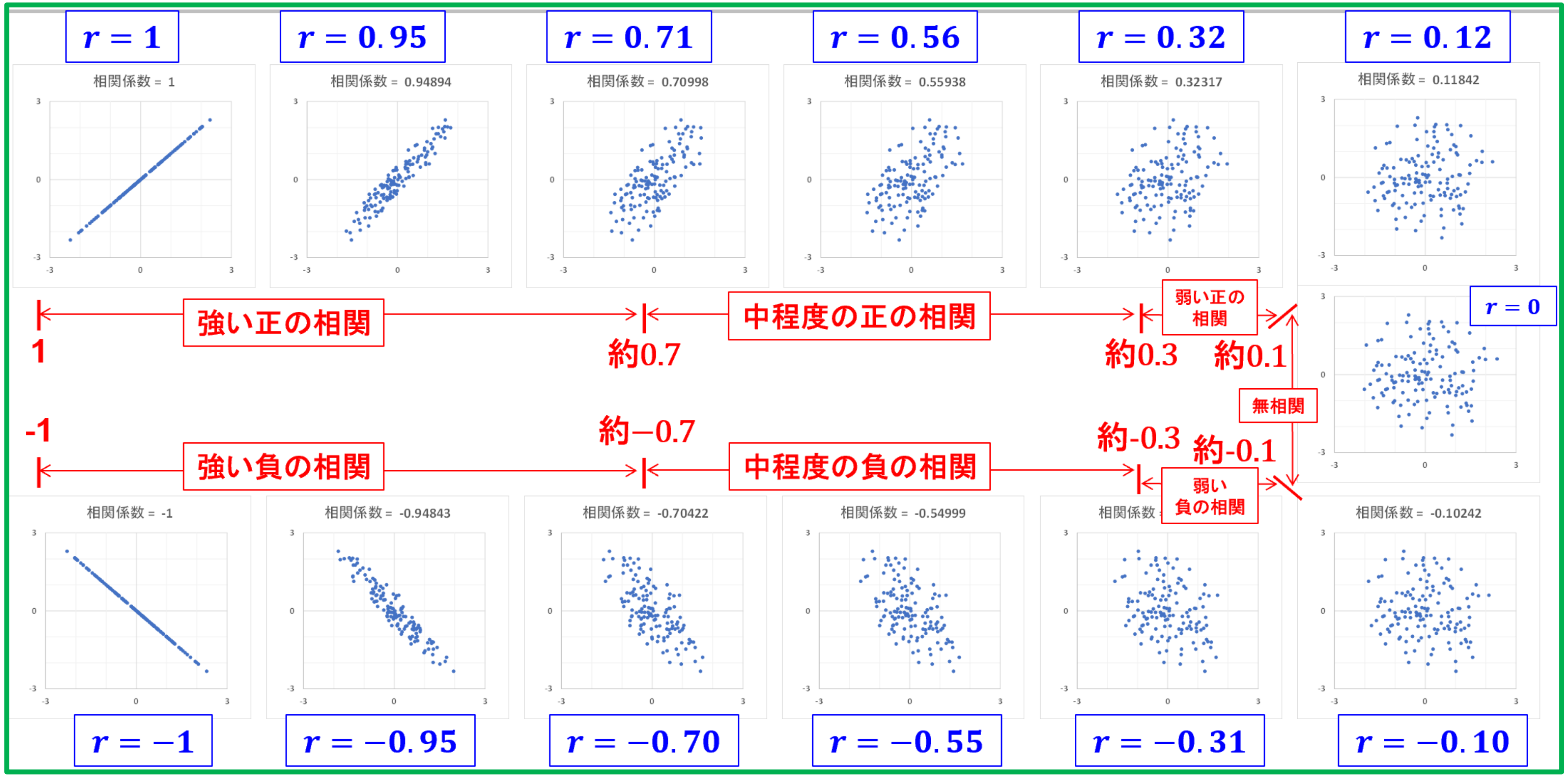
(p38.1d)[C1]問9.コレログラムの選択

(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

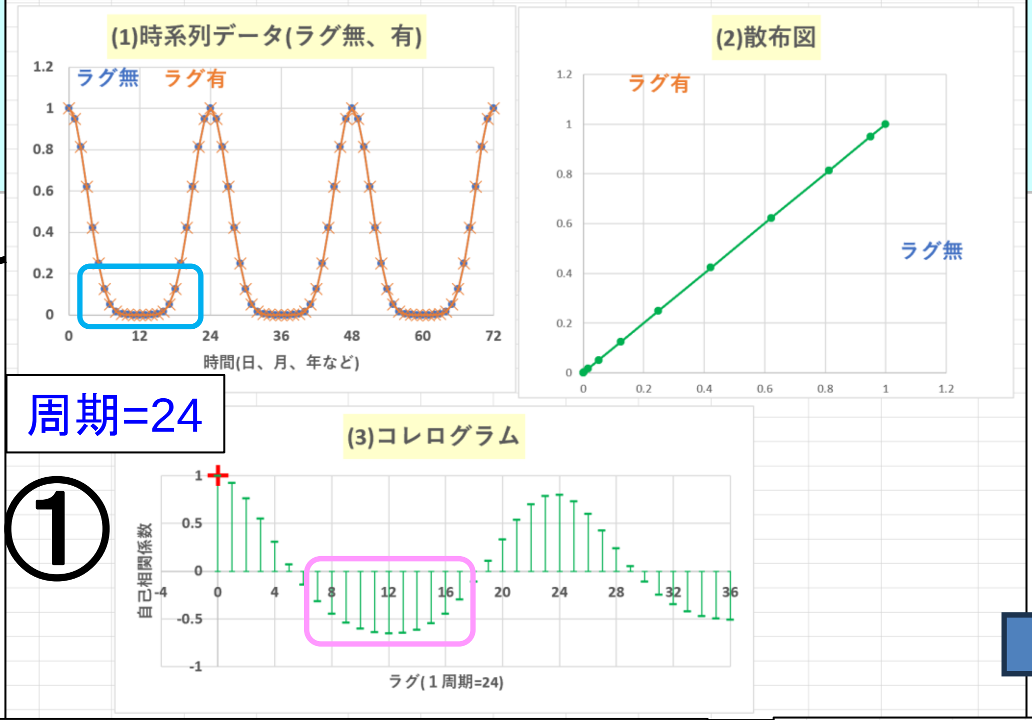
散布図と相関係数の例

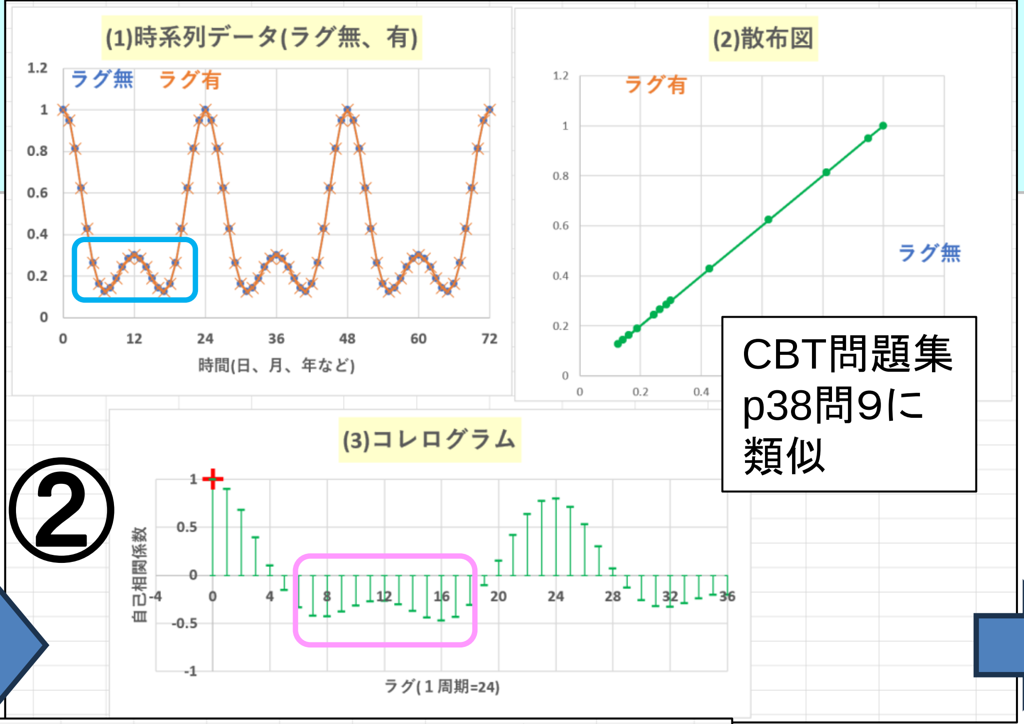


自己相関係数⇒最大値が1、最小値が-1にならない場合有

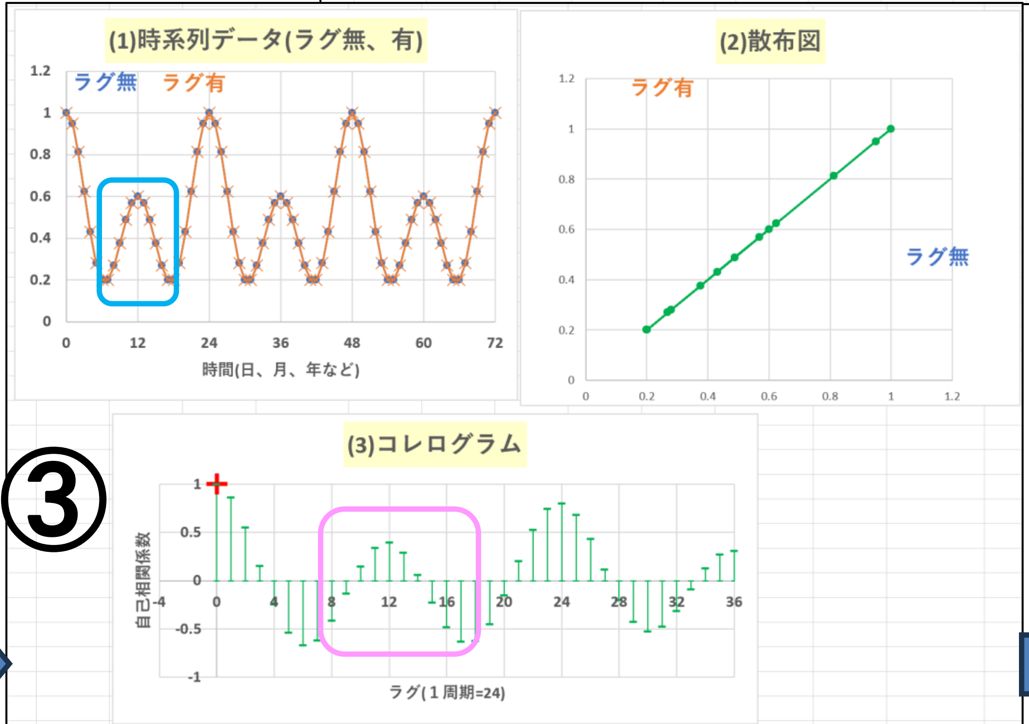
(p38.1e)

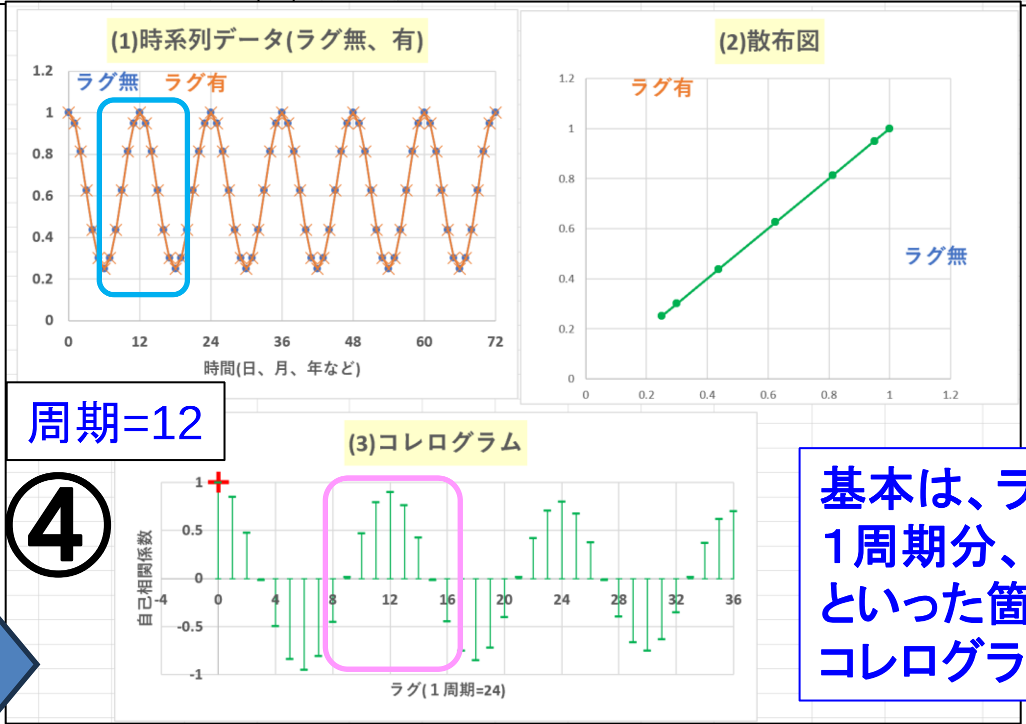
複雑な周期を持つ場合





CBT問題集
p38問9に
類似



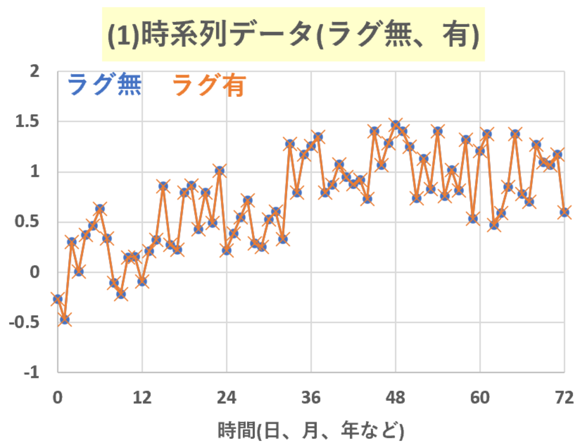


基本は、ラグの量が
1周期分、2周期分
といった箇所で、
コレログラムが大きくなる

(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

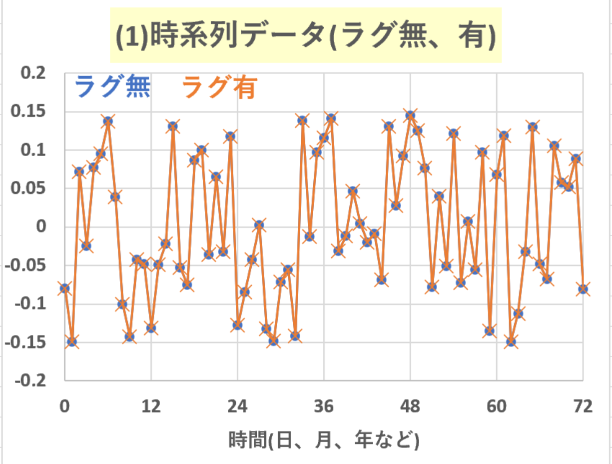
緩やかな変化＋ばらつき有

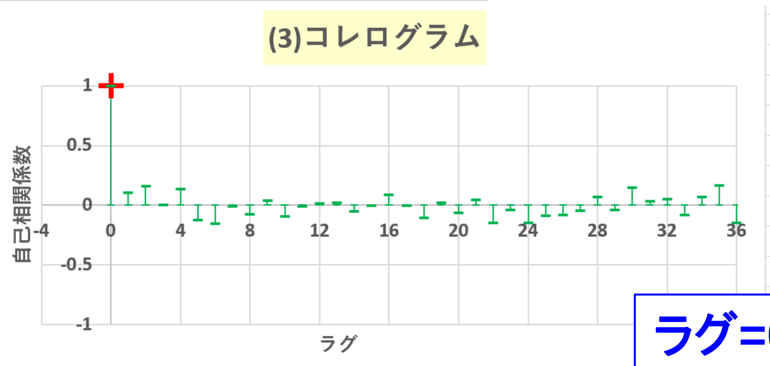


緩やかに変化する

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



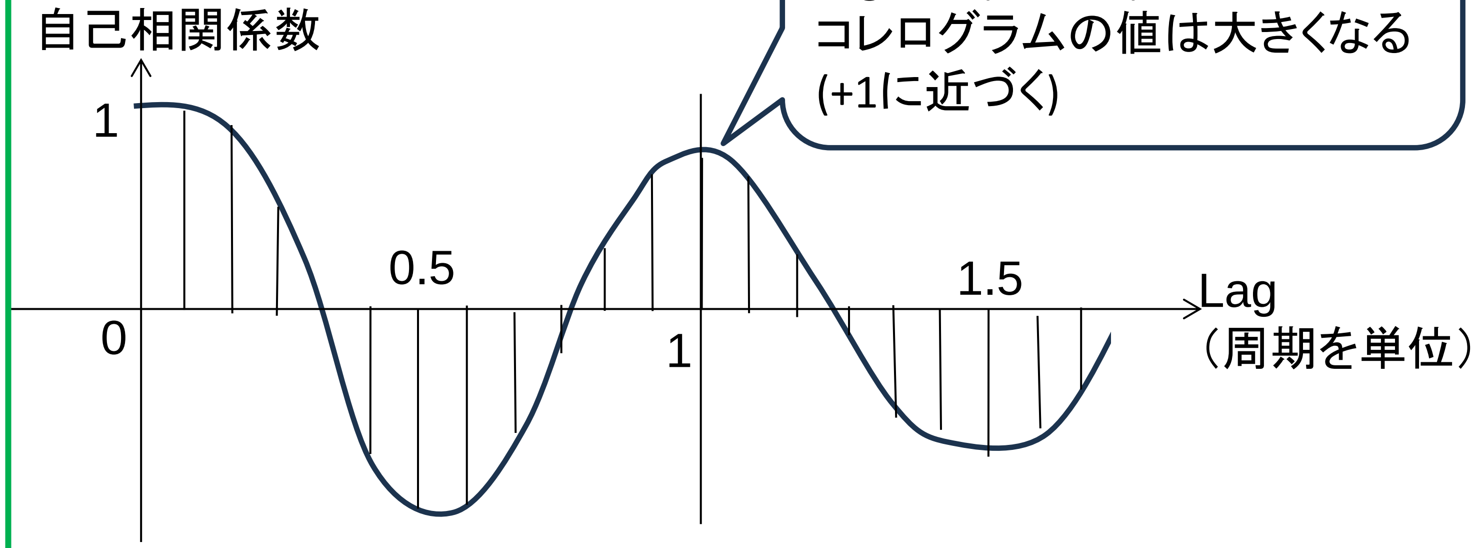


ラグ=0　以外は
自己相関係数≒0

(p38.2)[C1]問9.コレログラムの選択

(ABランク)

コレログラム

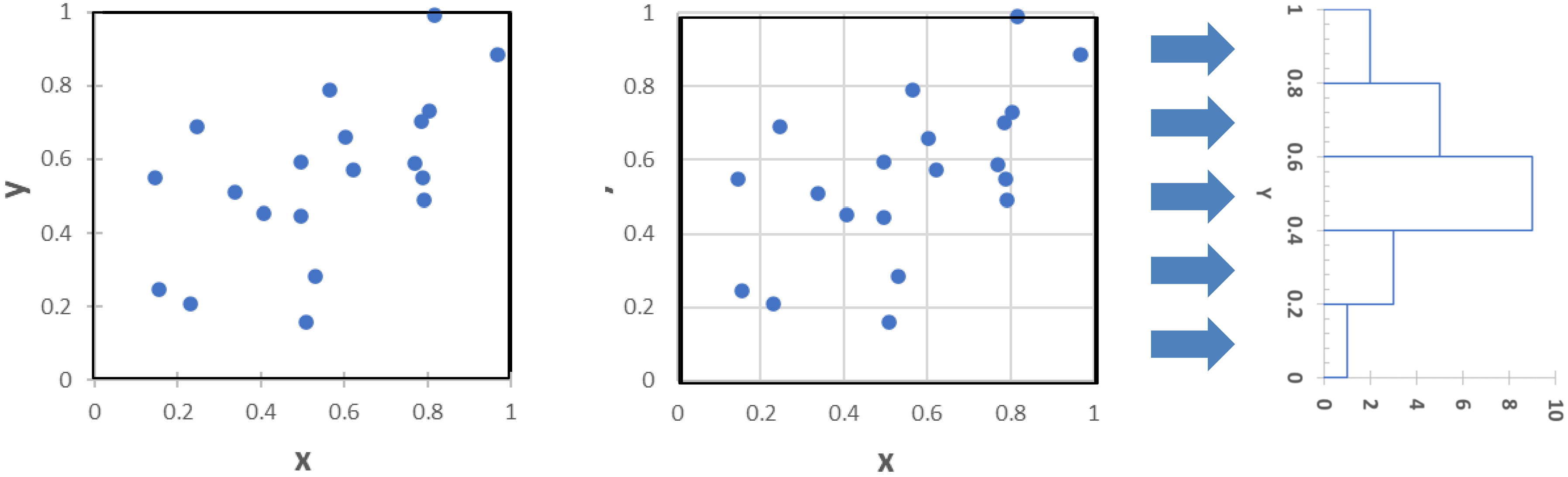


ひかり統計塾

統計検定２級
公式問題集(CBT対応版)の解説
カテゴリー2： 2変数記述統計の分野
問1（p42-43）@お試しサイト

(p42.1)[C2]問1. 散布図と度数分布

(例)　(注)CBT問題集のデータとは異なります



目盛線があればやりやすいですが、
問1の散布図にはありませんね

考え方：
散布図より、各階級の度数を調べる(対応付ける)
特に、端の点に着目する

(p42.2)[C2]問1. 散布図と度数分布

(Aランク)

p42の散布図の縦軸(女性の50歳時未婚率)⇔ヒストグラムの横軸

・散布図の縦軸の範囲・・・最小値:8～10の範囲にある　⇒②、④、⑤は最小値が10以上⇒×

最大値:18～20の範囲にある ⇒④、⑤は最大値が20以上⇒×

⇒②③④は×、　①③は〇

・①、③の違い？　⇒10～12の度数が①は少なく(5以下)、③は多い(10以上)

⇒散布図の縦軸で10%～12%の点:6個以上ある　⇒ ①③ の中で、①は×、　③は〇

考え方:
散布図より、各階級の度数を調べる(対応付ける)
特に、端の点に着目する

⇒　(答)③　済

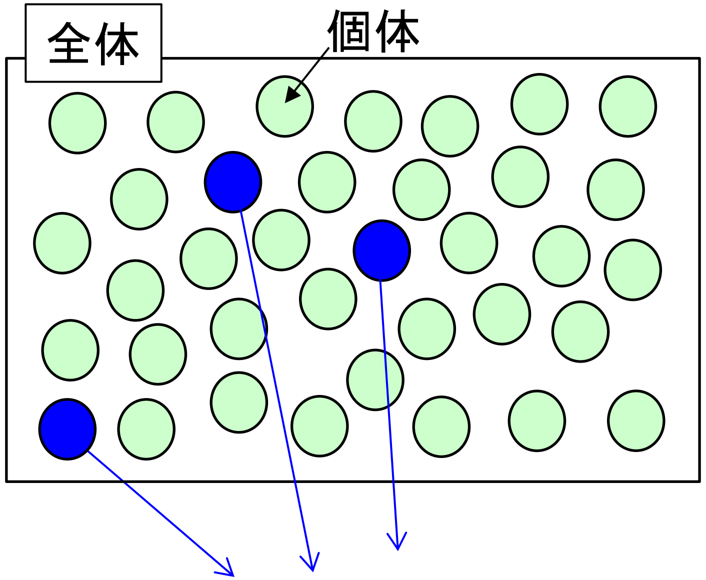
ひかり統計塾

統計検定２級
公式問題集(CBT対応版)の解説
カテゴリー3: データ収集の分野
問1 (p52-53) @お試しサイト

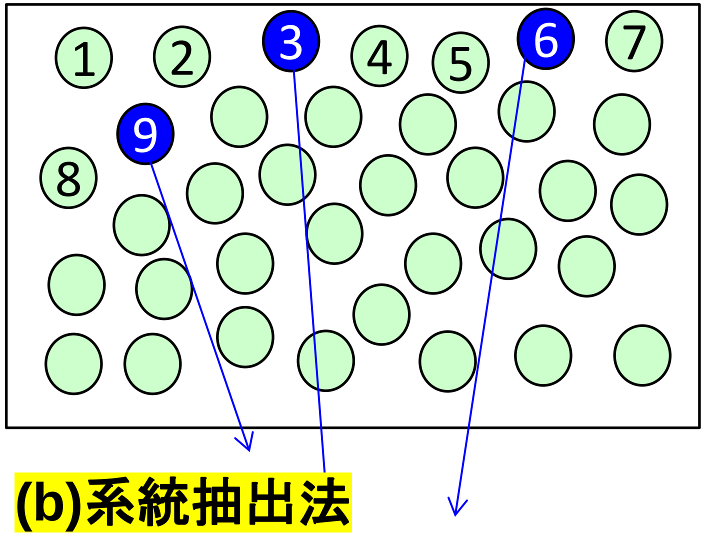
(p52.1)[C3]問1. 各標本抽出法の性質

(a)単純無作為抽出法
(b)系統抽出法
(c)層化無作為抽出法
(d)多段抽出法
(e)クラスター(集落)抽出法

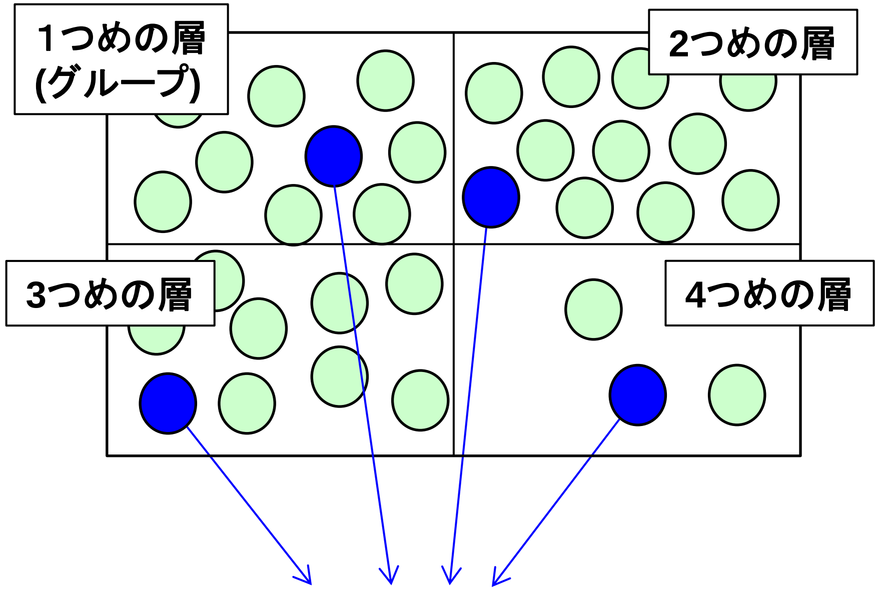
出題例:
・2021年6月問7
・2019年11月問6　(この問題)
・2019年6月問6
・2018年11月問5,問6
・2018年6月問6
（↑この期間、毎回出題あり）
・・・



(a)単純無作為抽出法

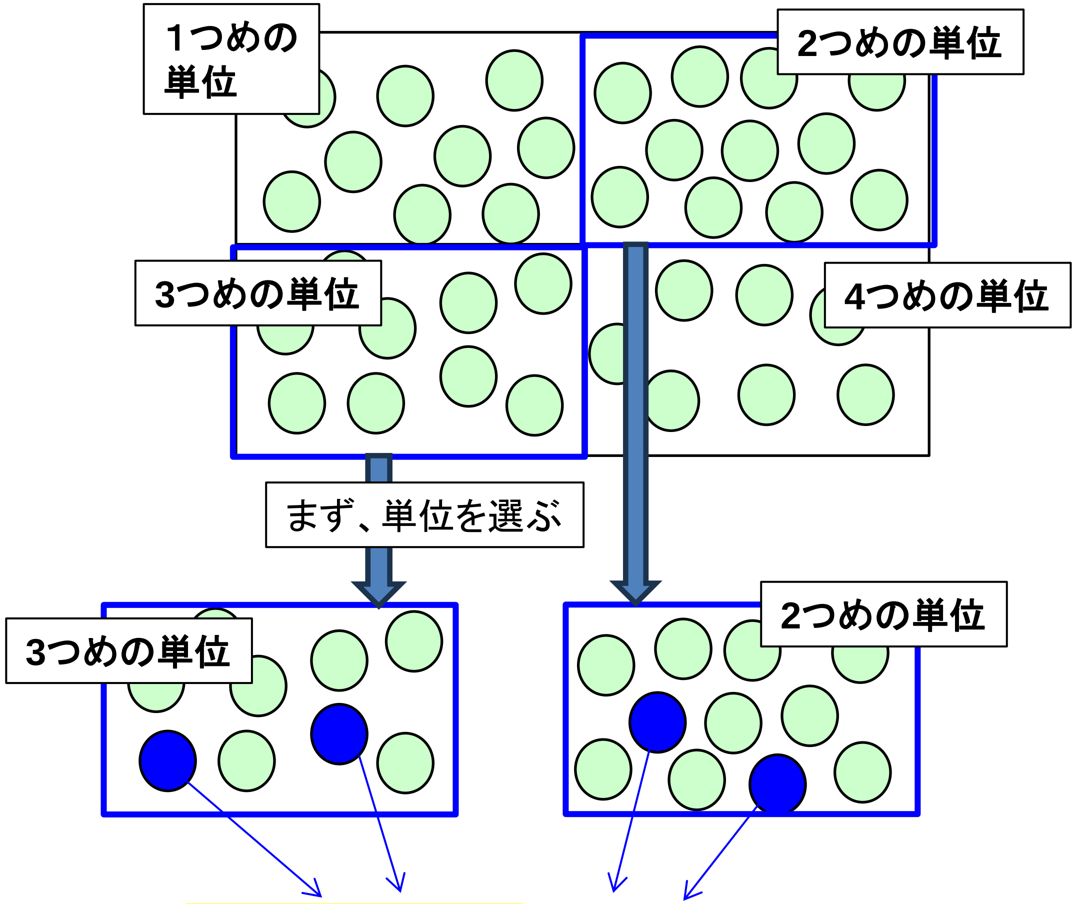


(b)系統抽出法
番号をつけ、同じ間隔で抽出する
(例:3の倍数のものを抽出）

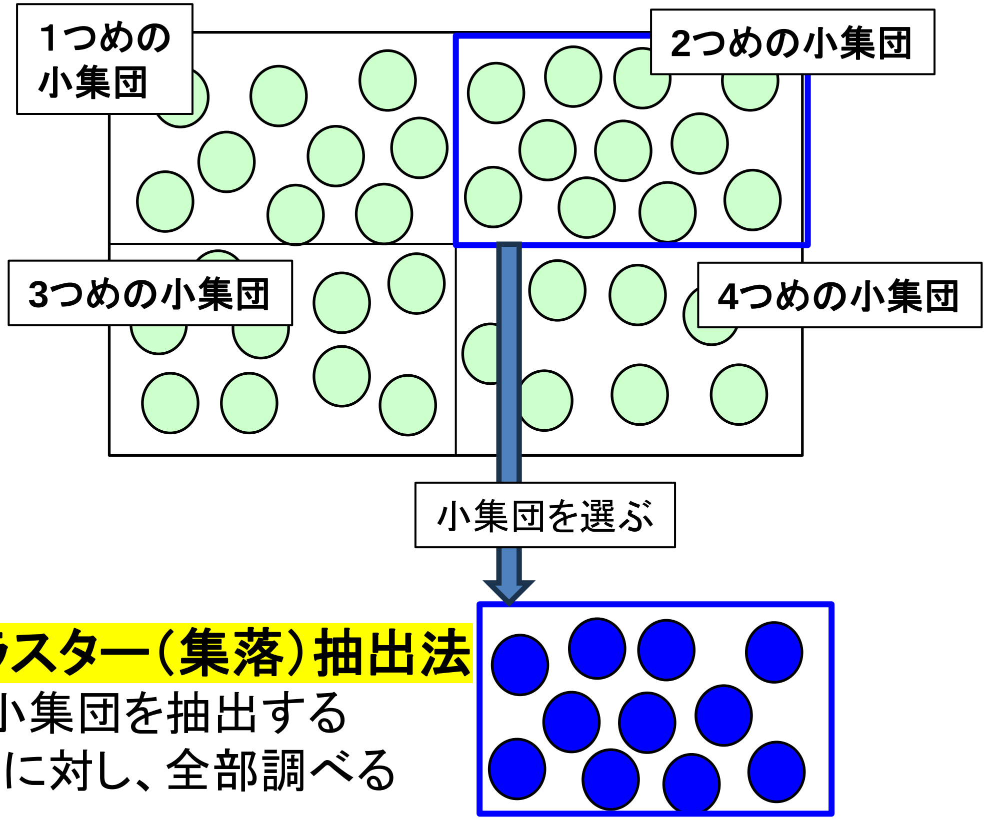


(c)層化無作為抽出法
層毎にランダムに抽出する

(p52.2)[C3]問1. 各標本抽出法の性質



(d)多段抽出法
まず、単位を抽出し、
（そこから小単位を抽出し）
そこから抽出する



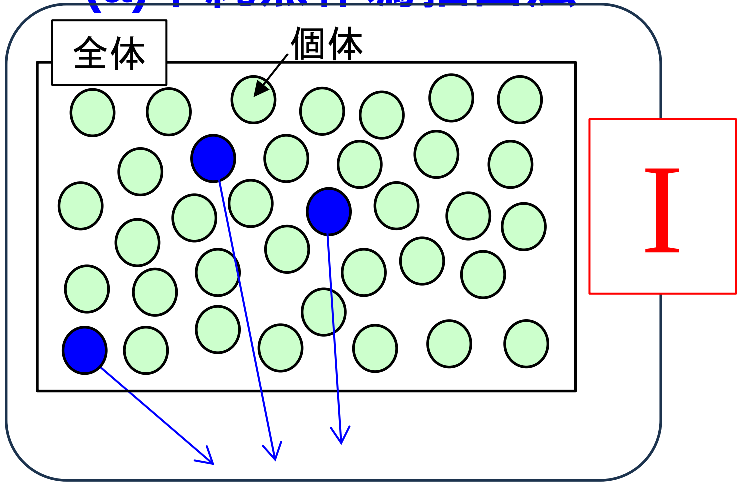
(e)クラスター（集落）抽出法
まず、小集団を抽出する
小集団に対し、全部調べる

(a)単純無作為抽出法
(b)系統抽出法
(c)層化無作為抽出法
(d)多段抽出法
(e)クラスター（集落）抽出法

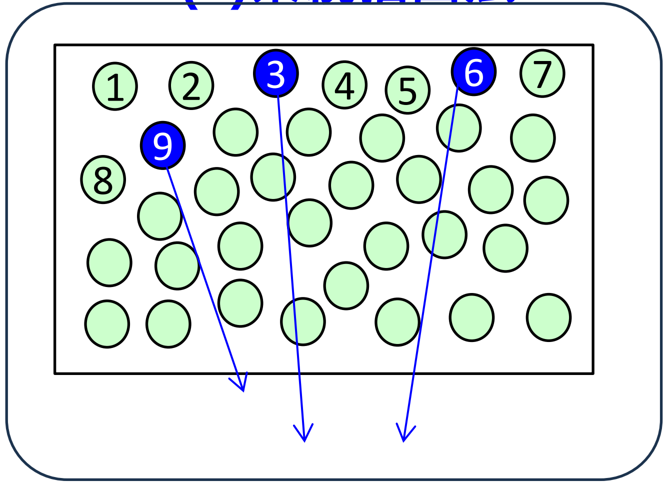
(p52.3)[C3]問1. 各標本抽出法の性質

(ABランク)

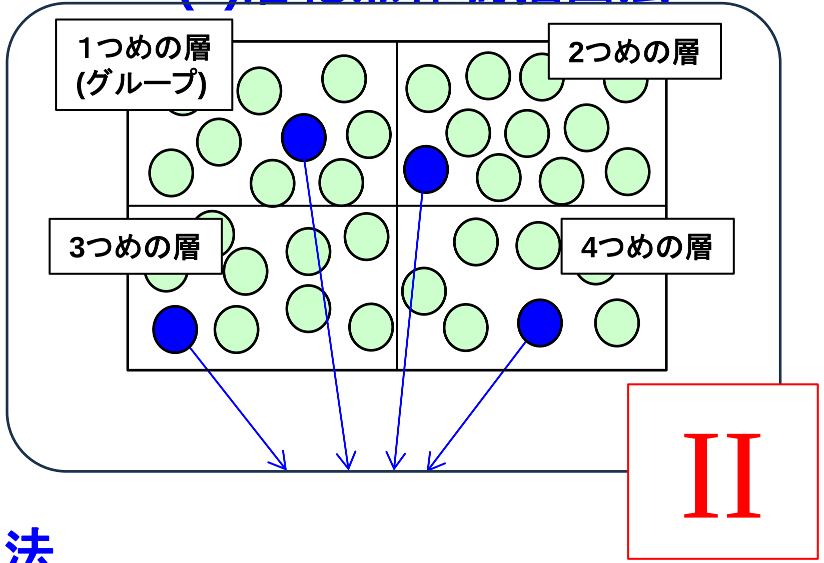
(a)単純無作為抽出法



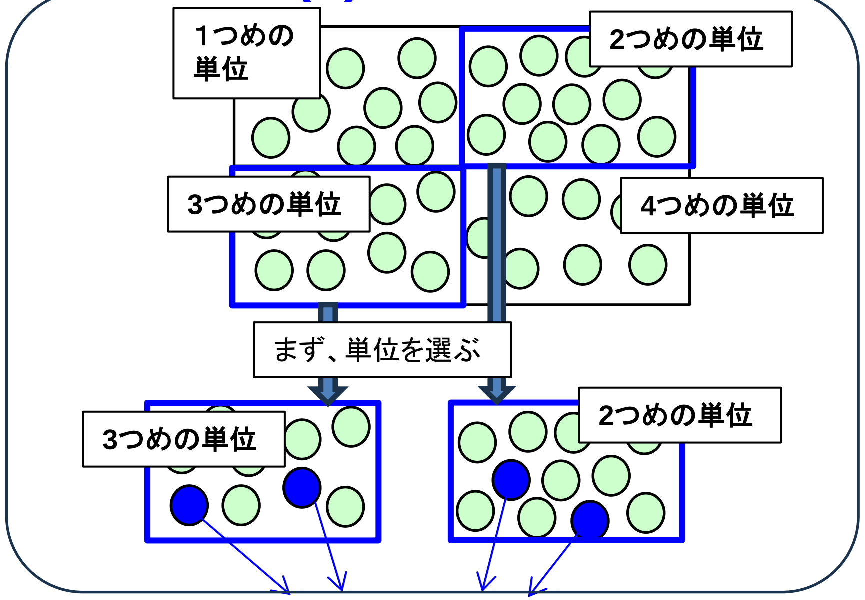
(b)系統抽出法



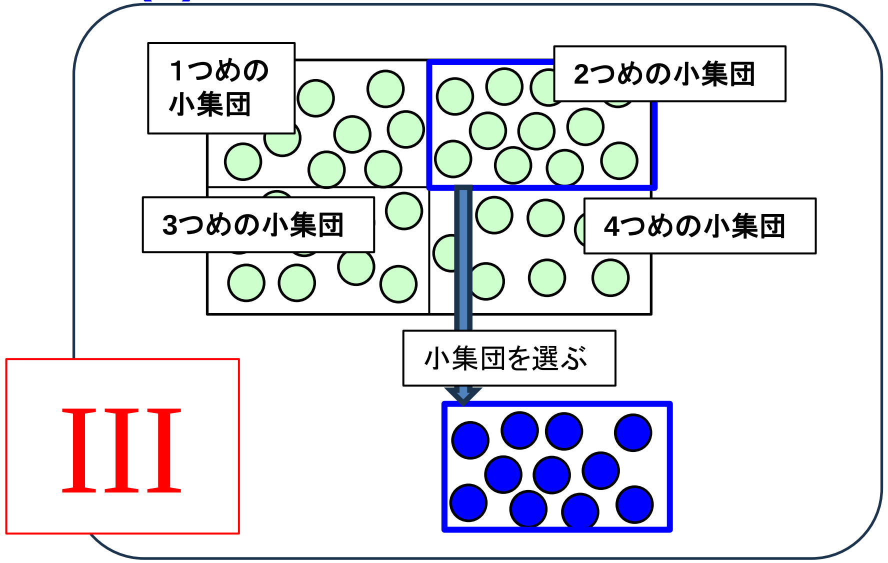
(c)層化無作為抽出法



(d)多段抽出法



(e)クラスター(集落)抽出法



（答）④

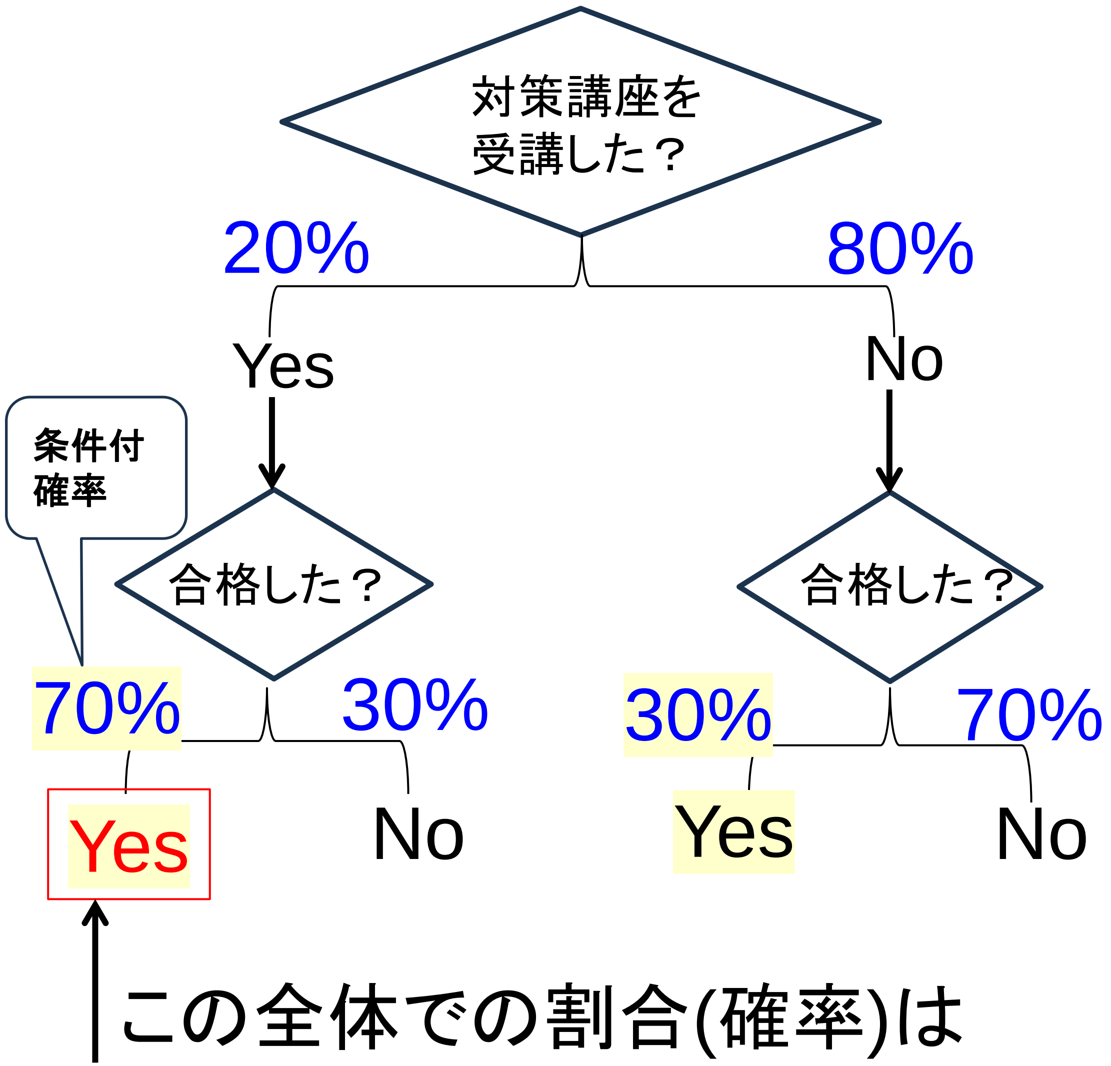
済

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4：確率の分野
問1,2 (p58-59) @お試しサイト

(p58.1)[C4]問1. 積事象の確率

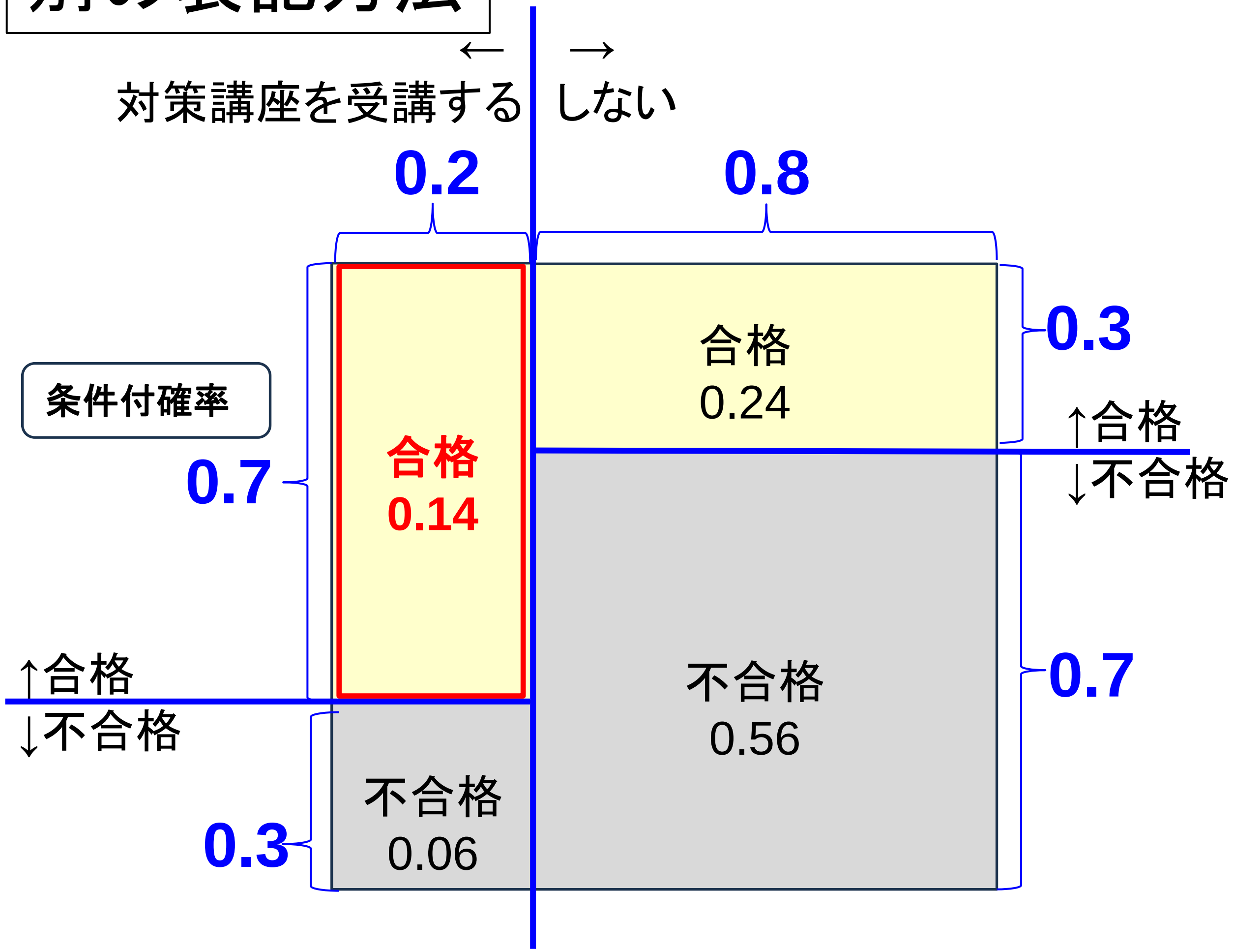
(Aランク)



この全体での割合(確率)は

$$0.2 \times 0.7 = 0.14$$

別の表記方法



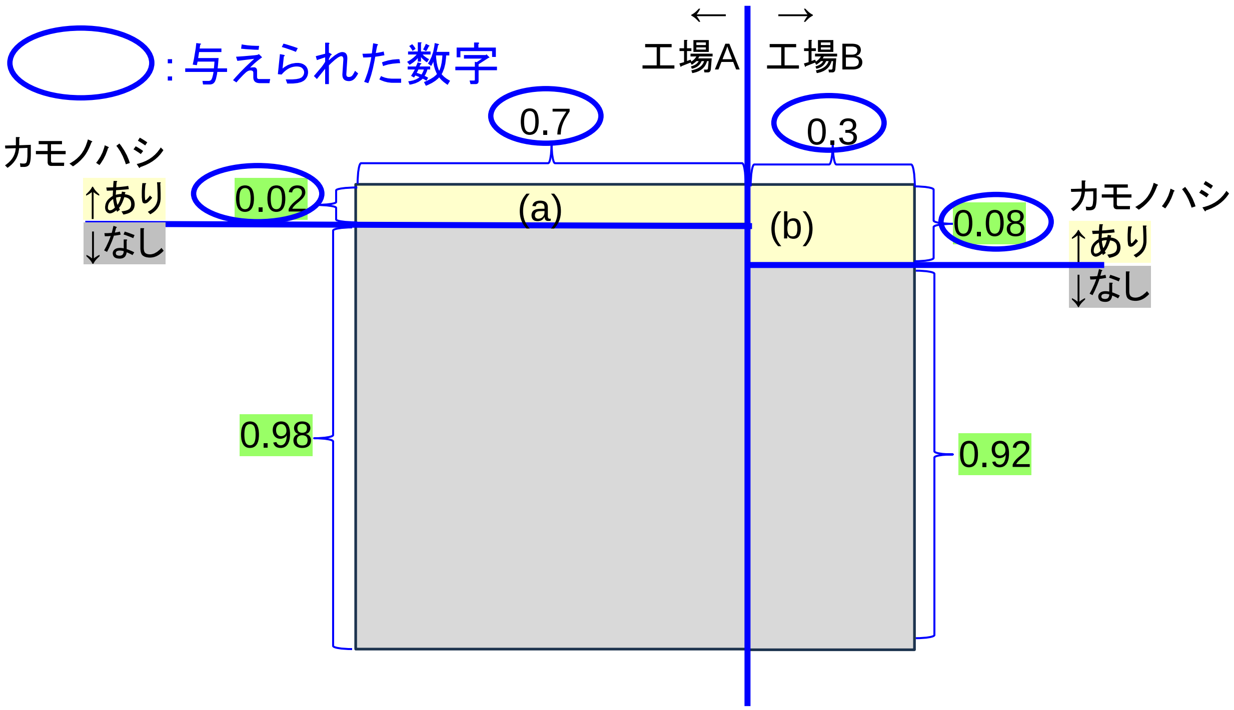
⇒（答）①

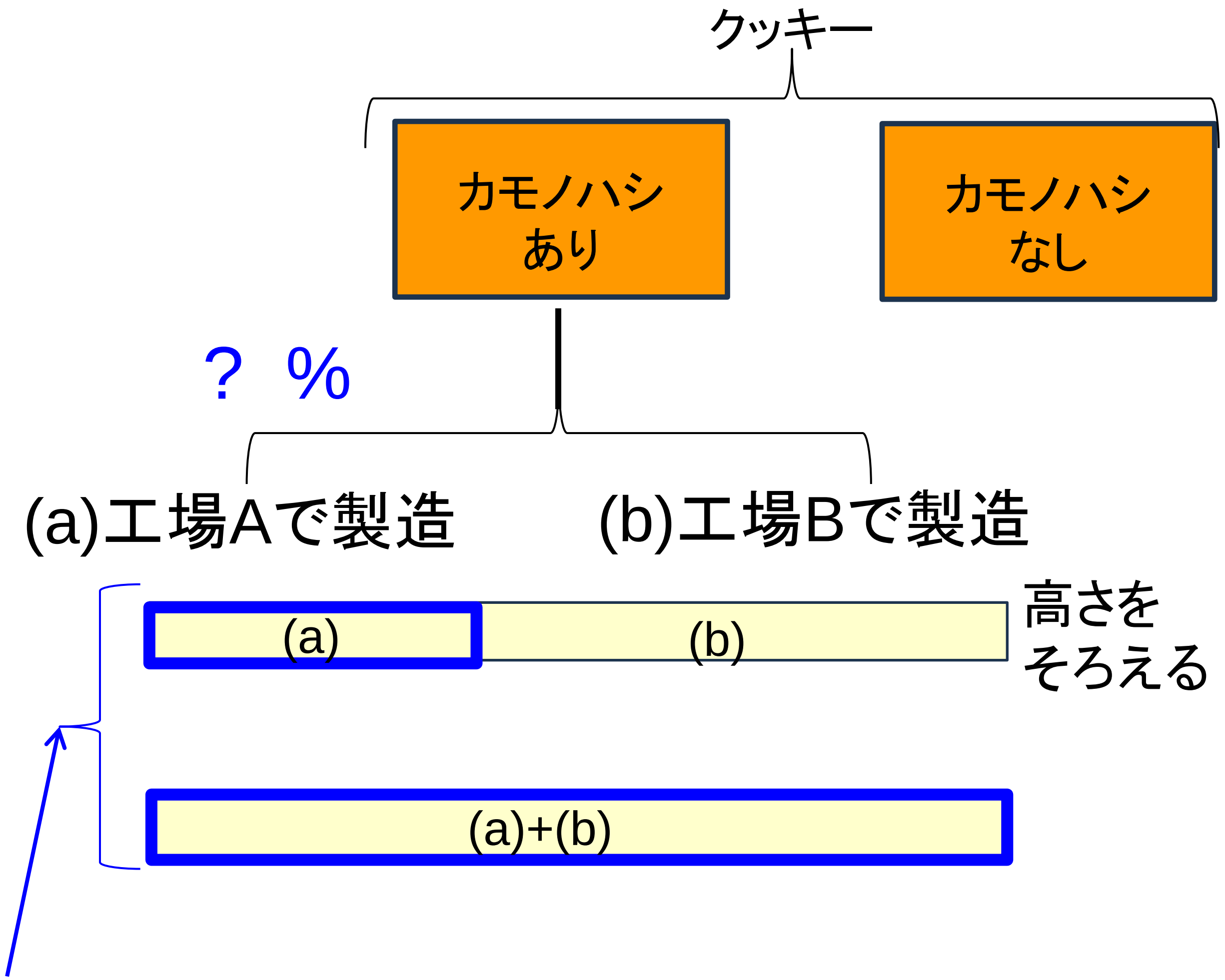
済

(p59.1)[C4]問2. ベイズの定理

(Aランク)

一辺の長さが1の
正方形を考えます





知りたいこと:
カモノハシありのクッキーが
工場Aで製造された割合(確率)

緑色の数字は
条件付確率です

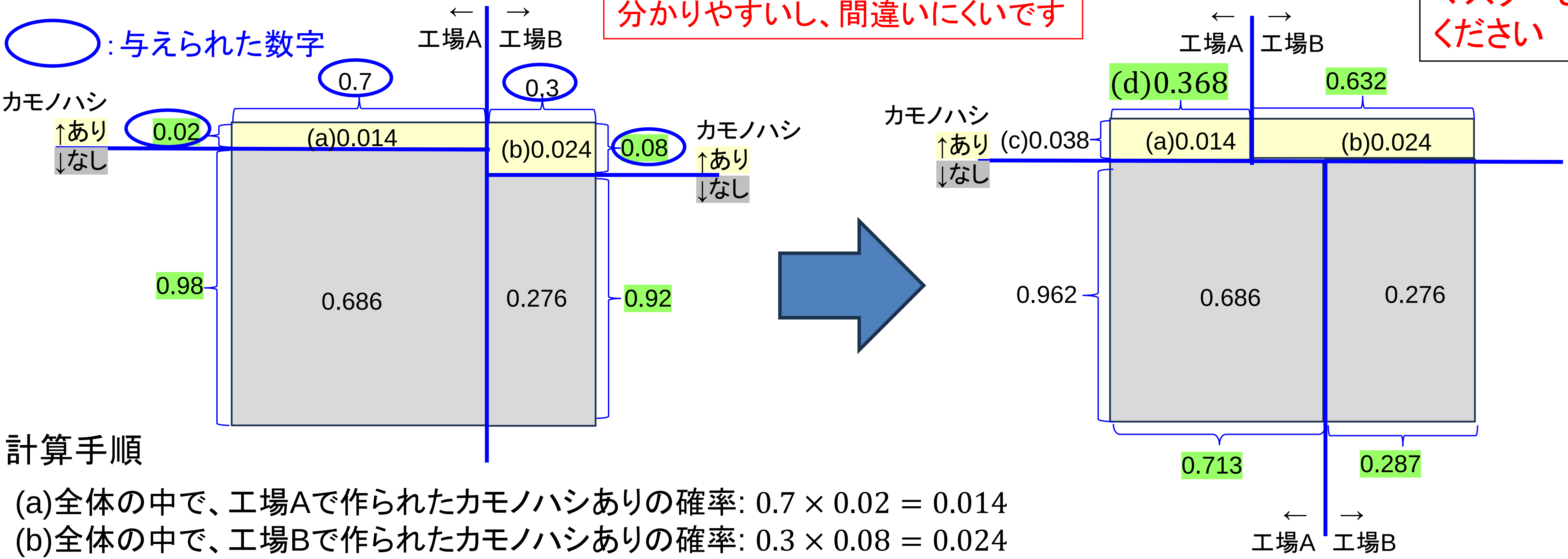
(p59.2)[C4]問2. ベイズの定理

(Aランク)

★よく出ます
マスターしてください

一辺の長さが1の正方形を考えます

「式」よりも「図」で考えた方が
分かりやすいし、間違いにくいです



計算手順

(a)全体の中で、工場Aで作られたカモノハシありの確率: $0.7 \times 0.02 = 0.014$

(b)全体の中で、工場Bで作られたカモノハシありの確率: $0.3 \times 0.08 = 0.024$

(c)全体の中で、カモノハシありの確率: $(c) = (a) + (b)$　$0.014 + 0.024 = 0.038$

(d)カモノハシありの中で、工場Aで作られた確率: $(d)=(a)/(c)$　$0.014/0.038=0.368$

緑色の数字は
条件付確率です

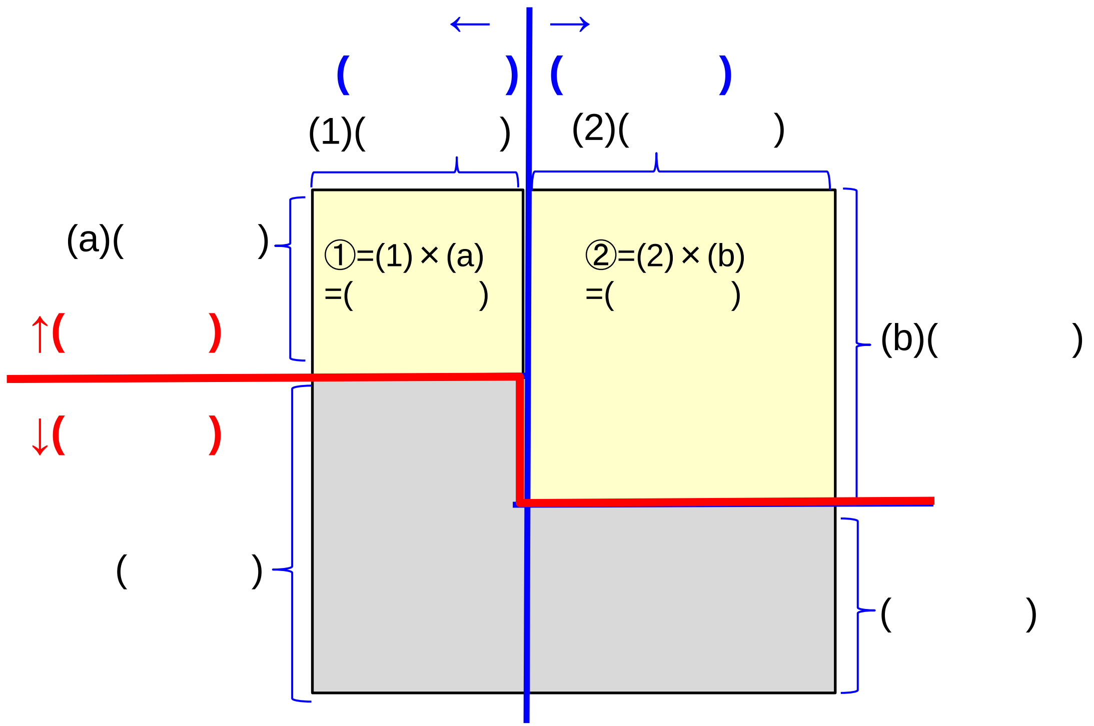
条件付確率を求めるには、上のような作図をして、
該当する辺の長さ(d)を求めるといいです。

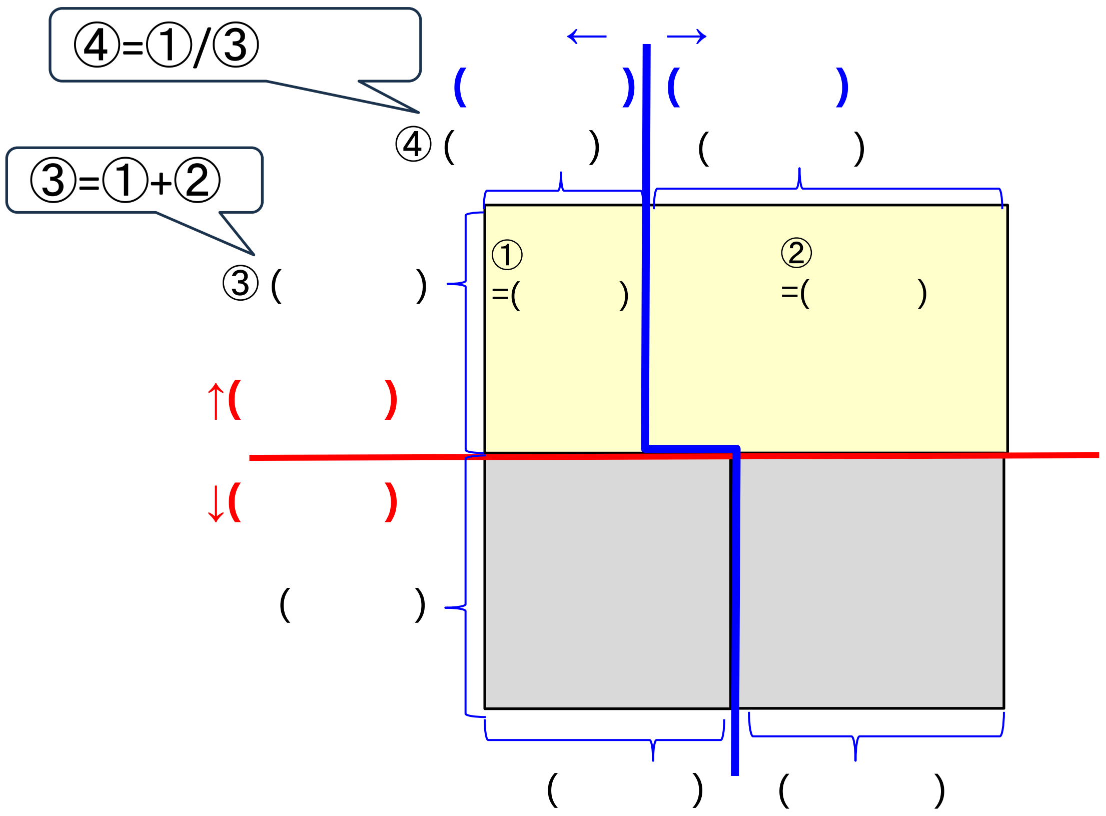
⇒(答)②

済

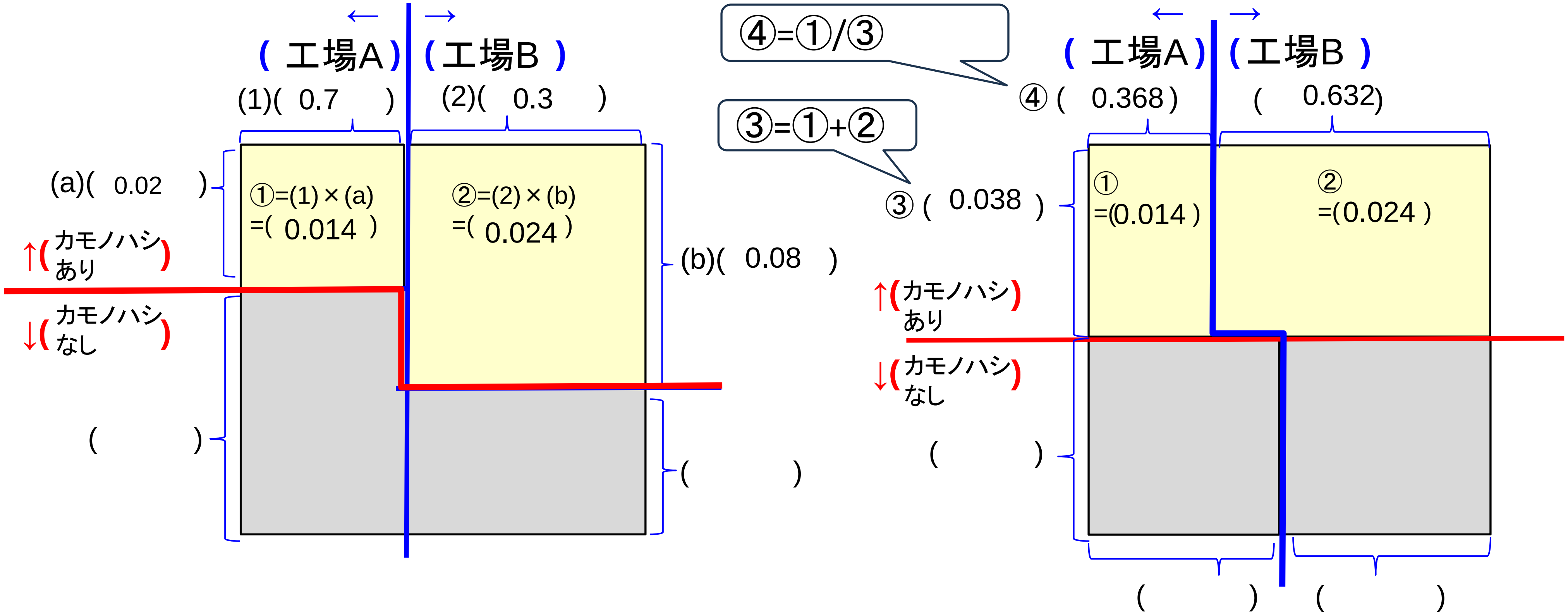
(p59.3)[C4]問2. ベイズの定理の問題用計算シート

もちろん、各数値により図の形・サイズは変わりますが、このシートではそこまで対応不可な点を、ご了承ください。ご注意ください。

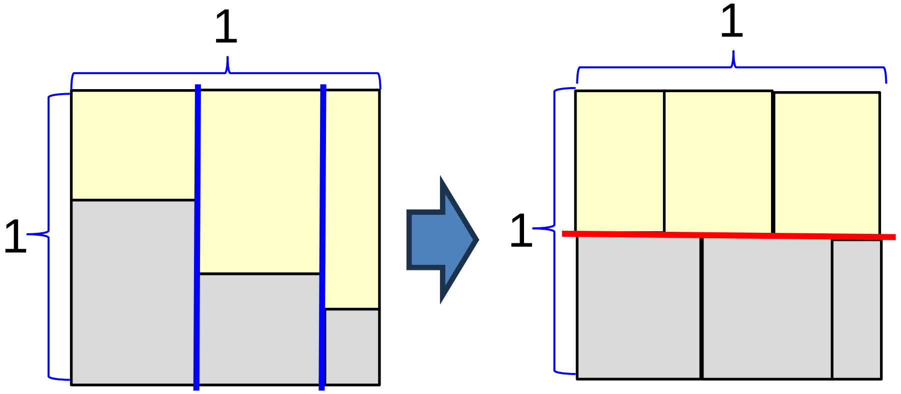




(p59.4)[C4]問2. ベイズの定理の問題用計算シート(例)

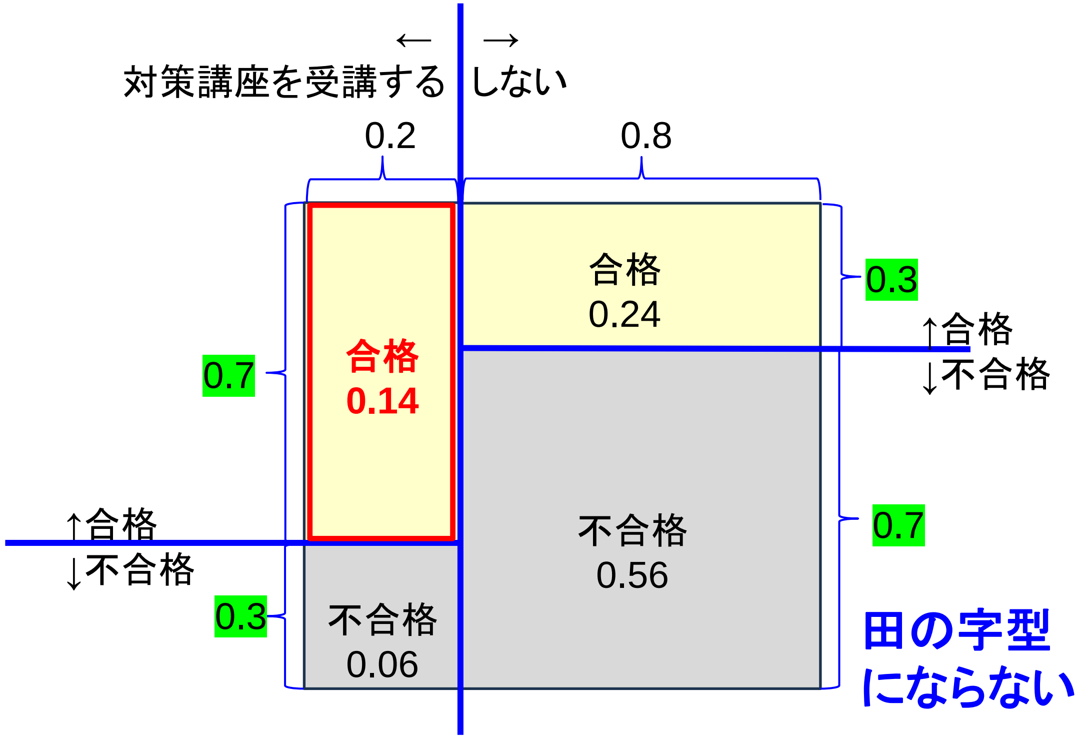


(例)2級過去問2017年6月問7
(漁港X,Y,Z)でとれた貝で(規格内・規格外)の貝に関する問題
もしも、(1),(2)以外に(3),...があったら、
②=(2)×(b)と同様に、②'=(3)×(c)、...も計算し、
③=①+②+②'+...を計算し、④=①/③を求めるといいです。

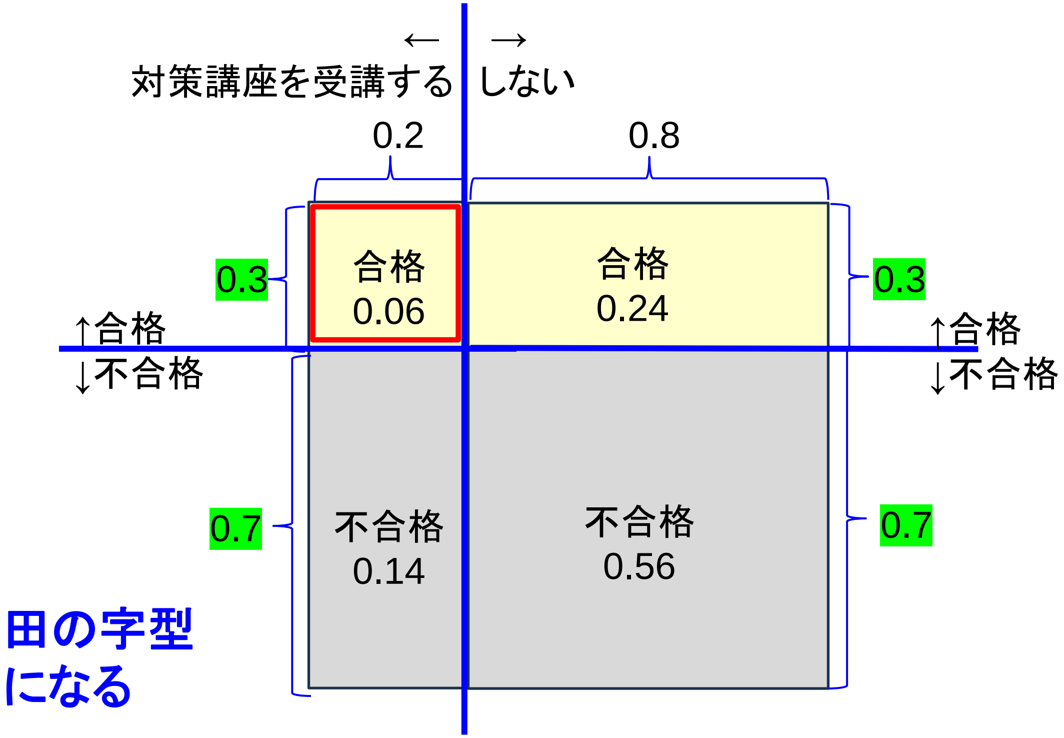


(p59.5) (脱線)「独立」について

A塾の対策講座の受講の有無と
試験の合否の割合(確率)は以下の通りと仮定



B塾の対策講座の受講の有無と
試験の合否の割合(確率)は以下の通りと仮定



A塾の対策講座の受講の有無と
試験の合否には、関係がありますか？

0.14+0.24=0.38

⇒関係ある。受講すると、合格しやすい
「A塾の対策講座の受講の有無」と「合否」は独立でない
P(受講後合格) ≠ P(受講) × P(合格) (0.14 ≠ 0.2 × 0.38=0.076)

B塾の対策講座の受講の有無と
試験の合否には、関係がありますか？

⇒関係ない。受講しても、合格のしやすさは変わらない
「B塾の対策講座の受講の有無」と「合否」は独立である
P(受講後合格) = P(受講) × P(合格) (0.06=0.2 × 0.3)が成立

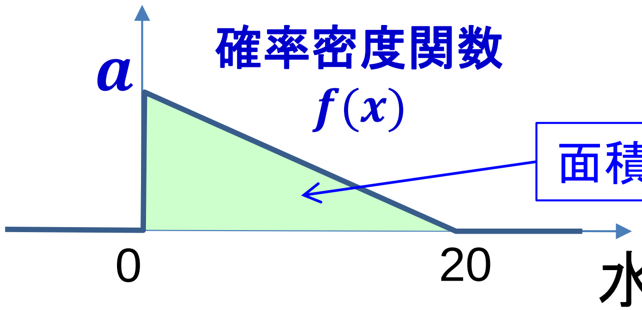
ひかり統計塾

統計検定２級
公式問題集(CBT対応版)の解説
カテゴリー5：確率分布の分野
問1,2（p66-69）@お試しサイト

(p66.1)[C5]問1. 確率分布の定数の決定

(Aランク)

水道使用量の分布・・・連続分布



$$f(x) = a\left(1 - \frac{x}{20}\right) \quad (0 \leq x < 20)$$
$$f(x) = 0 \quad (x < 0, x \geq 20)$$

⇒ a の値は？

$\int_{-\infty}^{\infty} f(x)\,dx = \int_{0}^{20} f(x)dx = 1$を

満たす必要がある

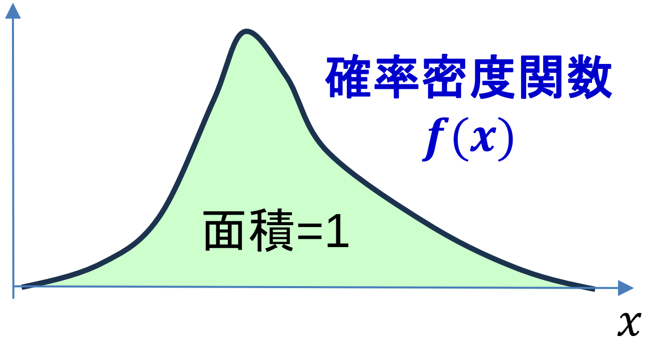
面積=$\frac{1}{2} \times 20 \times a = 10a = 1$

$a = \frac{1}{10} = 0.1$ (答)④

p67の様に、積分を使ってもできますが、
こちらの方がとても簡単です

公式

連続分布



$\int_{-\infty}^{\infty} f(x)\,dx = 1$を満たす必要がある

済

(p68.1)[C5]問2. 正規確率の計算

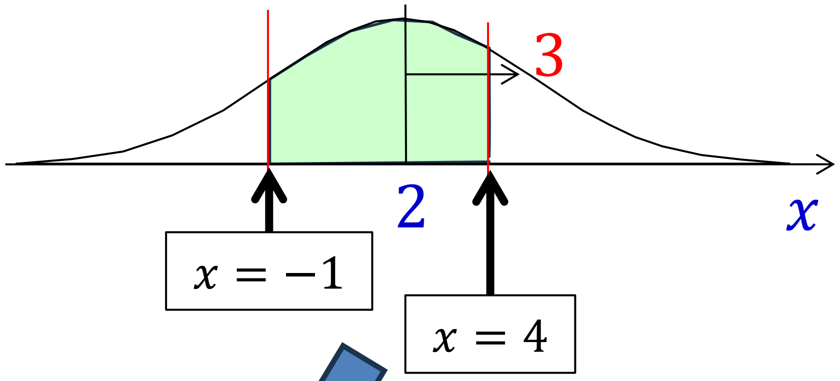
(Aランク)
超基本&超重要

$X \sim N(2, 3^2)$

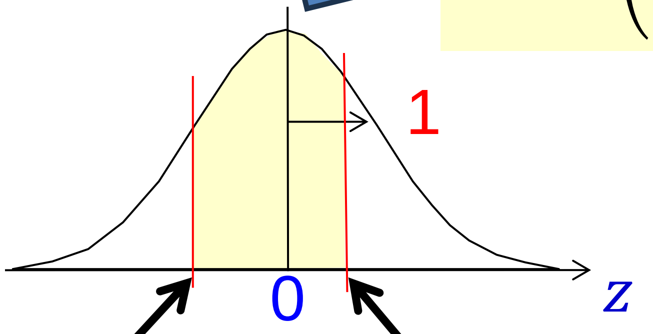
$$Z = \frac{X-2}{3}$$

$Z \sim N(0,1)$

Q: 確率$P(-1 < X \le 4)$はいくら？



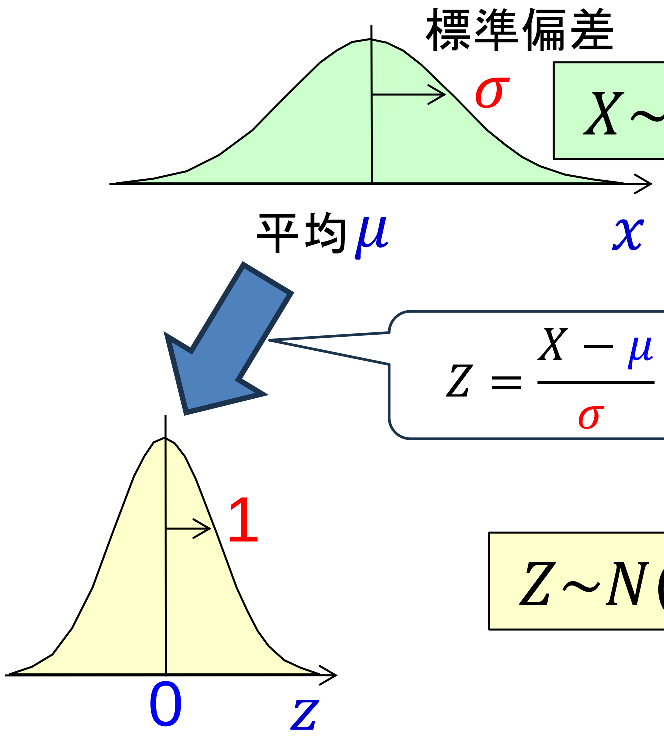
確率$P\left(-1 < Z \le \frac{2}{3}\right)$



$$z = \frac{-1-2}{3} = -1$$

$$z = \frac{4-2}{3} = \frac{2}{3} \fallingdotseq 0.67$$

公式(確率変数の標準化)



$X \sim N(\mu, \sigma^2)$

$$Z = \frac{X-\mu}{\sigma}$$

$Z \sim N(0,1)$

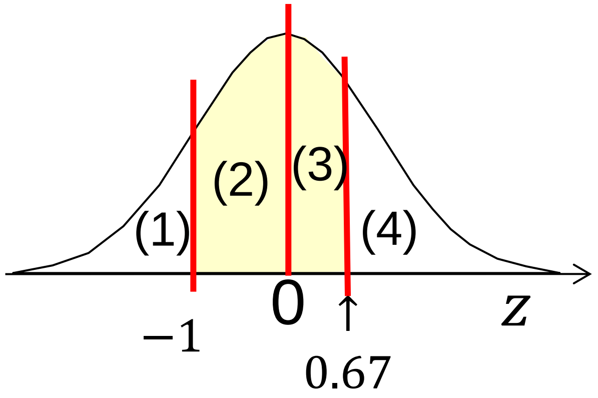
(公式)$X \sim N(\mu, \sigma^2)$の時、
$Z = \frac{X-\mu}{\sigma}$と変換すると、$Z \sim N(0,1)$となる

(p68.2)[C5]問2. 正規確率の計算

(Aランク)
超基本＆超重要

$Z\sim N(0,1)$

確率$P(-1 < Z \leq 0.67)$ =(2)+(3)



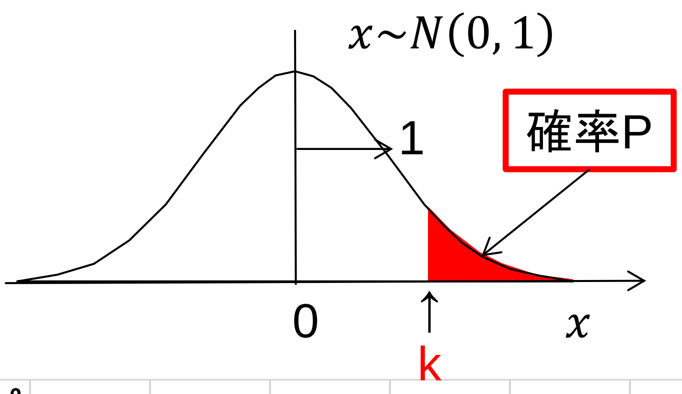
(補足)
連続分布では、
$P(-1 < Z < 0.67)$
$= P(-1 < Z \leq 0.67)$
$= P(-1 < Z \leqq 0.67)$

$P(Z = 0.67) = 0$

右の表で
k=1.00 (k1=1.0, k2=0.00)を見ると (1)=0.1587
k=0.67 (k2=0.6, k2=0.07)を見ると (4)=0.2514
また、(1)+(2)+(3)+(4)=1

従って、
確率$P = (2) + (3) = 1 - ((1) + (4))$
$= 1 - (0.1587 + 0.2514) = 0.590$

(答)⑤



p200 付表1
標準正規分布の
上側確率

k=k1+k2											
	k2=	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
k1=	0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641
	0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
	0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
	0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
	0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
	0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
	0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
	0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
	0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
	0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
	1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
	1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
	1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
	1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
	1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
	1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
	1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
	1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
	1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
	1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
	2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
	2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
	2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	[illegible]
	2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	[illegible]
	2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	[illegible]
	2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	[illegible]

済

ひかり統計塾

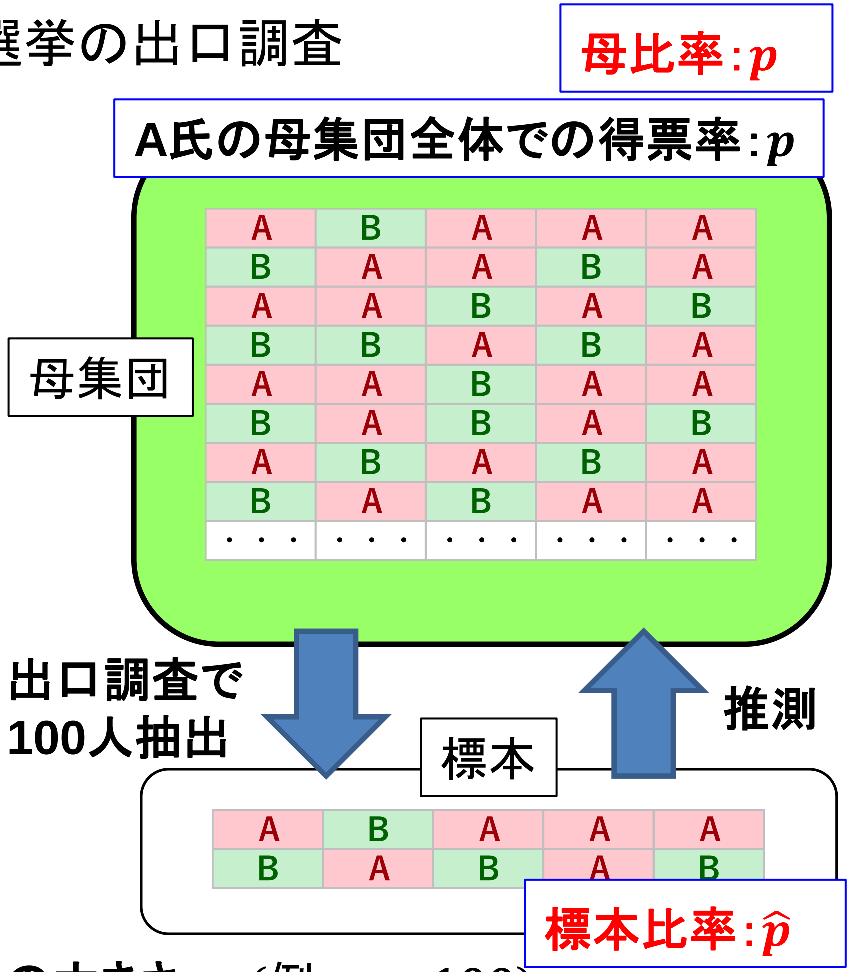
統計検定２級
公式問題集(CBT対応版)の解説
カテゴリー6：標本分布の分野
問1 (p78-79) @お試しサイト

(p78.1) [C6]問1. 標本割合$\hat{p}$の標本分布

(A～Bランク)

関連問題:p72,問4
p74,問5

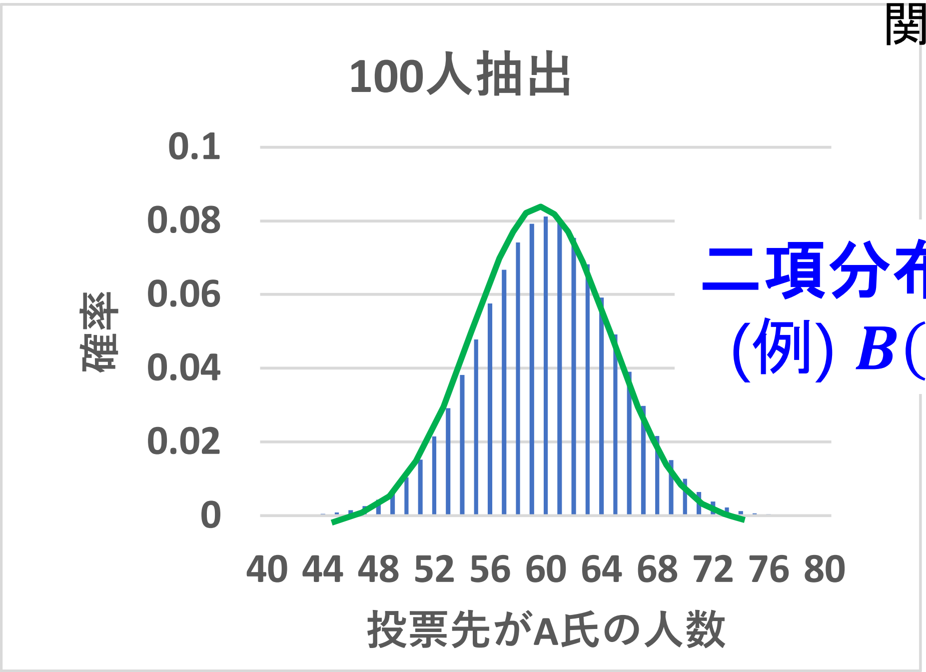
(例)選挙の出口調査



標本の大きさ:n (例: $n = 100$)

A氏に54人が投票　$\hat{p} = \frac{54}{100} = 0.54$

(標本(n人)中、A氏に投票した人数をXとする)



二項分布:$B(n, p)$
(例) $B(100, 0.6)$

A氏に投票した人数X

二項分布:　$X \sim B(n, p)$

正規分布:　$X \sim N(np, np(1-p))$

A氏の標本得票率:

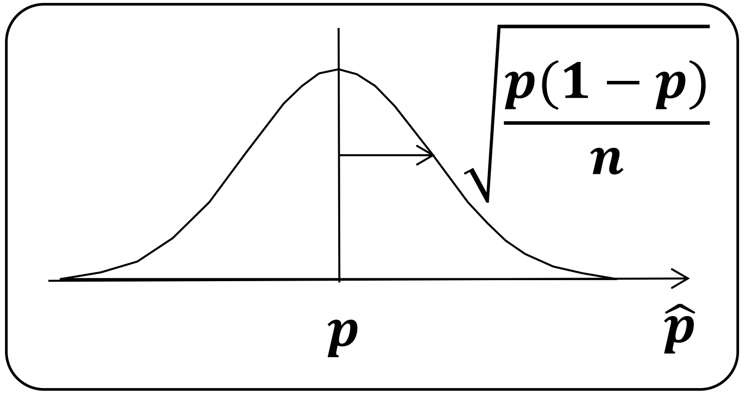
$$\hat{p} = \frac{X}{n} \sim N\left(p, \frac{p(1-p)}{n}\right)$$

(p78.2)[C6]問1. 標本割合$\hat{p}$の標本分布

(A～Bランク)

標本得票率：
$\hat{p}$は正規分布
に従う

$$\hat{p}\sim N\left(p,\frac{p(1-p)}{n}\right)$$



(注)テキストに間違いあり
（初版第3刷まで？）
（誤）$1.96 \leq (ア) \leq 1.96$
ではなく
(正)$-1.96 \leq (ア) \leq 1.96$

(公式)$X\sim N(\mu,\sigma^2)$の時、
$Z=\frac{X-\mu}{\sigma}$と変換すると、
$Z\sim N(0,1)$となる

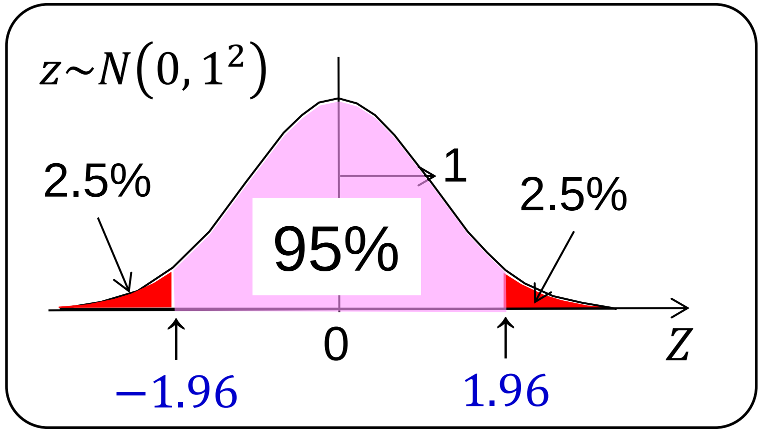
$\hat{p}$を標準化する

$$Z=\frac{\hat{p}-p}{\sqrt{\frac{p(1-p)}{n}}}\sim N(0,1)$$

$$(ア)=\frac{\hat{p}-p}{\sqrt{\frac{p(1-p)}{n}}}$$

注

②または③



95%の確率で、$-1.96 \leq Z \leq 1.96$

$$-1.96 \leq \frac{\hat{p}-p}{\sqrt{\frac{p(1-p)}{n}}} \leq 1.96$$

nが大きくなると、
標本⇒母集団に近づく
推定の精度が良くなる
$p \fallingdotseq \hat{p}$
①(イ)、④(イ)はおかしい

$$\hat{p}-1.96\sqrt{\frac{p(1-p)}{n}} \leq p \leq \hat{p}+1.96\sqrt{\frac{p(1-p)}{n}}$$

$$(イ)=1.96\sqrt{\frac{p(1-p)}{n}}$$

③または⑤

(答)③

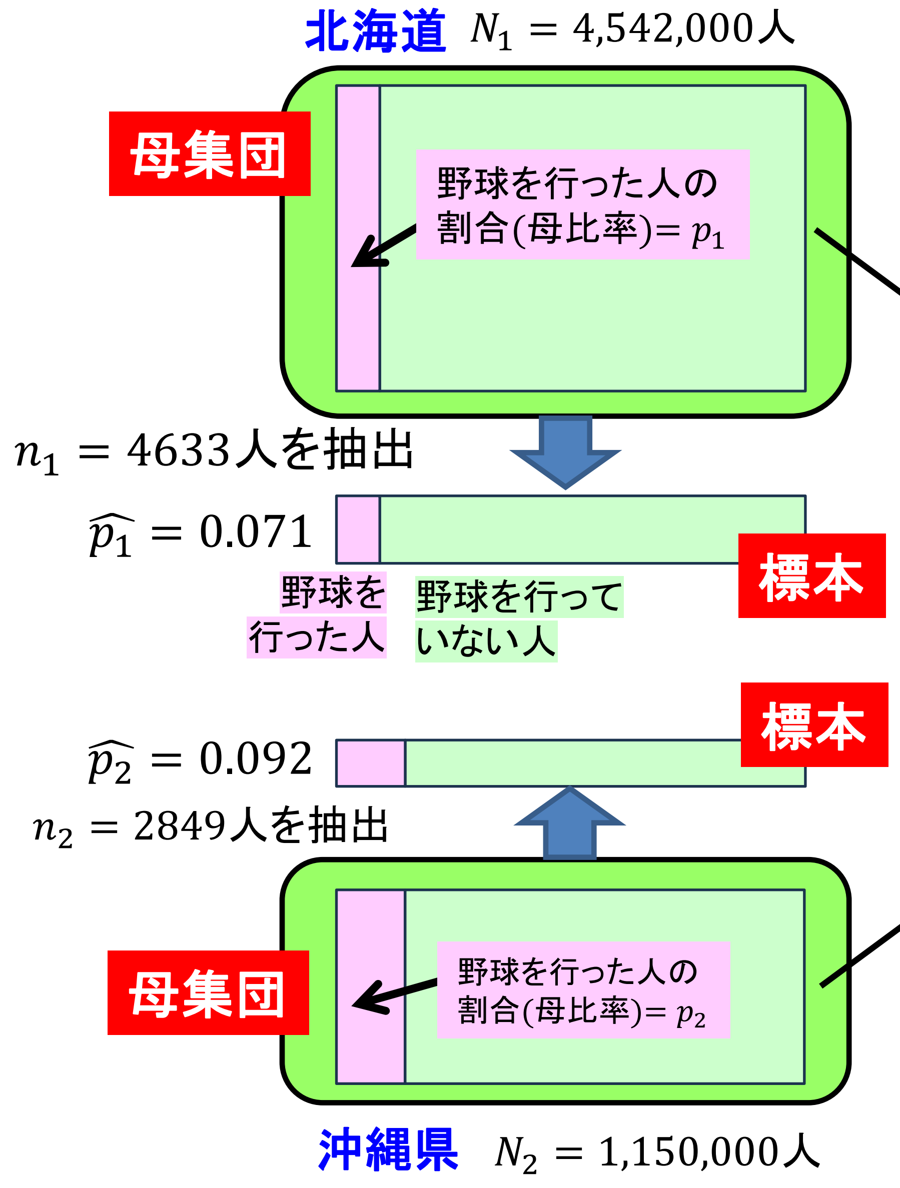
済

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7：推定の分野
問1 (p94-96) @お試しサイト

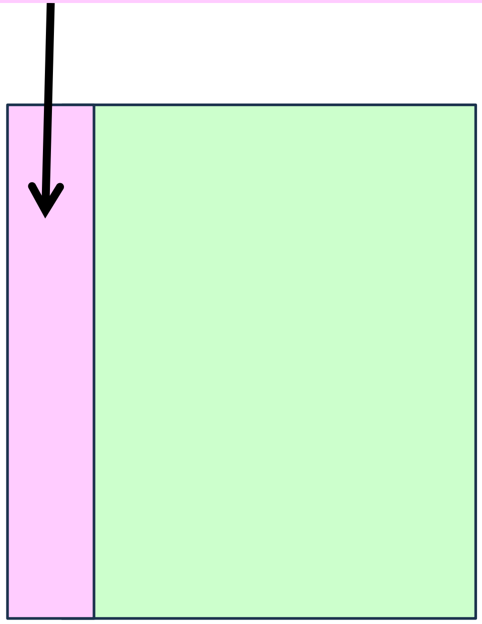
(p94.1)[C7]問1. 推定値と標準誤差

(Bランク)



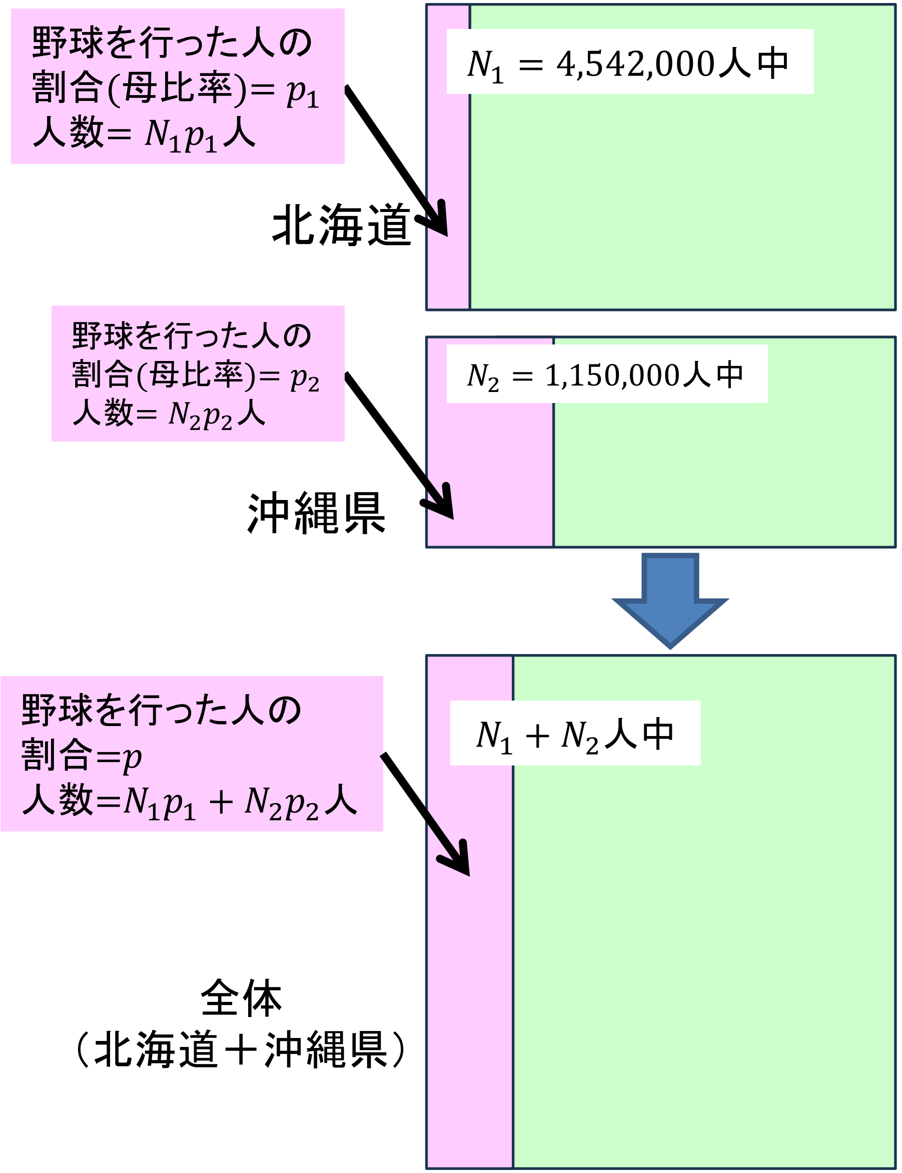
全体(北海道＋沖縄県、15歳以上：約570万人)での
過去1年間に**野球を行った人の**
割合を知りたい

母比率pの推定値、標準誤差を知りたい



(p94.2)[C7]問1. 推定値と標準誤差

(Bランク)



(1)全体(北海道+沖縄県)の人数: $N_1 + N_2$

(2)全体で野球を行っていた人の人数: $N_1p_1 + N_2p_2$

(1)の中での(2)の比率(全体で野球を行った人の割合):

$$p = \frac{N_1p_1 + N_2p_2}{N_1 + N_2}$$

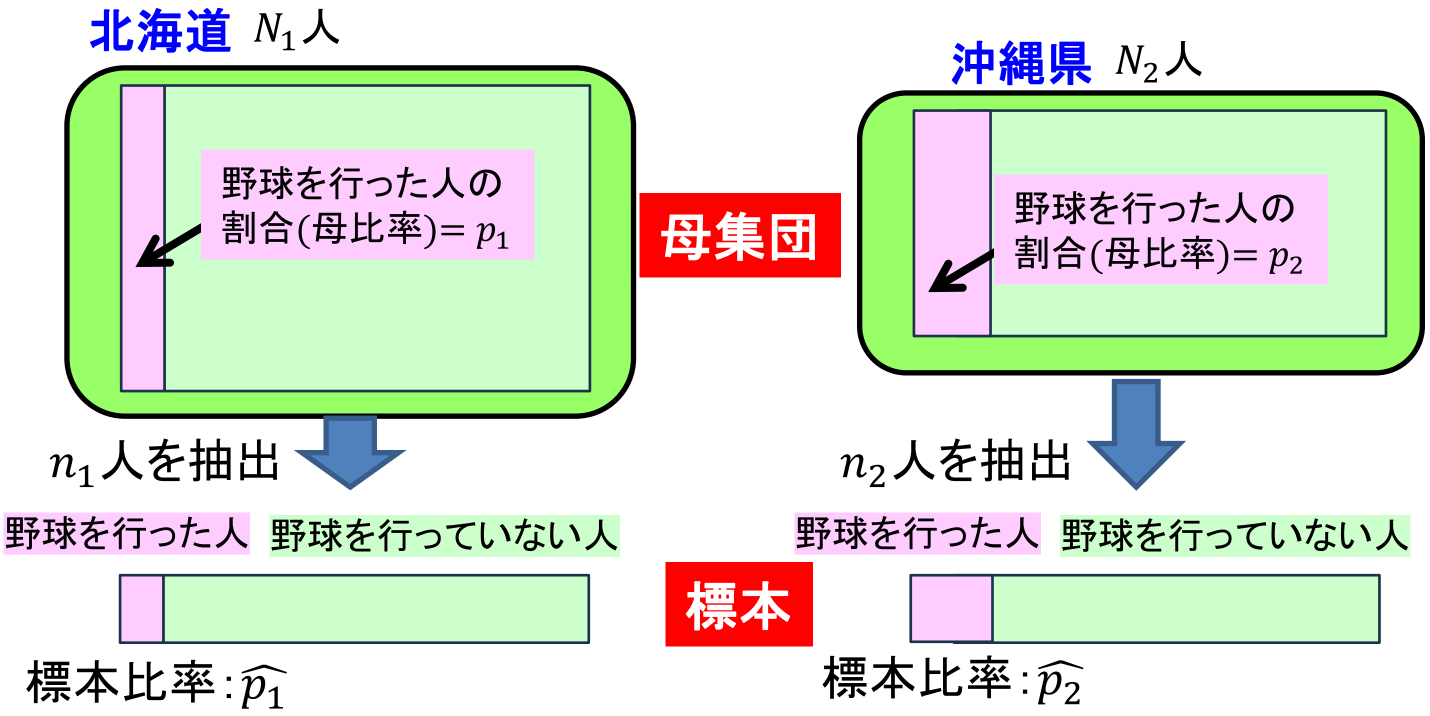
北海道での母比率: p_1の推定値を$\widehat{p_1}$、
沖縄県での母比率: p_2の推定値を$\widehat{p_2}$ とすると、
全体 での母比率: pの 推定値 $\hat{p}$ は、

$$\hat{p} = \frac{N_1\widehat{p_1} + N_2\widehat{p_2}}{N_1 + N_2}$$ に書けます

(答)(ア)の部分の候補: ①②

(p94.3)[C7]問1. 推定値と標準誤差

(Bランク)



$$\widehat{p_1} \sim N\left(p_1, \frac{p_1(1-p_1)}{n_1}\right)$$

$$\widehat{p_2} \sim N\left(p_2, \frac{p_2(1-p_2)}{n_2}\right)$$

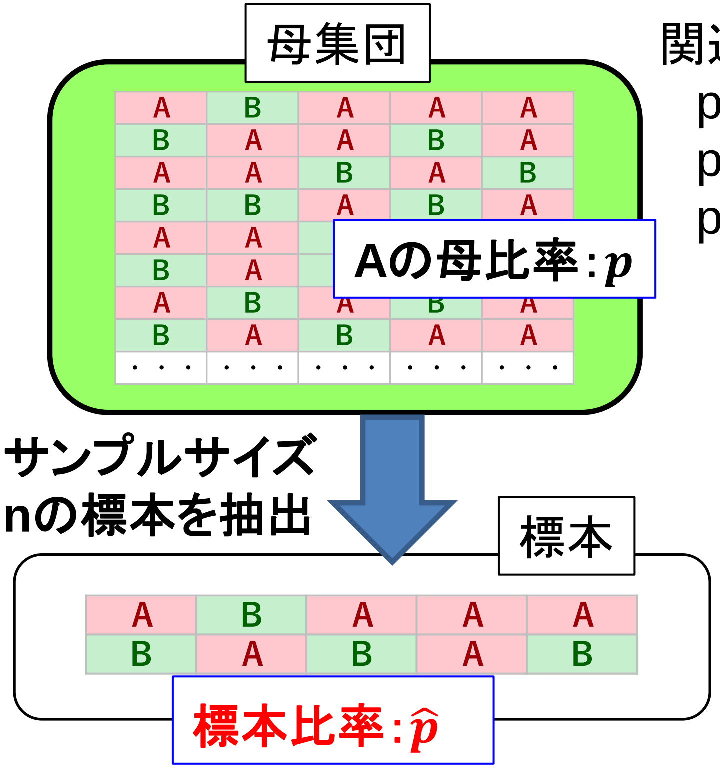
全体での

母比率の推定値: $\hat{p} = \dfrac{N_1\widehat{p_1} + N_2\widehat{p_2}}{N_1 + N_2} = \dfrac{N_1}{N_1 + N_2}\widehat{p_1} + \dfrac{N_2}{N_1 + N_2}\widehat{p_2}$

⇒ $\hat{p}$はどんな分布に従う?

公式:母比率と標本比率の関係

関連問題:
p72,問4
p74,問5
p78,問1



$$\hat{p} \sim N\left(p, \frac{p(1-p)}{n}\right)$$

(p94.4)[C7]問1. 推定値と標準誤差

(Bランク)

北海道と沖縄県全体で、野球をやった人の比率は、

$N_1=4{,}542{,}000$人

$N_2=1{,}150{,}000$人

$$\hat{p}=\frac{N_1}{N_1+N_2}\widehat{p_1}+\frac{N_2}{N_1+N_2}\widehat{p_2}=A\widehat{p_1}+B\widehat{p_2}$$

但し、$A=\frac{N_1}{N_1+N_2}, B=\frac{N_2}{N_1+N_2}$ $(A,B:$ 定数$)$

$$\widehat{p_1}\sim N\left(p_1,\frac{p_1(1-p_1)}{n_1}\right)\qquad \widehat{p_2}\sim N\left(p_2,\frac{p_2(1-p_2)}{n_2}\right)$$

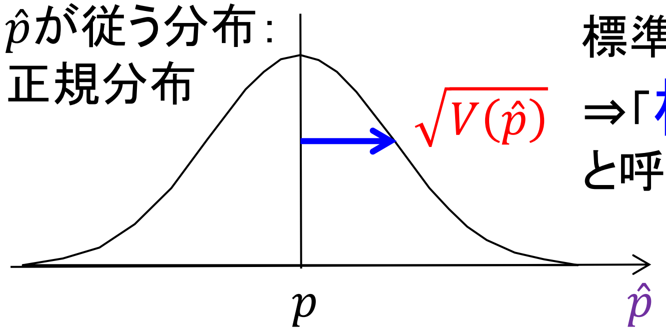
分散：$V(\widehat{p_1})=\frac{p_1(1-p_1)}{n_1}$　分散：$V(\widehat{p_2})=\frac{p_2(1-p_2)}{n_2}$

(公式)　確率変数X,Yが独立な時、
$V(X+Y)=V(X)+V(Y)$
$V(aX)=a^2V(X)$ $(a:$ 定数$)$

分散：$V(\hat{p})=V(A\widehat{p_1}+B\widehat{p_2})=V(A\widehat{p_1})+V(B\widehat{p_2})=A^2V(\widehat{p_1})+B^2V(\widehat{p_2})$

$$=\left(\frac{N_1}{N_1+N_2}\right)^2V(\widehat{p_1})+\left(\frac{N_2}{N_1+N_2}\right)^2V(\widehat{p_2})=\left(\frac{N_1}{N_1+N_2}\right)^2\frac{p_1(1-p_1)}{n_1}+\left(\frac{N_2}{N_1+N_2}\right)^2\frac{p_2(1-p_2)}{n_2}$$

$\hat{p}$が従う分布：
正規分布



標準偏差：
⇒「標準誤差」
と呼びます

$$\text{標準誤差}=\sqrt{V(\hat{p})}=\sqrt{\left(\frac{N_1}{N_1+N_2}\right)^2\frac{\widehat{p_1}(1-\widehat{p_1})}{n_1}+\left(\frac{N_2}{N_1+N_2}\right)^2\frac{\widehat{p_2}(1-\widehat{p_2})}{n_2}}$$

(ア)の候補：①②

未知のp_1,p_2の代わりに$\widehat{p_1},\widehat{p_2}$を使った。　(イ)の候補：②　⇒(答)②

済

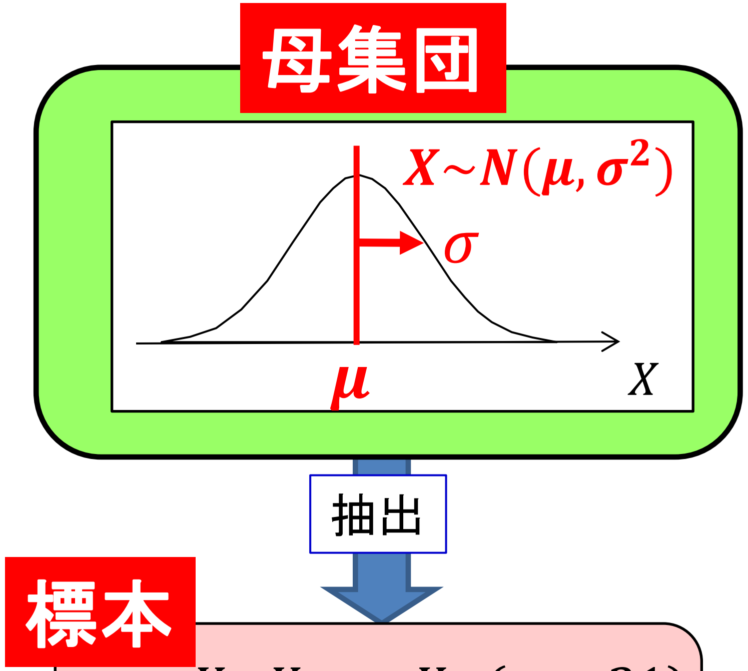
ひかり統計塾

統計検定２級
公式問題集(CBT対応版)の解説
カテゴリー8: 検定の分野
問1（p106-107）@お試しサイト

(p106.1)[C8]問1. 母平均の検定の考え方

(Aランク)

ある金融資産の日次収益率(%)



$X_1, X_2, \ldots, X_n$ $(n = 21)$

⇒標本平均$\bar{X}$

(不偏分散$V = \widehat{\sigma^2}$)

例

0.16	0.03	-0.49	0.47	0.43	-0.42	0.49
-0.29	0.29	0.50	0.29	-0.31	0.50	0.14
0.69	-0.19	0.35	-0.19	0.37	-0.06	-0.03

⇒ $\mu = 0$ と言える？

公式

$X_i \sim \boxed{N(\mu, \sigma^2)}$ の時、

標本平均 $\bar{X} \sim \boxed{N\left(\mu, \dfrac{\sigma^2}{n}\right)}$

($X_1, X_2, \ldots, X_n$の平均)

(公式)$X \sim N(\mu, \sigma^2)$の時、
$Z = \dfrac{X-\mu}{\sigma}$と変換すると、
$Z \sim N(0,1)$となる

$$Z = \frac{\bar{X} - \mu}{\sqrt{\frac{\sigma^2}{n}}} \sim N(0, 1)$$

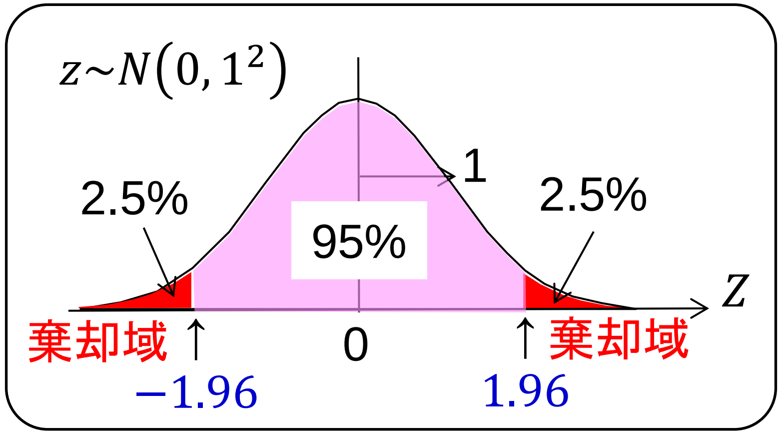
帰無仮説H_0: $\mu = \mu_0$ $(\mu_0 = 0)$

対立仮説H_1: $\mu \neq \mu_0$ (両側検定)

H_0が正しいと仮定する

また、問題では、n=21なので、

⇒検定統計量: $Z = \dfrac{\bar{X} - 0}{\sqrt{\frac{\sigma^2}{n}}} = \dfrac{\bar{X}}{\sqrt{\frac{\sigma^2}{21}}}$



(ア) 棄却域は、$|Z| > 1.96$

(p106.2)[C8]問1. 母平均の検定の考え方

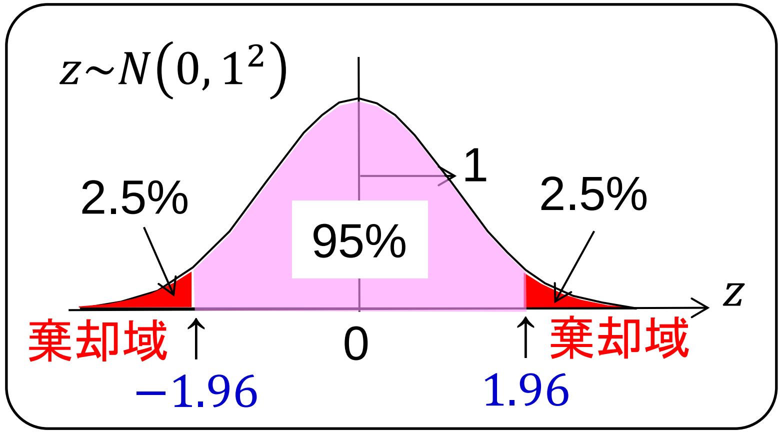
(Aランク)

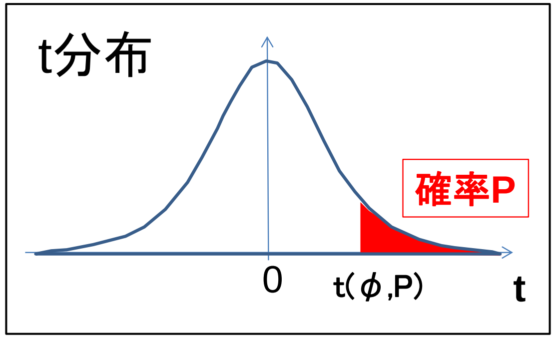
母分散σ^2が既知の時

$$Z=\frac{\bar{X}-\mu}{\sqrt{\frac{\sigma^2}{n}}}\sim N(0,1)$$

検定統計量: $Z=\dfrac{\bar{X}}{\sqrt{\frac{\sigma^2}{21}}}$

済





母分散σ^2が未知の時

不偏分散: $\widehat{\sigma^2}=\dfrac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2$

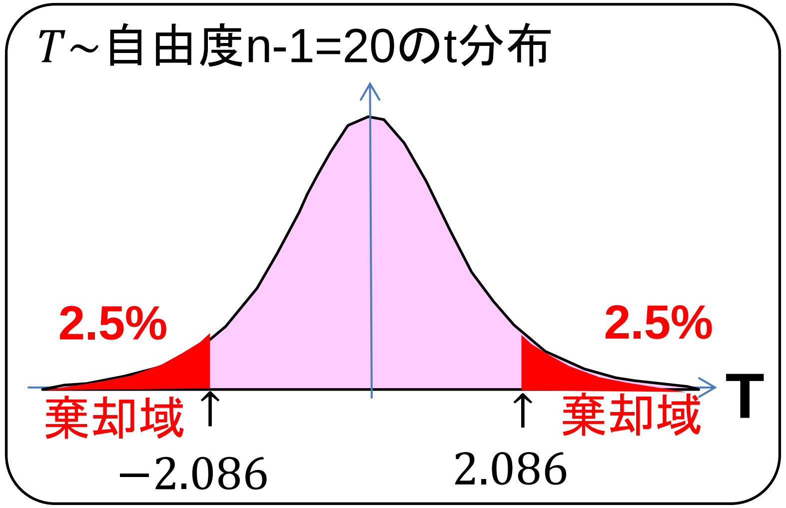
片側P=	0.250	0.200	0.05	0.025	0.01
(両側2P=)	0.500	0.400	0.10	0.05	0.02
ϕ= 1	1.000	1.376	6.314	12.706	31.821
2	0.816	1.061	2.920	4.303	6.965
3	0.765	0.978	2.353	3.182	4.541
4	0.741	0.941	2.132	2.776	3.747
5	0.727	0.920	2.015	2.571	3.365
6	0.718	0.906	1.943	2.447	3.143
7	0.711	0.896	1.895	2.365	2.998
8	0.706	0.889	1.860	2.306	2.896
18	0.688	0.862	1.734	2.101	2.552
19	0.688	0.861	1.729	2.093	2.539
20	0.687	0.860	1.725	2.086	2.528
30	0.683	0.854	1.697	2.042	2.457
40	0.681	0.851	1.684	2.021	2.423

$$T=\frac{\bar{X}-\mu}{\sqrt{\frac{\widehat{\sigma^2}}{n}}}\sim\text{自由度n-1のt分布}$$

n=21なのでn-1=20

(帰無)仮説: $\mu=\mu_0\ \ (\mu_0=0)$
が正しいと仮定する

検定統計量: $T=\dfrac{\bar{X}}{\sqrt{\frac{\widehat{\sigma^2}}{21}}}$



(イ) $|T|>2.086$

(p106.3)[C8]問1. 母平均の検定の考え方

(Aランク)

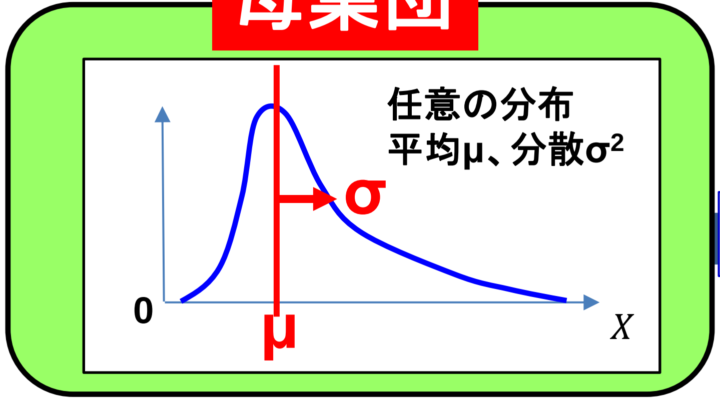
これまで、X_iは正規分布に従うと仮定していた

p102.問5の解説を参照してください

X_iが正規分布に従わない場合は･･･

⇒中心極限定理を使います

母集団



抽出

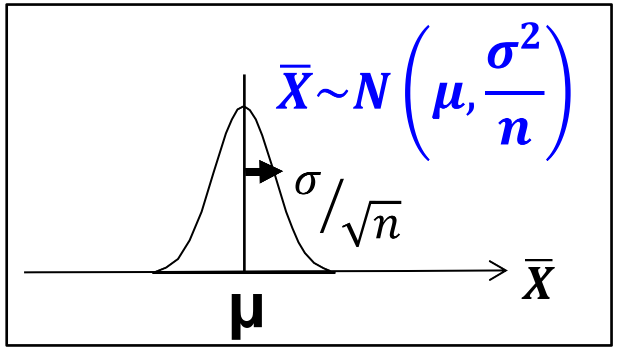
標本

$X_1, X_2, \ldots, X_n$

⇒標本平均$\bar{X}$

⇒不偏分散$\widehat{\sigma^2}$

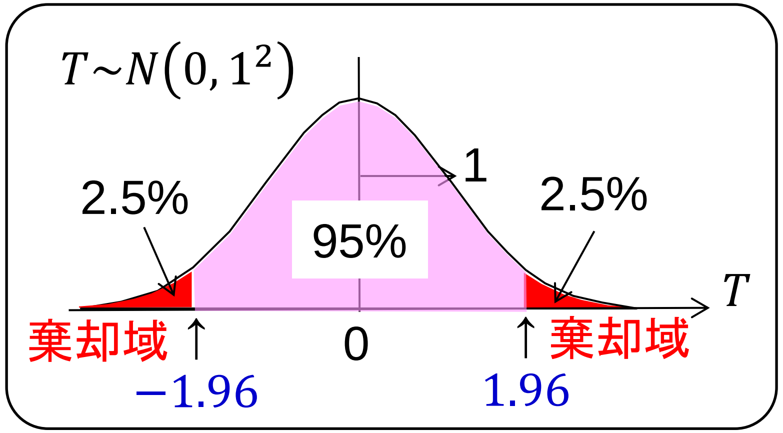
nが十分に大きい時、
$\bar{X}$は近似的に正規分布に従う



$$\Rightarrow Z = \frac{\bar{X} - \mu}{\sqrt{\sigma^2/n}} \sim N(0,1)$$

$$\Rightarrow T = \frac{\bar{X} - \mu}{\sqrt{\widehat{\sigma^2}/n}} \sim N(0,1)$$

(ウ) $|T| > 1.96$



⇒(ア) $|Z| > 1.96$, (イ)$|T| > 2.086$、
(ウ) $|T| > 1.96$ (答)①

済

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8: 検定の分野
問10 (p123-125) @お試しサイト
第１種･第２種の過誤、検出力の基本

(p123.2a) [C8]問10.(準備)第1種、第２種の過誤、検出力

(例)動物(犬または猫)が写った写真から、「犬」と「猫」を判別するシステムがあります。

写真に写った動物(犬または猫)のパーツ・特徴(耳、目、ひげ、前脚、胴体、後脚、尻尾、毛並み、大きさ、色など)から、「猫度指数」を算出します。猫の「猫度指数」は大きく、犬の「猫度指数」は小さいです。

帰無仮説(H_0): 写真の動物は犬である

対立仮説(H_1): 写真の動物は猫である　　を考えます

「猫」、「犬」の写真に基づく「猫度指数」の分布は
右の通り(連続型一様分布)だったと仮定します。

帰無仮説(H_0)の棄却域を

$x \geq 0.6$　とした時に、

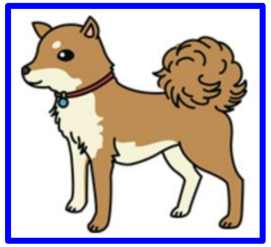
(1)犬を猫と誤判定してしまう確率は？　⇔第1種の過誤の確率

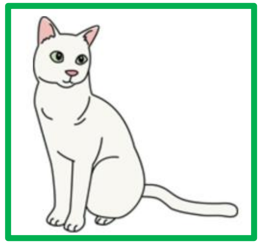
(2)猫を犬と誤判定してしまう確率は？　⇔第2種の過誤の確率

(3)猫を猫と正しく判定する確率は？　⇔検出力

$$\int_0^1 f_D(x)dx = 1$$

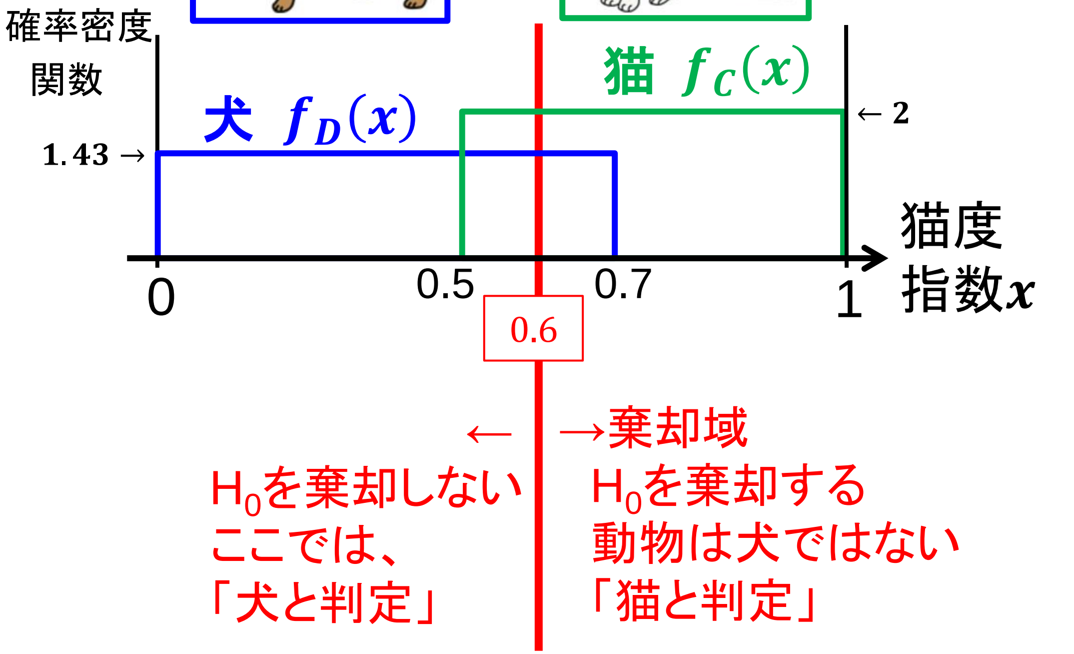
$$1.43 \times 0.7 = 1$$





$$\int_0^1 f_C(x)dx = 1$$

$$2 \times 0.5 = 1$$



(p123.2b) [C8]問10.(準備)第1種、第２種の過誤、検出力

帰無仮説(H_0):
写真の動物は犬である

対立仮説(H_1):
写真の動物は猫である

(1)犬を猫と誤判定してしまう確率は？

第1種の過誤

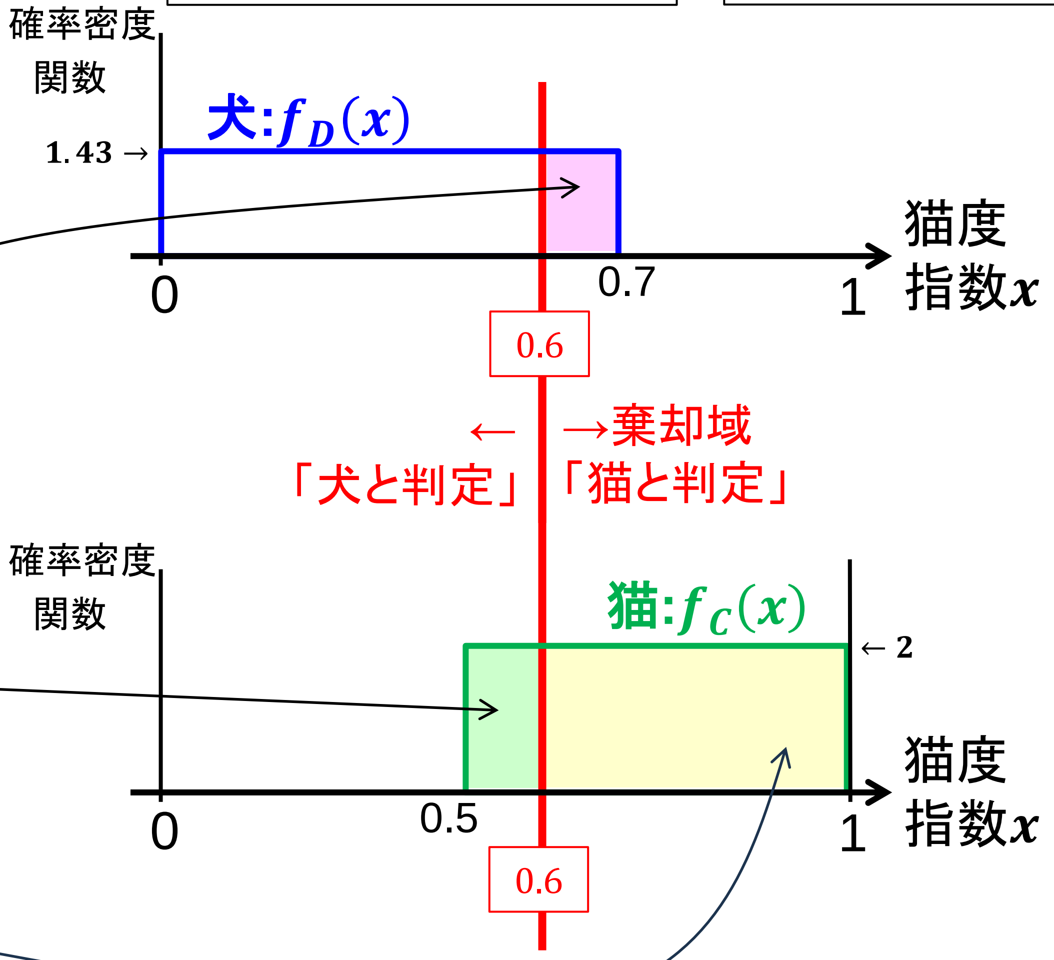
あ(A)わて者の誤り

Q1:どの部分の確率を求めたらいいでしょう？

⇒右図の

Q2:確率はいくらでしょう？

$\alpha = 1.43 \times 0.1 = 0.143$



(2)猫を犬と誤判定してしまう確率は？

第2種の過誤

ぼ(Bo)んやり者の誤り

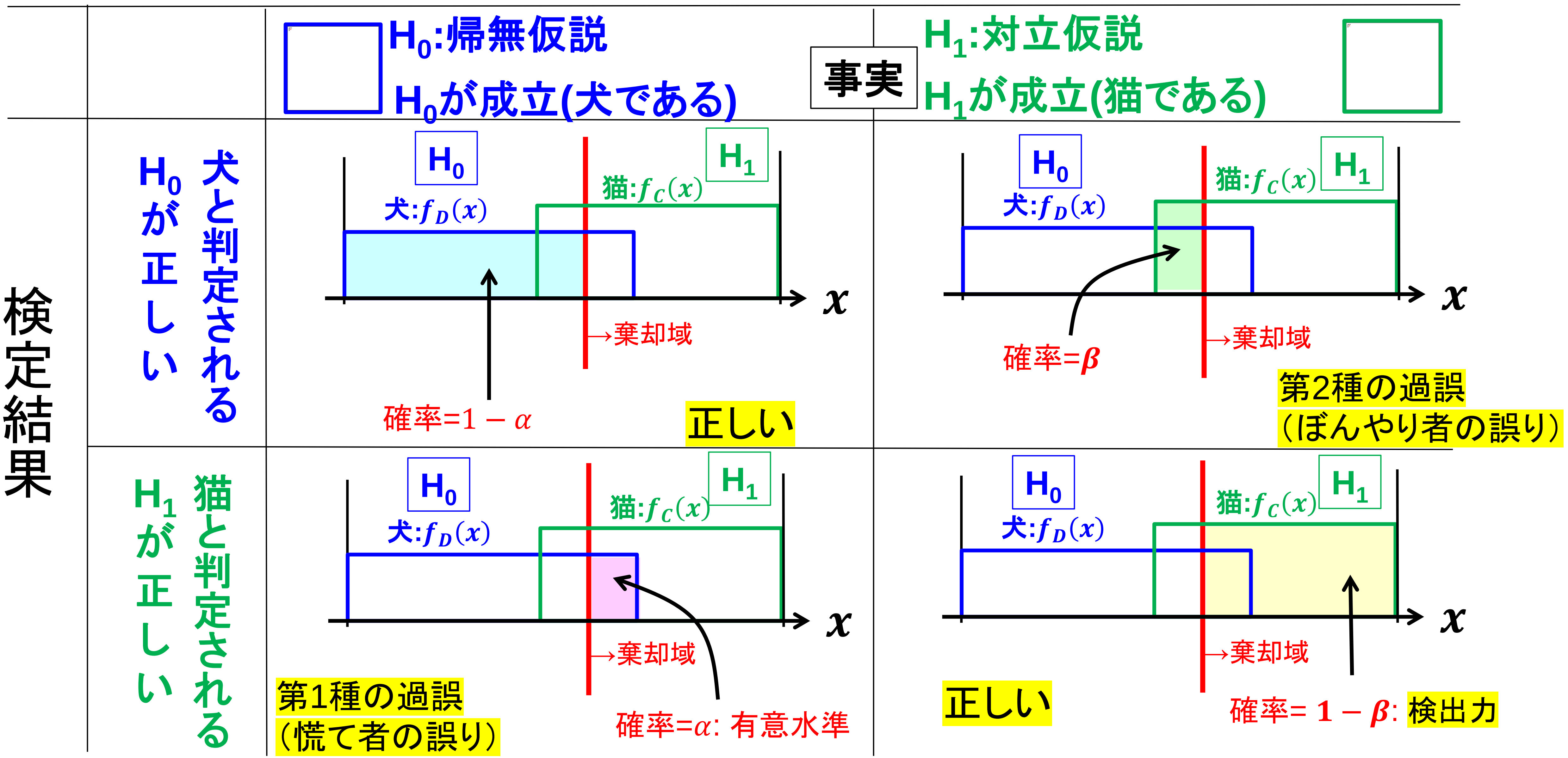
$\beta = 2 \times 0.1 = 0.20$

(3)猫を猫と正しく判定する確率は？

検出力

$1 - \beta = 2 \times 0.4 = 0.80$

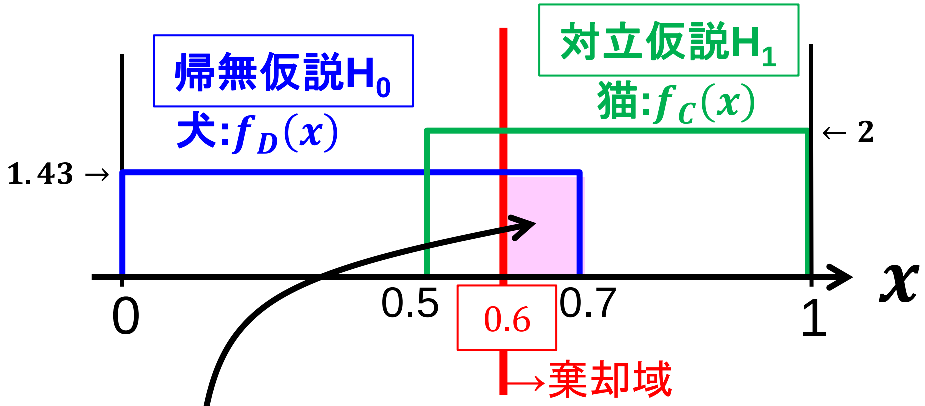
(p123.2c) [C8]問10.(準備)第1種、第２種の過誤、検出力



(p123.2d) [C8]問10.(準備)第1種、第２種の過誤、検出力

Q: 分布が、「一様分布」でなく、
「正規分布」の場合、
第１種、第2種の過誤の確率、検出力は
計算できる？

⇒計算できます！



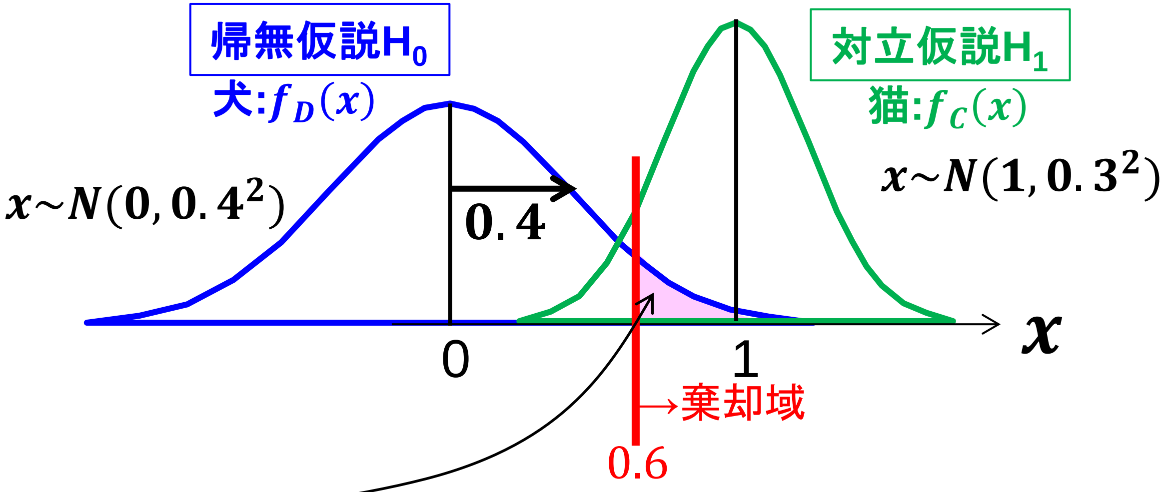
$\alpha = 1.43 \times 0.1 = 0.143$

帰無仮説(H_0): 写真の動物は犬である
⇒　(例)$x \sim N(0, 0.4^2)$　の場合を考えます
　(例)第1種の過誤（慌て者の誤り）の場合：

$\alpha = Pr(x \geq 0.6)$

$= Pr\left(z \geq \frac{0.6}{0.4} = 1.5\right) = 0.0668$

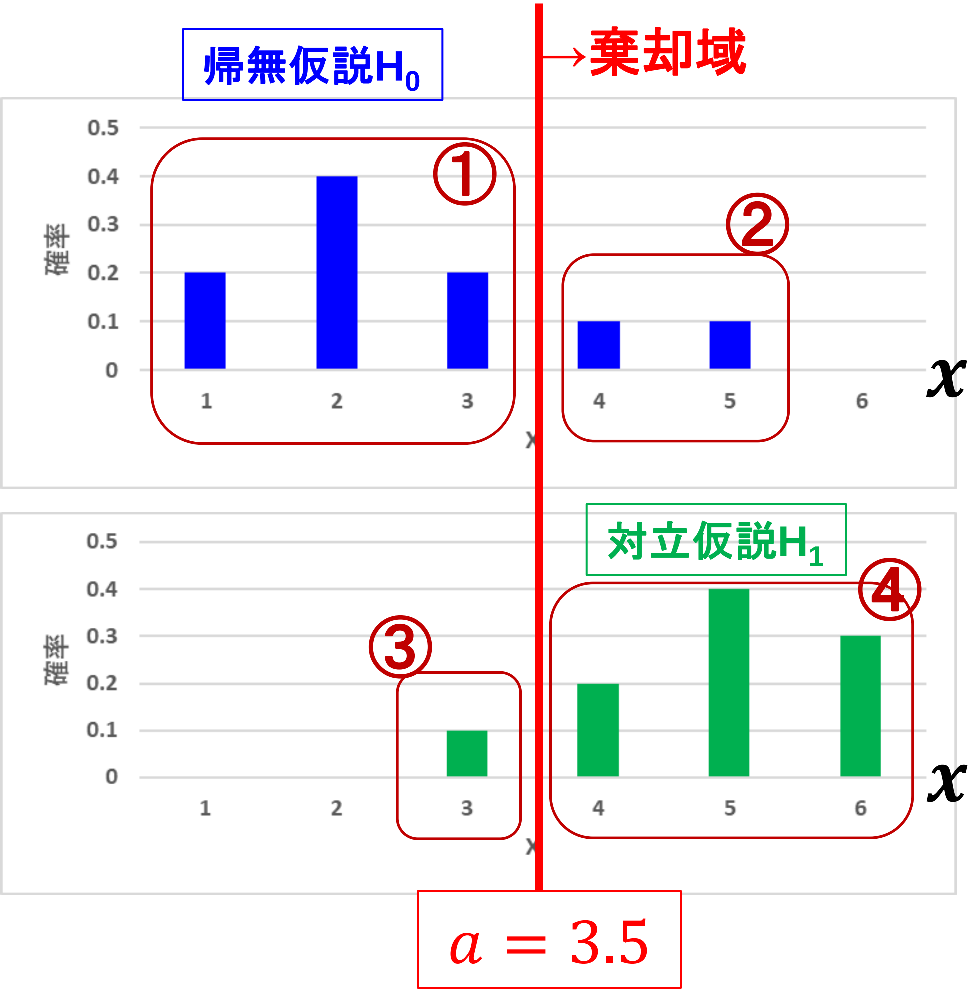
(解説) $x \sim N(0, 0.4^2)$　より、標準化: $z = \frac{x}{0.4} \sim N(0,1)$すると
正規分布表より確率が求まります



(p123.2e) [C8]問10.(準備)第1種、第２種の過誤、検出力

Q. 連続分布でなく、離散分布の場合は、計算できる？　⇒計算できます！

(例)出題例:2018年6月問13



		事実: H_0が成立	事実: H_1が成立
検定結果	H_0が正しい	① 正しい 確率= $1-\alpha=0.8$	③ 第2種の過誤 （ぼんやり者の誤り） 確率= $\boldsymbol{\beta}=\mathbf{0.1}$
検定結果	H_1が正しい	② 第1種の過誤 （慌て者の誤り） 確率= $\alpha=0.2$	④ 正しい 確率= $\mathbf{1}-\boldsymbol{\beta}=\mathbf{0.9}$ ：検出力

ひかり統計塾

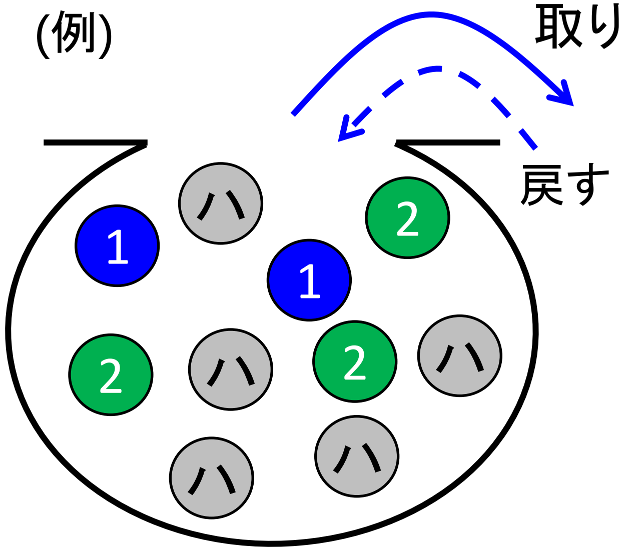
統計検定２級
公式問題集(CBT対応版)の解説
カテゴリー9： カイ二乗検定の分野
問1 (p126-127) @お試しサイト

(p126.1)[C9-1]問1. 適合度検定の基本

(ABランク)

くじ引き

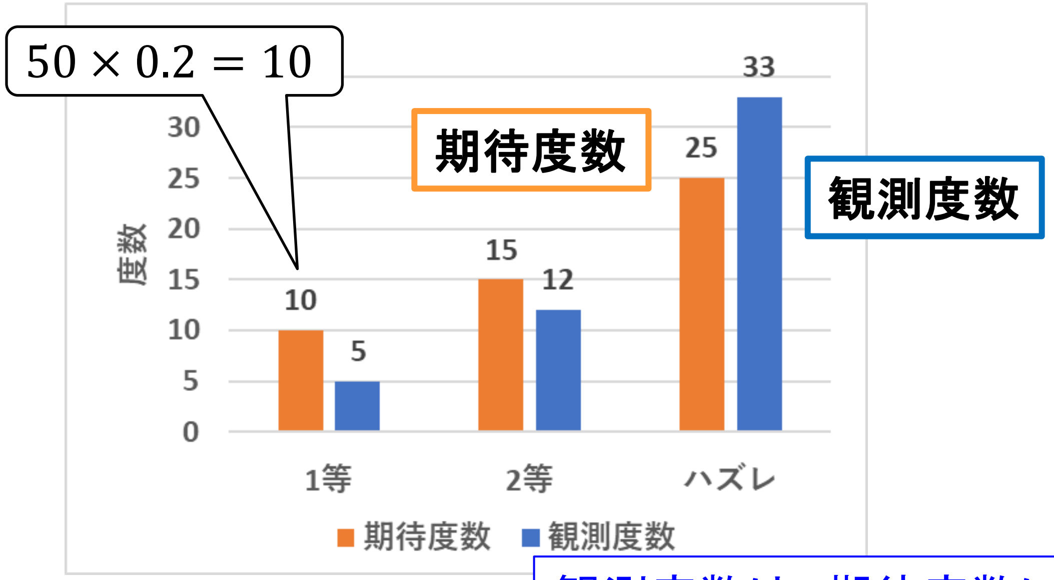
(例)



想定されている割合

1等:20%	2個	袋中の玉が10個の場合
2等:30%	3個	
ハズレ:50%	5個	

50人がくじを引いた時、



観測度数は、期待度数に比べ、
・当たりが少ない
・ハズレが多い

ずれがある。その可能性:
・くじの比率=想定比率だったが、たまたま、ずれが生じた
・くじの比率≠想定比率だったので、その結果、ずれが生じた
⇒検定します

(p126.2)[C9-1]問1. 適合度検定の基本

(ABランク)

帰無仮説(H_0):くじ引きの当たり外れの発生割合は、期待される割合と同じ
対立仮説(H_1):くじ引きの当たり外れの発生割合は、期待される割合と異なる

カテゴリー数=k個 (k=3)

	1等	2等	ハズレ	合計
観測度数	**5**	**12**	**33**	**50**
期待される割合	**0.2**	**0.3**	**0.5**	
期待度数	**(50×0.2=)10**	**(50×0.3=)15**	**(50×0.5=)25**	**50**
$\frac{(観測度数-期待度数)^2}{期待度数}$	$\frac{(5-10)^2}{10}=2.5$	$\frac{(12-15)^2}{15}=0.6$	$\frac{(33-25)^2}{25}=2.56$	**5.66**

(公式)
- カテゴリー数:kの時、自由度 $\phi=k-1$
- 検定統計量(χ^2統計量)は、以下で計算

$$\chi_0^2=\frac{(観測度数-期待度数)^2}{期待度数}の和$$

- 自由度ϕのχ^2分布を用いて検定する

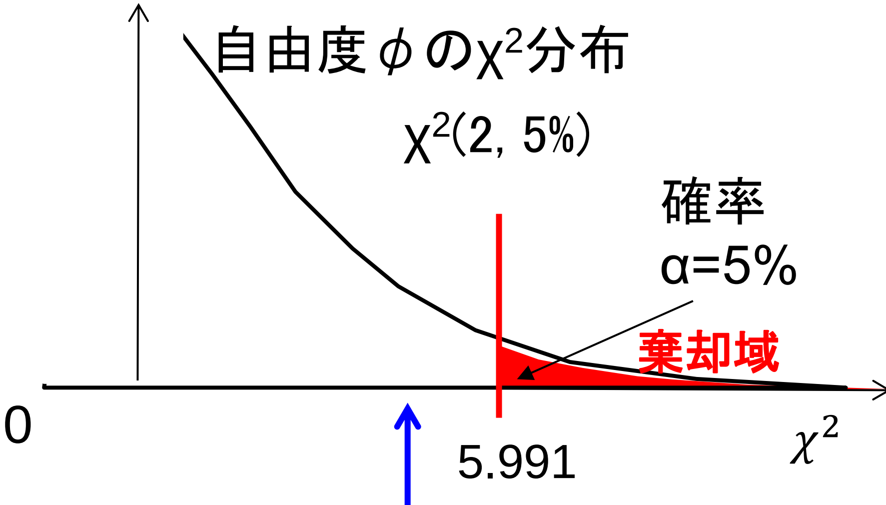
(ア) 自由度ϕ= **2**

(イ)検定統計量: χ_0^2 = **5.66**

(答) ④

χ²分布の上側P点

ϕ=	P=	0.975	0.950	0.050	0.025
	1	0.001	0.004	3.841	5.024
	2	0.051	0.103	5.991	7.378
	3	0.216	0.352	7.815	9.348
	4	0.484	0.711	9.488	11.143
	5	0.831	1.145	11.070	12.833
	6	1.237	1.635	12.592	14.449
	7	1.690	2.167	14.067	16.013
	8	2.180	2.733	15.507	17.535



$\chi_0^2=5.66$

帰無仮説(H_0)は棄却できない ⇒(ウ)
対立仮説(H_1)が正しいかどうか
何も言えない

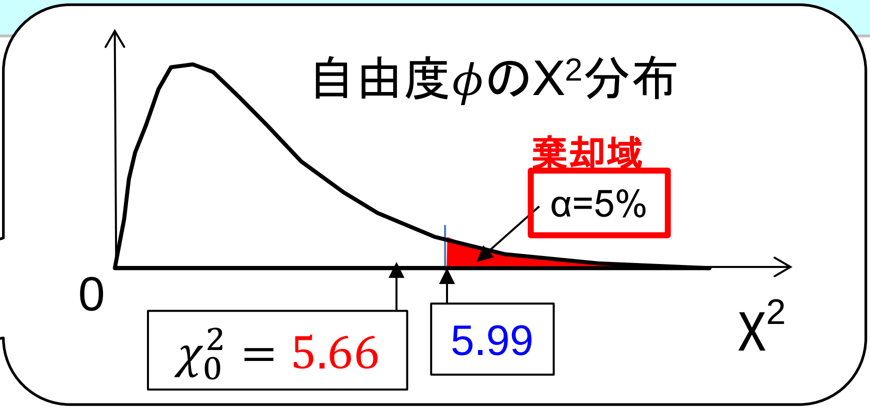
(答) ④

済

(p126.3)[C9-1]問1. 適合度検定の基本

(ABランク)

(1): 2つの仮説を立てます
帰無仮説(H_0):くじ引きの当たり外れの発生割合は、期待される割合と同じ
対立仮説(H_1):くじ引きの当たり外れの発生割合は、期待される割合と異なる



(2): 有意水準αを定めます
α=0.05=5%

(3):棄却域を定めます
カテゴリー数(k)= 3
自由度(ϕ)=k-1= 3-1=2
棄却域: $\chi^2 > \chi^2(\phi,\alpha) = \chi^2(2,0.05) = 5.99$

(4):検定統計量:$\chi_0^2 = \sum \frac{(観測度数-期待度数)^2}{期待度数}$
を計算します
$\Rightarrow \chi_0^2 = 5.66$

(5): H_0を棄却する･しないの判定
$\chi^2 = \chi_0^2$が棄却域にあれば、 H_0を棄却する
⇒ H_1は正しいと言える
$\chi^2 = \chi_0^2$が棄却域になければ、H_0は棄却できない
⇒ H_1は正しいかどうか何も言えない

カテゴリー数=k個 (k=3)

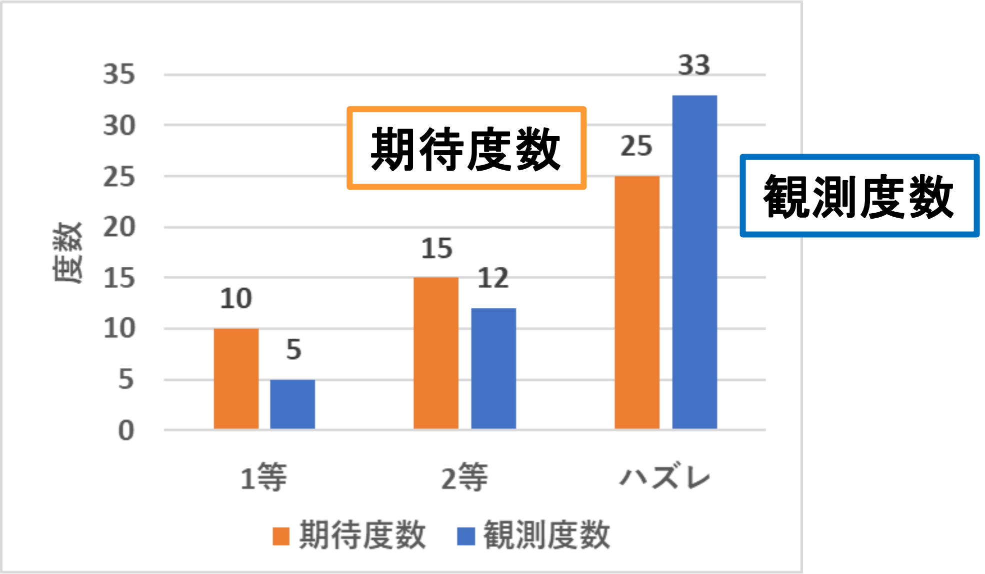
	1等	2等	ハズレ	合計
観測度数	**5**	**12**	**33**	**50**
期待されている割合	**0.2**	**0.3**	**0.5**	
期待度数	(50×0.2=)**10**	(50×0.3=)**15**	(50×0.5=)**25**	**50**
$\frac{(観測度数-期待度数)^2}{期待度数}$	$\frac{(5-10)^2}{10}$	$\frac{(12-15)^2}{15}$	$\frac{(33-25)^2}{25}$	**5.66**

出題のポイント:
･期待度数、検定統計量:χ_0^2
･自由度、棄却域、(ごくまれに)P値
･ H_0が棄却されるかどうか

(p126.4)[C9-1](補足)問1. 適合度検定とは

適合度の検定:
仮定された理論上の確率分布から得られる期待度数に対して、
データとして得られた観測値の度数の当てはまりの良さを検定する

例:
・一様分布(例:p128問2)
・特定の比率
　・今回のくじ引き
　・メンデルの法則(AA:Aa:aa=1:2:1)
・ポアソン分布(例:p132問4)
・正規分布
・・・



(p126.5)[C9-1]問1.(補足)適合度検定における自由度

カテゴリーの数：k個　(この例では、k=3)

期待度数

	1等	2等	ハズレ	合計
人数	10	15	25	50

-5　-3　$+8$

観測度数

	1等	2等	ハズレ	合計
人数	**5**	**12**	**33**	50

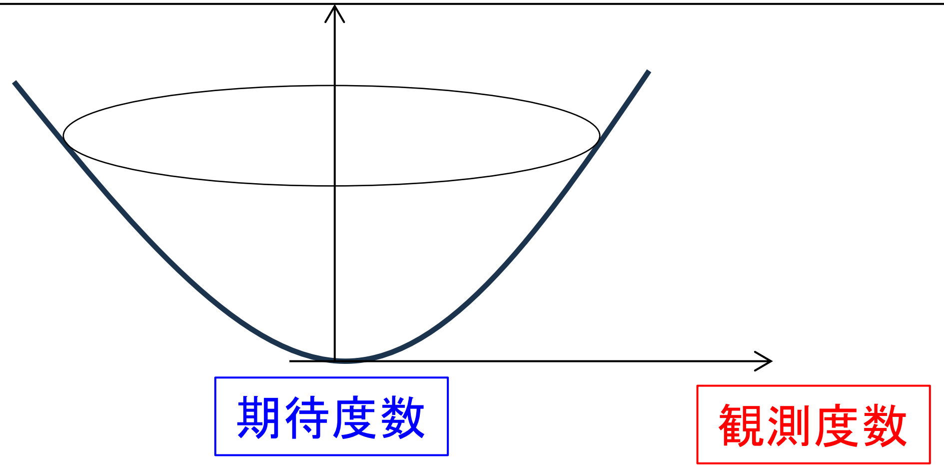
	1等	2等	ハズレ	合計
人数	★	★		150

2つの値は、独立に変化しうる

残り1個は自動的に決まる

⇒　自由度ϕ=2

検定統計量：$\chi_0^2 = \dfrac{(\text{観測度数} - \text{期待度数})^2}{\text{期待度数}}$ の和

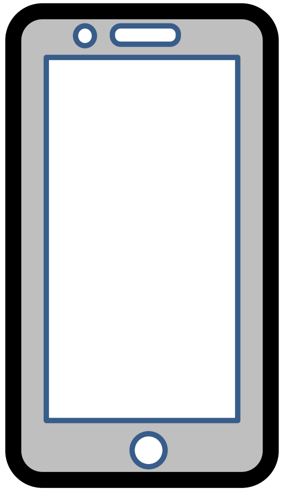


一般には、
カテゴリーの数：k個
(この例ではk=3)の時、
自由度ϕ=k-1
(この例ではϕ=3-1=2)

ひかり統計塾

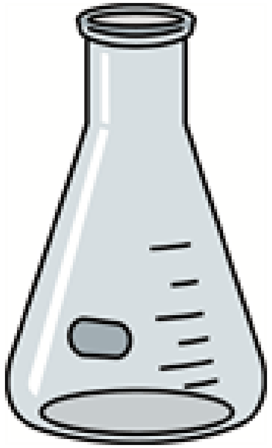
統計検定２級
公式問題集(CBT対応版)の解説
サブカテゴリー10-1．回帰分析の分野
(p142)回帰分析の基礎 @お試しサイト

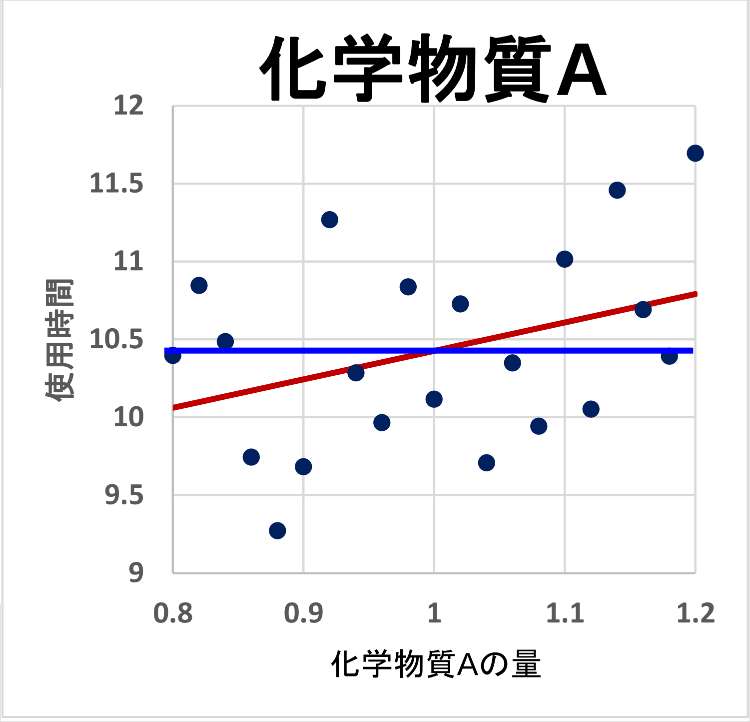
(p142.1a)[C10-1]回帰分析の基礎: 対象とする問題

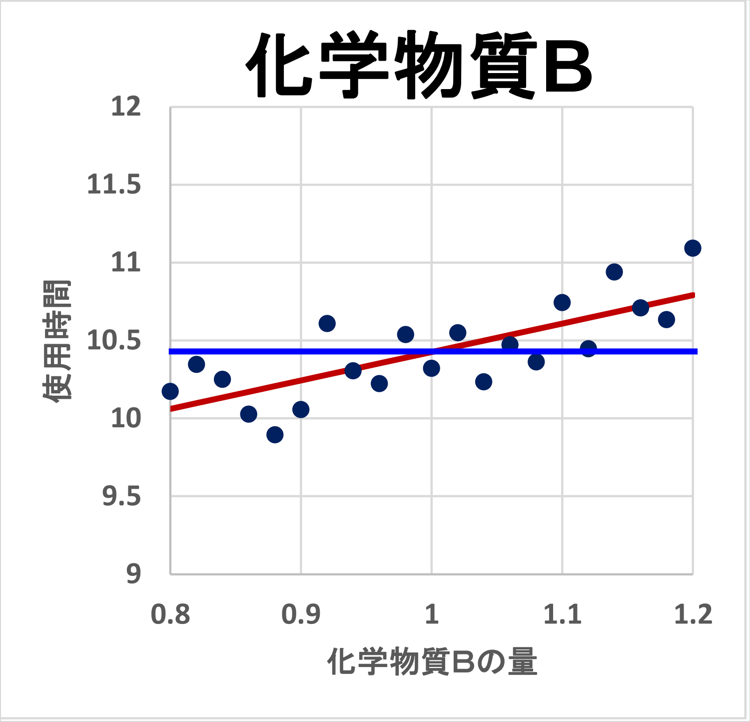


(問題)スマホの1充電当たりの使用時間を長くする検討を実施中と仮定。
・スマホのある素子作製工程で使う化学物質A、または化学物質Bの量を大きくすると、使用時間が長くなりそうな傾向がみられました。

(1)「化学物質の量(x)」と「使用時間(y)」の関係(式)は？　直線を想定すると？
(2)化学物質の量を変えると、使用時間が変わると言えるか？　　　を考えます



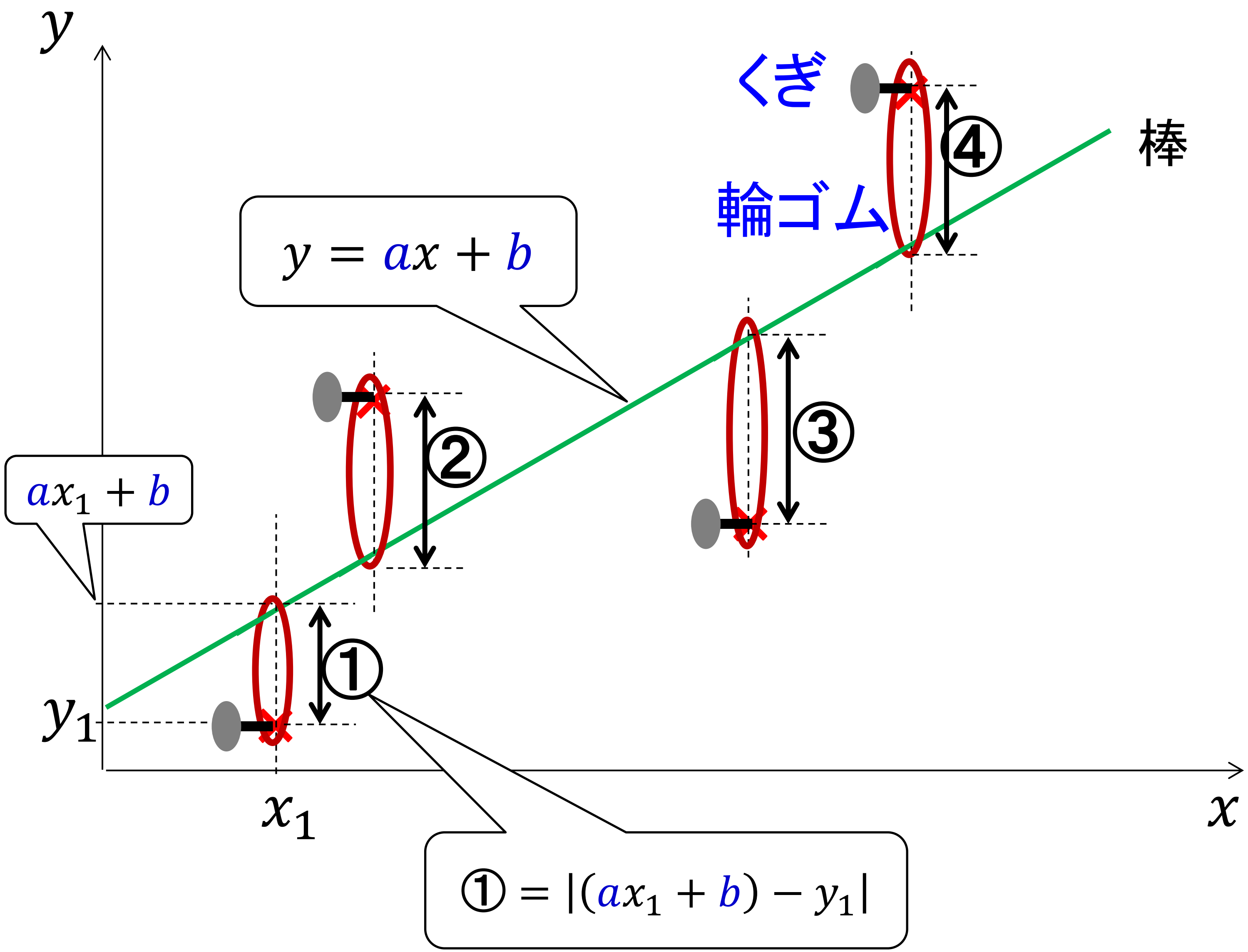




(1)直線を引く場合
直線をどう決める？
(回帰直線)

(2)傾き=0ですか？
傾き≠0ですか？

(p142.1b)[C10-1]回帰分析の基礎: 回帰直線の求め方



（大雑把な説明）
左図で、釣り合う「棒」の位置を探す

（より正確な説明）
①^2+②^2+ ③^2+ ④^2 を最小にするように、
直線($y = ax + b$)を決める
⇒係数a, bを決める
「最小二乗法」と呼ばれます

(p142.1c)[C10-1]回帰分析の基礎: 回帰直線の式

以降、$y = ax + b$ ⇒ $y = \beta_1 x + \beta_0$ と書きます。

⇓ 最小二乗法により決定

$y = \widehat{\beta_1} x + \widehat{\beta_0}$

$\widehat{\beta_1}, \widehat{\beta_0}$は、それぞれ$\beta_1, \beta_0$の「推定値」

公式(回帰直線、回帰係数):

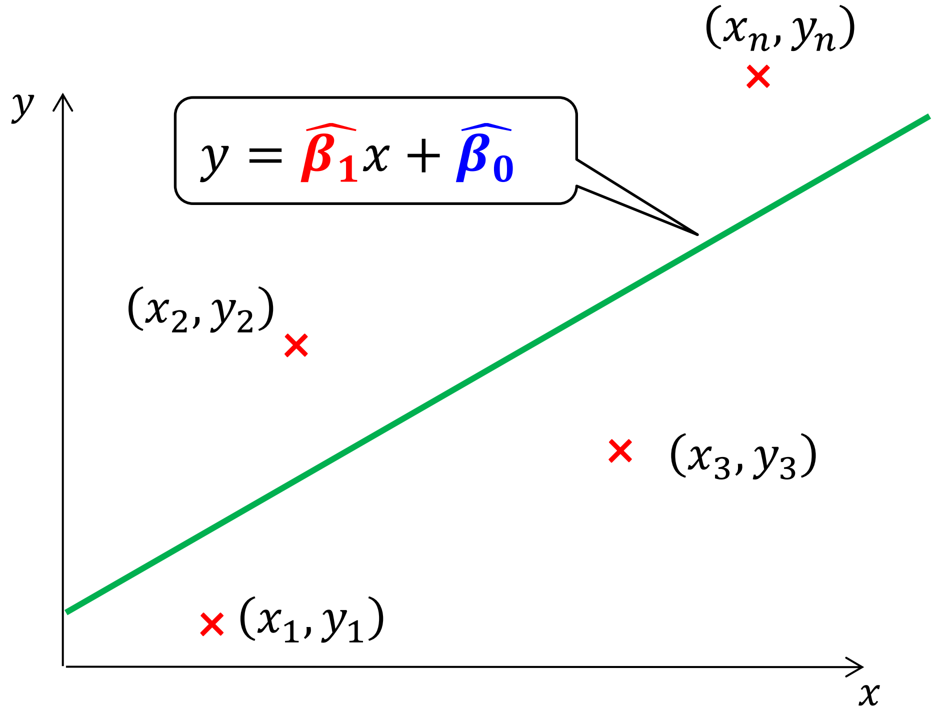
n個の点:$(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$ に対する回帰直線: $y = \widehat{\boldsymbol{\beta}_1} x + \widehat{\boldsymbol{\beta}_0}$

回帰係数: $\widehat{\boldsymbol{\beta}_1} = \dfrac{\boldsymbol{S_{xy}}}{\boldsymbol{S_{xx}}}$, $\quad \widehat{\boldsymbol{\beta}_0} = \bar{\boldsymbol{y}} - \widehat{\boldsymbol{\beta}_1} \bar{\boldsymbol{x}}$

x, yの平均: $\bar{x} = \dfrac{1}{n}(x_1 + x_2 + \cdots + x_n)$, $\bar{y} = \dfrac{1}{n}(y_1 + y_2 + \cdots + y_n)$

偏差積和: $\boldsymbol{S_{xy}} = (\boldsymbol{x_1} - \bar{\boldsymbol{x}})(\boldsymbol{y_1} - \bar{\boldsymbol{y}}) + (\boldsymbol{x_2} - \bar{\boldsymbol{x}})(\boldsymbol{y_2} - \bar{\boldsymbol{y}}) + \cdots + (\boldsymbol{x_n} - \bar{\boldsymbol{x}})(\boldsymbol{y_n} - \bar{\boldsymbol{y}})$

偏差平方和: $\boldsymbol{S_{xx}} = (\boldsymbol{x_1} - \bar{\boldsymbol{x}})^2 + (\boldsymbol{x_2} - \bar{\boldsymbol{x}})^2 + \cdots + (\boldsymbol{x_n} - \bar{\boldsymbol{x}})^2$



(p142.1d)[C10-1]回帰分析の基礎: 回帰直線の特徴,憶え方

$$y=\widehat{\beta_1}x+\widehat{\beta_0},\quad \widehat{\beta_1}=\frac{S_{xy}}{S_{xx}},\ \widehat{\beta_0}=\bar{y}-\widehat{\beta_1}\bar{x}$$

特徴①: 回帰直線は点$(\bar{x},\bar{y})$を通る

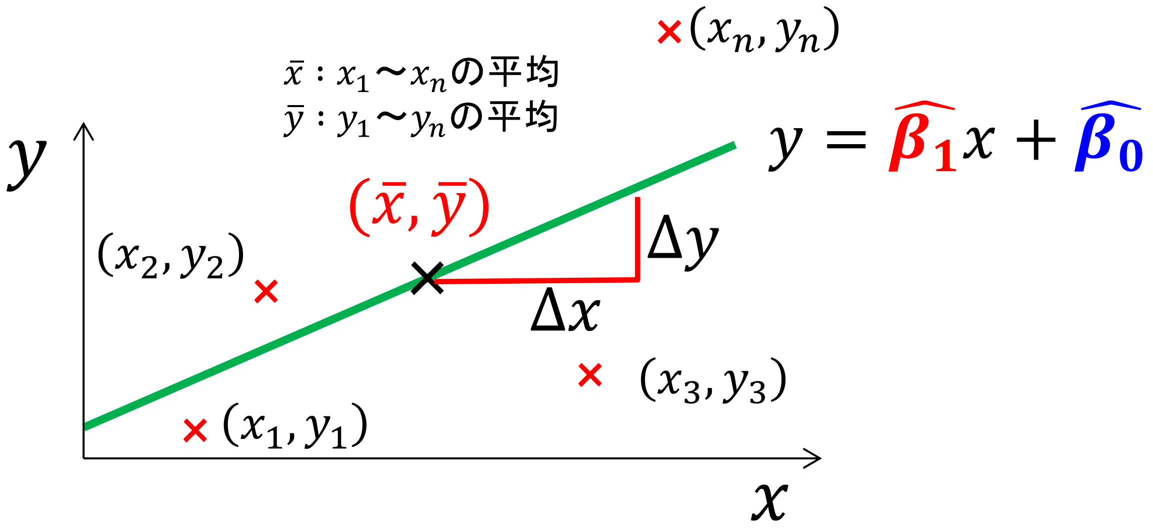
役立つ問題例:
・回帰直線を選ぶ
(例:p143問1[2])

$$\widehat{\beta_0}=\bar{y}-\widehat{\beta_1}\bar{x}$$

$$\Downarrow$$

$$\bar{y}=\widehat{\beta_1}\bar{x}+\widehat{\beta_0}$$

点$(\bar{x},\bar{y})$は回帰直線 $y=\widehat{\beta_1}x+\widehat{\beta_0}$ 上にある



特徴②: $\widehat{\beta_1}$は、「傾き」に対応

$$\widehat{\beta_1}=\frac{S_{xy}}{S_{xx}}$$

$$\widehat{\beta_1}\sim\frac{\Delta y}{\Delta x}\sim\frac{(y-\bar{y})}{(x-\bar{x})}\sim\frac{(x-\bar{x})(y-\bar{y})}{(x-\bar{x})(x-\bar{x})}\sim\frac{\Sigma(x-\bar{x})(y-\bar{y})}{\Sigma(x-\bar{x})(x-\bar{x})}\sim\frac{\sum_{i=1}^{n}(x_i-\bar{x})\boxed{(y_i-\bar{y})}}{\sum_{i=1}^{n}(x_i-\bar{x})\boxed{(x_i-\bar{x})}}=\frac{S_{xy}}{S_{xx}}$$

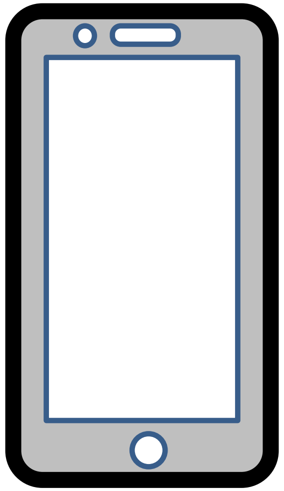
傾き

重み

偏差積和: $S_{xy}=\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})$

偏差平方和: $S_{xx}=\sum_{i=1}^{n}(x_i-\bar{x})(x_i-\bar{x})$

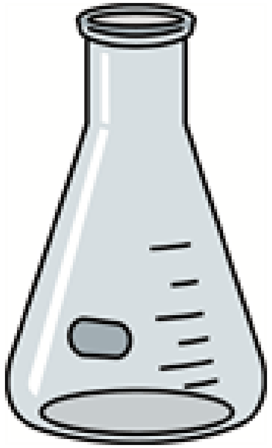
(p142.1e)[C10-1]回帰分析の基礎: 対象とする問題

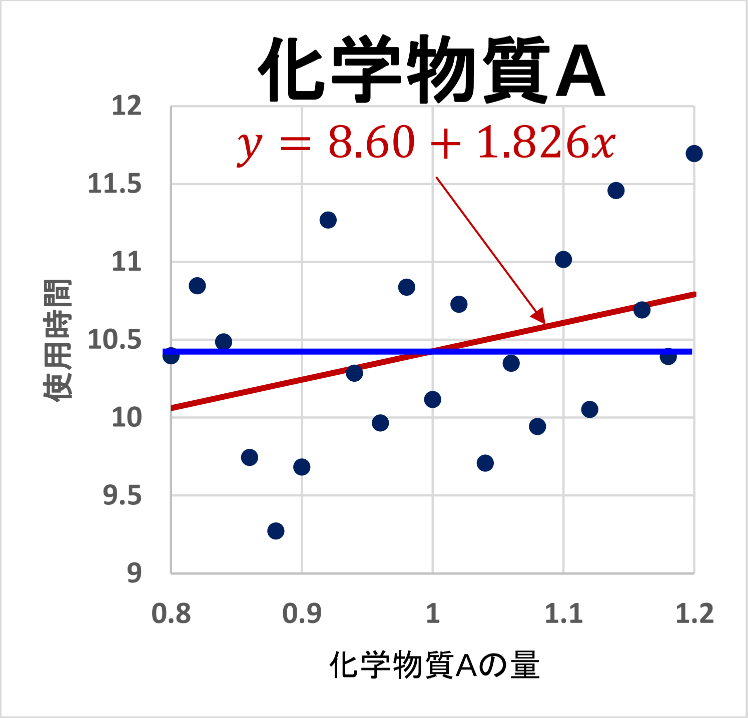


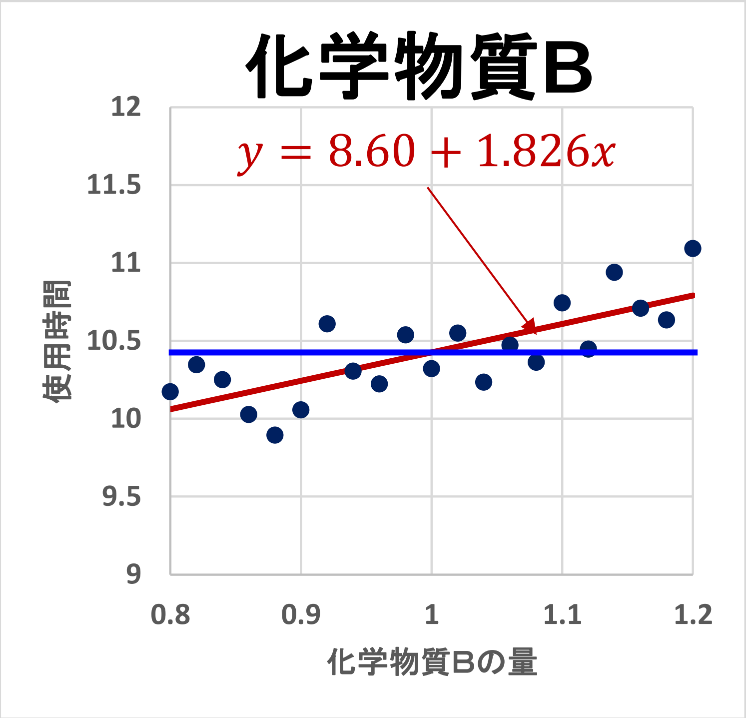
(1)「化学物質の量(x)」と「使用時間(y)」の関係（式）は？　直線を想定すると？

(2)化学物質の量を変えると、使用時間が変わると言えるか？　　　　　　を考えます

済







(1)直線を引く場合
直線をどう決める？
⇒同じ回帰直線を得た

(2)傾き≠0と言える？

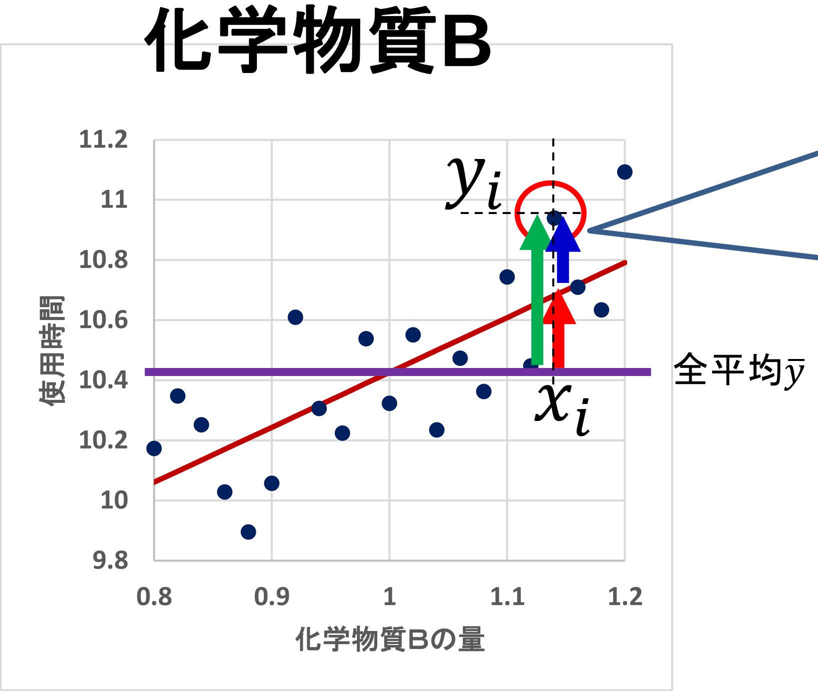
回帰の評価

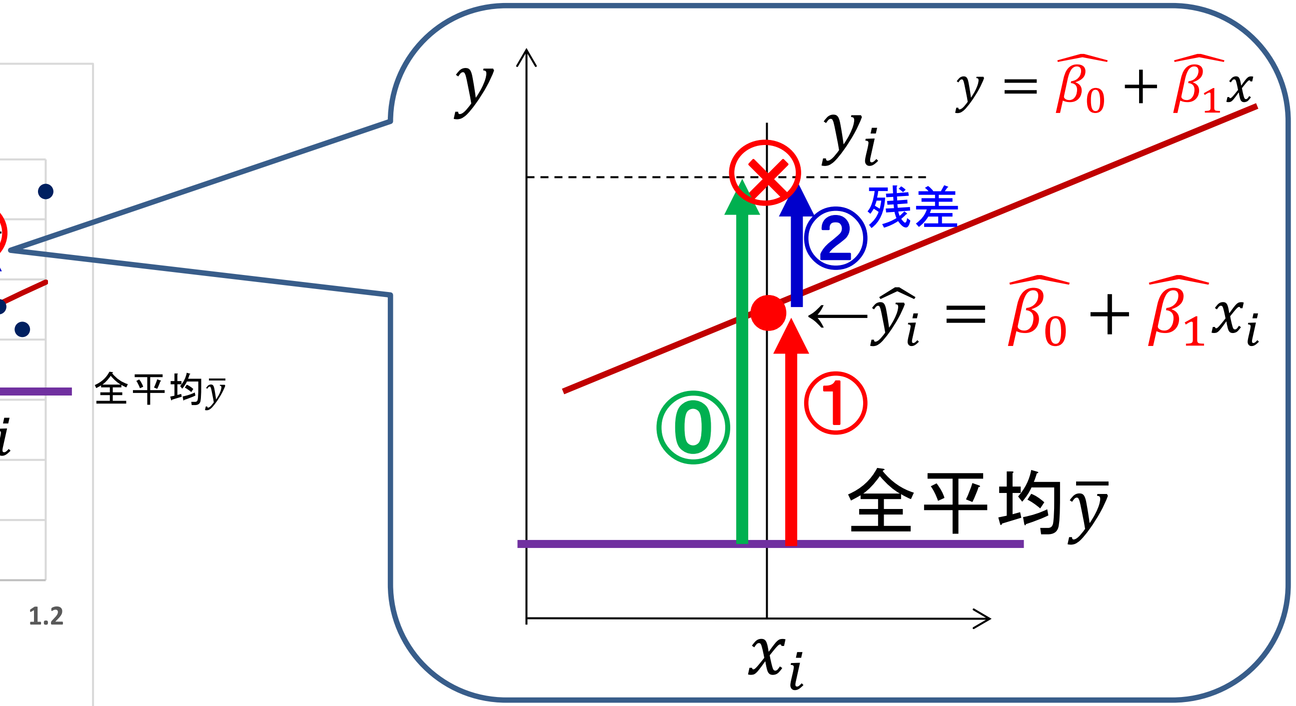
$y = 8.60 + 1.826x \Rightarrow \ x = 1.1$の時、$y = 10.61$と言い切れる？

(p142.1f)[C10-1]回帰分析の基礎: 回帰の評価、平方和の分解

分散分析の場合と同様にばらつきを①回帰、②残差の2つに分けます

(注)回帰では、
「誤差」の代わりに
「残差」を使います
(参考)入門統計解析法
p188 [注8.5]





⓪総平方和: $S_T = \sum_{i=1}^{n} (y_i - \bar{y})^2 = \sum_{i=1}^{n} \left((y_i - \hat{y}_i) + (\hat{y}_i - \bar{y}) \right)^2$

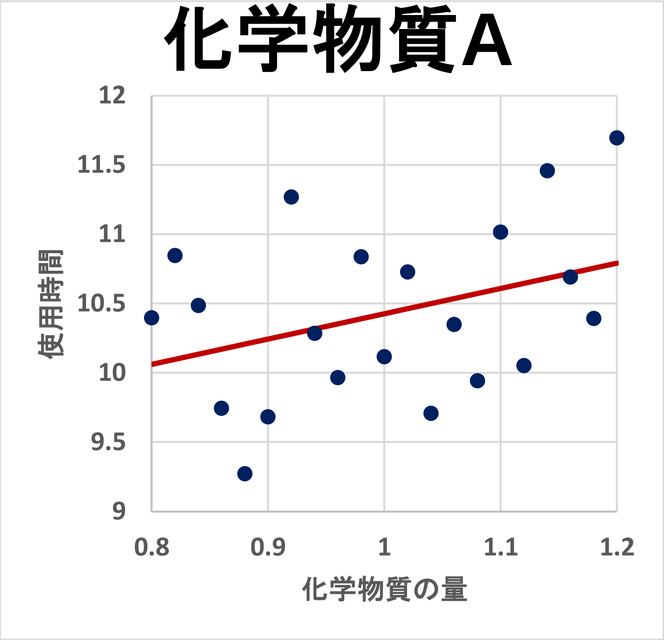
①回帰(Regression)による平方和: $S_R = \sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2$

②残差平方和: $S_e = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$

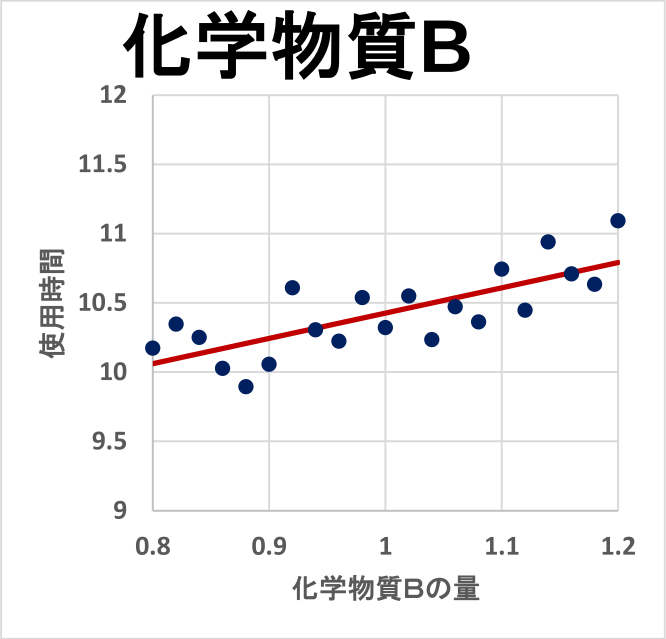
↓(参考)入門統計解析法
p186 (8.17)式

分散分析の場合と同様に $S_T = S_R + S_e$ が成立します！

(p142.1g)[C10-1]回帰分析の基礎: 決定係数(寄与率)



化学物質	A	B
回帰: S_R	1.027	1.027
残差: S_e	6.789	0.754
計: S_T	7.816	1.781
決定係数	0.131	0.577



①回帰 13.1%
②残差 86.9%
⓪総平方和

総平方和 S_Tのうち、
回帰の寄与 S_Rの割合が
決定係数(寄与率)です。

①回帰 57.7%
②残差 42.3%
⓪総平方和

化学物質	A	B
相関係数	0.362	0.759
相関係数^2	0.131	0.577
決定係数	0.131	0.577

$$\text{決定係数} = \frac{S_R}{S_T} = r^2 \quad (r\text{: 相関係数})$$

(p142.1h)[C10-1]回帰分析の基礎: 分散分析のステップ

理論：帰無仮説が正しい時、検定統計量: $F_0 = \frac{V_R}{V_e}$は自由度$(\boldsymbol{\phi_R}, \boldsymbol{\phi_e})$のF分布に従う

(1): 2つの仮説:
帰無仮説 H_0: $\beta_1 = 0$
対立仮説 H_1: $\beta_1 \neq 0$

回帰式:
$y = \beta_0 + \beta_1 x$

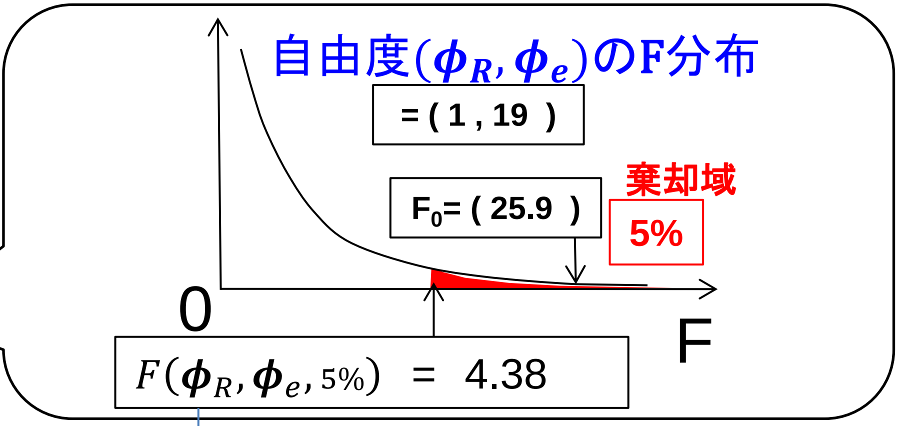
$\bar{x} = (\ 1.0000\ )$ $S_{xx} = (\ 0.3080\ )$
$\bar{y} = (\ 10.4261\ )$ $S_{xy} = (\ 0.5624\ )$

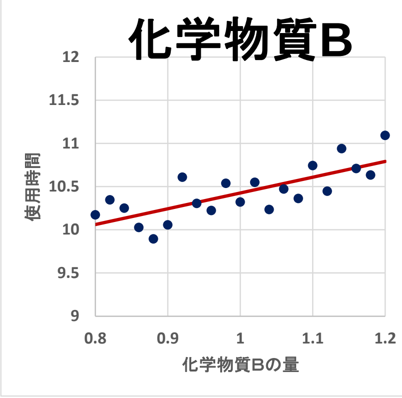
$\widehat{\beta_1} = \frac{S_{xy}}{S_{xx}}$
$\widehat{\beta_0} = \bar{y} - \widehat{\beta_1}\bar{x}$

$\widehat{\beta_1} = \frac{S_{xy}}{S_{xx}} = 1.826$
$\widehat{\beta_0} = \bar{y} - \widehat{\beta_1}\bar{x} = 8.600$

(2): 有意水準αを定めます
α=0.05=5%

(3): 自由度を求め、棄却域を定めます
データ総数n=(21),
全自由度$\phi_T = n - 1 = (\ 20\ )$
回帰の自由度: (単回帰では) ⇒ $\phi_R = 1$
残差の自由度: $\phi_e = \phi_T - \phi_R = (\ 19\ )$
棄却域: $F > F(\phi_R, \phi_e, 5\%)$
$F(1, (\ 19\ ), 5\%) = (\ 4.38\ )$





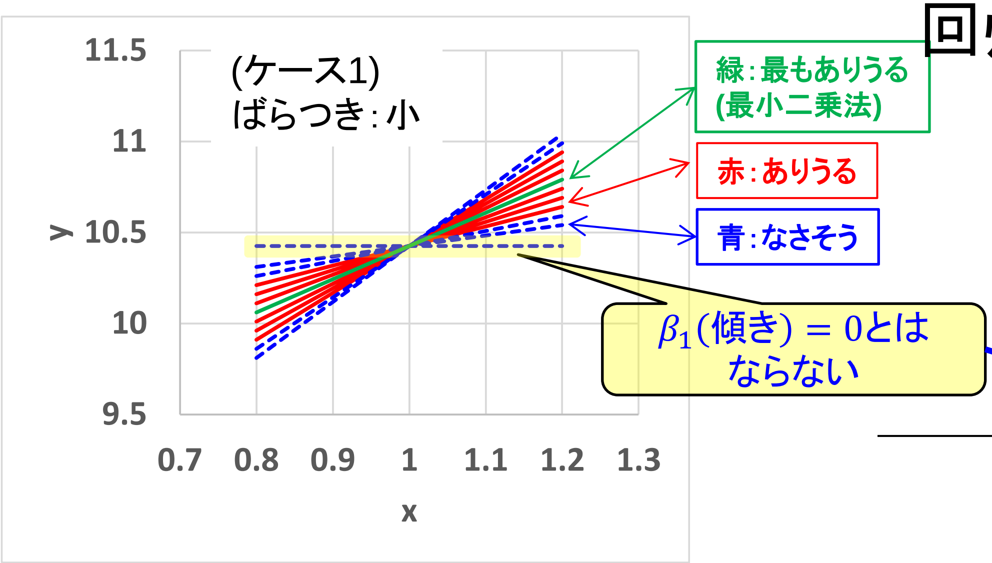
(4)分散分析表

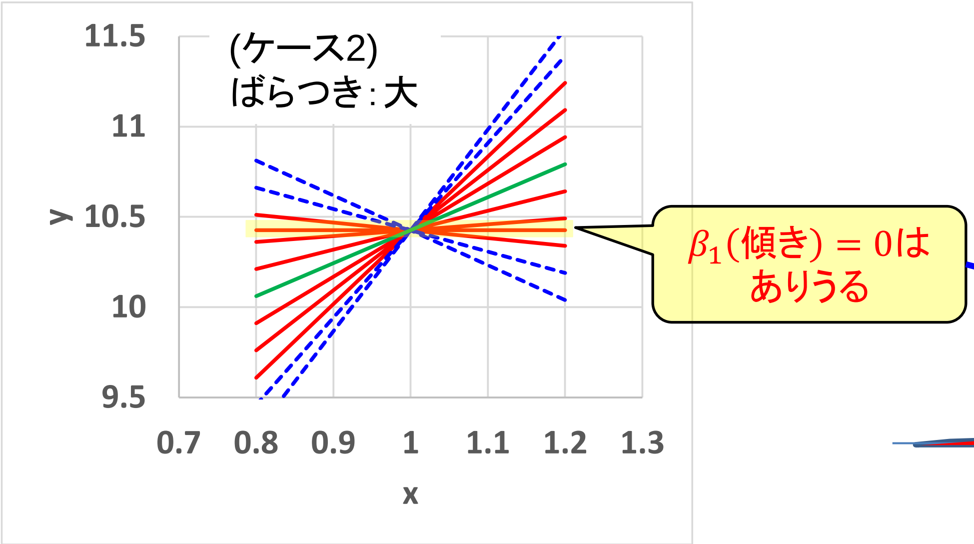
要因	(a)平方和 S		(b)自由度 φ		(c)平均平方 V=S/φ	F_0 =V_R/V_e	棄却域の下限
R(回帰)	1.027	÷	1	=	1.027	25.9	4.38
e(残差)	0.754	÷	19	=	0.0397	>	
計	1.781		20				

大小関係(<,>)

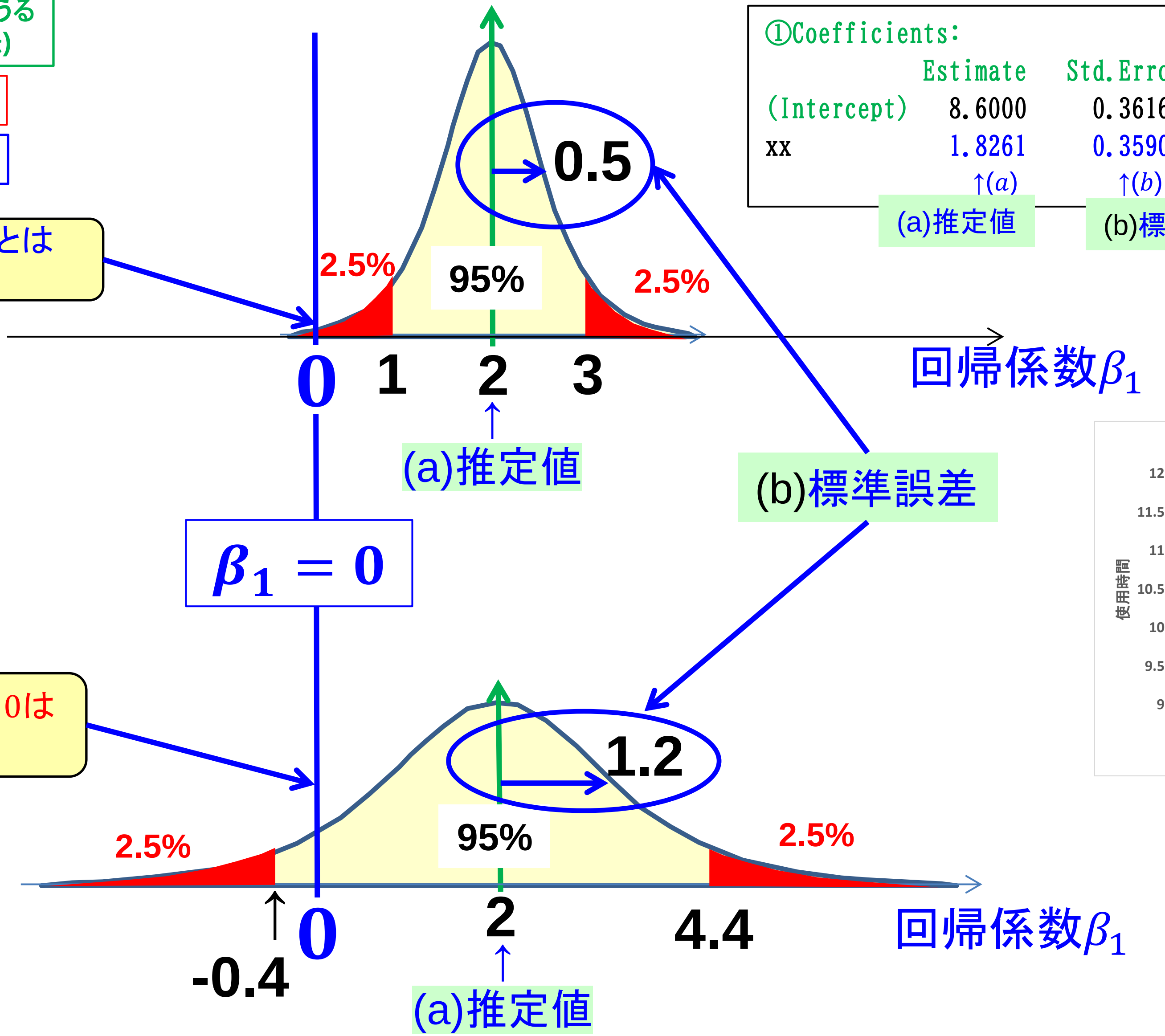
(5): H_0を棄却する・しないの判定
F_0 =(25.9) は棄却域に(ある ・ ない)
どっちですか？
(●) $F = F_0$が棄却域にあるので、H_0を棄却する
⇒ H_1は正しいと言える
( ) $F = F_0$が棄却域にないので、H_0は棄却できない
⇒ H_1は正しいかどうか何も言えない

(p142.1i)[C10-1]回帰分析の基礎：回帰係数≠0の検定



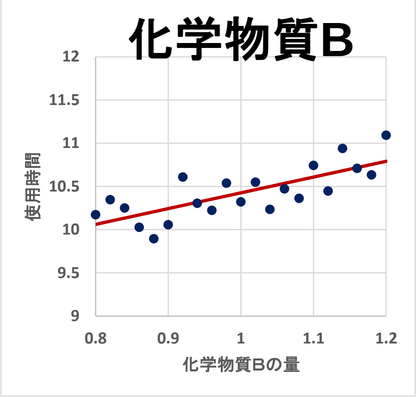


回帰係数の取りうる範囲(イメージ)




```
①Coefficients:
            Estimate   Std.Error t value Pr(>|t|)
(Intercept)   8.6000     0.3616   23.780 1.34e-15 ***
xx            1.8261     0.3590    5.086 6.56e-05 ***
                ↑(a)       ↑(b)
```

(a)推定値　(b)標準誤差

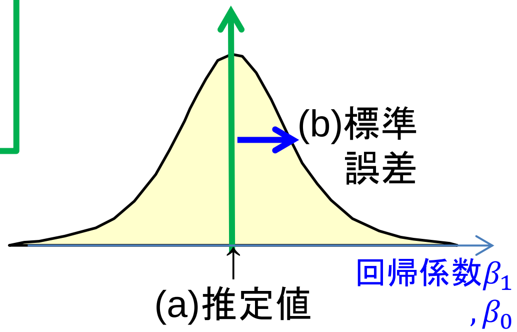


(p142.1j)[C10-1]回帰分析の基礎: 回帰係数≠0の検定

(公式) $t = \dfrac{(a)\text{推定値} - \beta_1}{(b)\text{標準誤差}} \sim$ 自由度ϕ_eのt分布

単回帰分析での残差の自由度$\phi_e = n - 2$

(n: データ組数) $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$



注目する回帰係数: (例) β_1

(a)推定値(Estimate)=1.8261

(b)標準誤差(Std. Error)=0.3590

(c)残差の自由度$= 21 - 2 = 19$

$$t = \frac{(a)\text{推定値} - \beta_1}{(b)\text{標準誤差}} = \frac{1.8261 - \beta_1}{0.3590} \sim \text{自由度}\phi_e = 19\text{の}t\text{分布}$$

統計ソフトRによる出力結果

```
①Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   8.6000     0.3616  23.780 1.34e-15 ***
xx            1.8261     0.3590   5.086 6.56e-05 ***
              ↑(a)       ↑(b)    ↑(e)
```

日本語版(ワンコピペエクセルシートより)

		①Estimate	①Std. Error	①t value	①Pr(>|t|)		
	①Coefficients 回帰係数	推定値	標準誤差	検定統計量 (t 値)	P値	区間推定 (下限)	区間推定 (上限)
(①Intercept) (切片)	β_0 (^)=	8.6000	0.3616	23.7805	1.33227E-15	7.8430	9.3569
(傾き)	β_1 (^)=	1.8261	0.3590	5.0864	0.000066	1.0747	2.5775
		↑(a)	↑(b)	↑(e)		↑(g2)	↑(g1)

$$y = \widehat{\beta_1}x + \widehat{\beta_0} = 1.8261\,x + 8.6$$

検定

両側検定

(1) 帰無仮説(H_0): $\beta_1 = 0$
対立仮説(H_1): $\beta_1 \neq 0$

(2) 有意水準
α=0.05

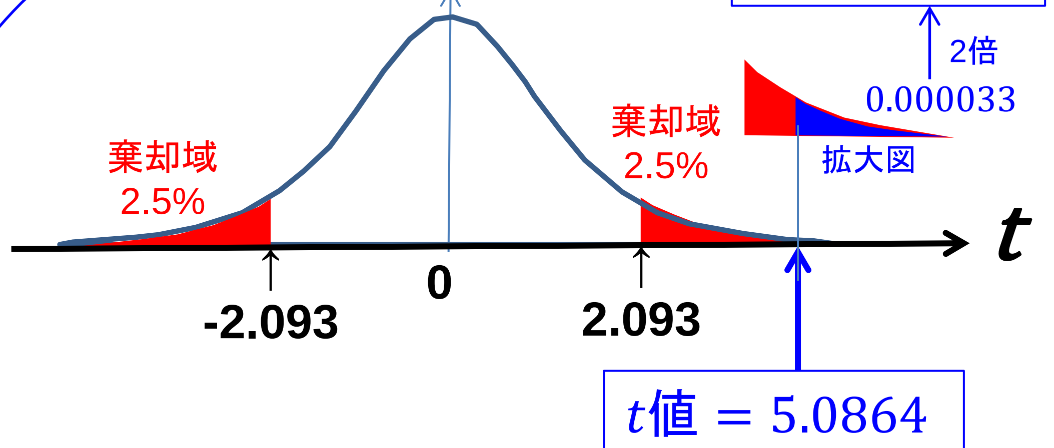
(3) 棄却域
|t|≧2.093

(4) 帰無仮説($\beta_1 = 0$)が正しい時の検定統計量

$$t = \frac{1.8261 - \beta_1}{0.3590} = \frac{1.8261 - 0}{0.3590} = 5.0864$$

(5) 検定統計量(t値)は棄却域にある
⇒帰無仮説(H_0)を棄却する⇒対立仮説(H_1)($\beta_1 \neq 0$)は正しいと言える

自由度$\phi_e = 19$のt分布



(p142.1k)[C10-1]回帰分析の基礎: 回帰係数の区間推定

前ページと同じ

(公式) $t = \dfrac{(a)\text{推定値} - \beta_1}{(b)\text{標準誤差}} \sim$ 自由度ϕ_eのt分布

単回帰分析での残差の自由度$\phi_e = n - 2$

(n: データ組数) $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$

注目する回帰係数: β_1

(a)推定値(Estimate)=1.8261

(b)標準誤差(Std. Error)=0.3590

(c)残差の自由度$= 21 - 2 = 19$

$$t = \frac{(a)\text{推定値} - \beta_1}{(b)\text{標準誤差}} = \frac{1.8261 - \beta_1}{0.3590} \sim \text{自由度}\phi_e = 19\text{の}t\text{分布}$$

統計ソフトRによる計算結果

```
①Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)     8.6000     0.3616  23.780 1.34e-15 ***
xx              1.8261     0.3590   5.086 6.56e-05 ***
                ↑(a)       ↑(b)     ↑(e)
```

日本語版(ワンコピペエクセルシートより)

	①Coefficients	①Estimate	①Std. Error	①t value	①Pr(>|t|)		
	回帰係数	推定値	標準誤差	検定統計量 (t 値)	P値	区間推定 (下限)	区間推定 (上限)
(①Intercept) (切片)	β_0 (^)=	8.6000	0.3616	23.7805	1.33227E-15	7.8430	9.3569
(傾き)	β_1 (^)=	1.8261	0.3590	5.0864	0.000066	1.0747	2.5775
		↑ (a)	↑ (b)	↑ (e)		↑ (g2)	↑ (g1)

区間推定の場合

$t \sim$ 自由度$\phi_e = 19$のt分布 の時、信頼確率95%で

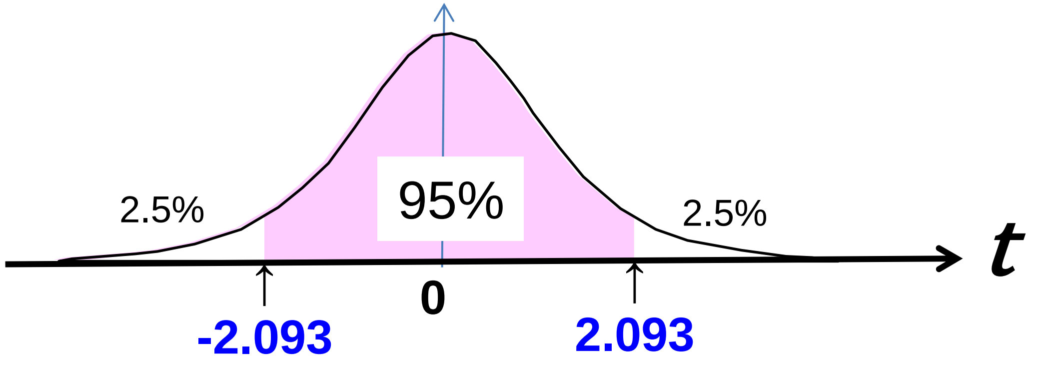
$$|t| \leq 2.093$$

$$|t| = \left|\frac{1.8261 - \beta_1}{0.3590}\right| \leq 2.093$$

$$|1.8261 - \beta_1| \leq 2.093 \times 0.3590 = 0.7514$$

$$1.0747 \leq \beta_1 \leq 2.5775$$

自由度$\phi_e = 19$のt分布



(p142.1L)[C10-1]回帰分析の基礎: 統計ソフト「R」の結果

統計検定2級の回帰の問題では、統計ソフトウェア(R)による出力結果が示され、
これに基づく解釈に関する問題が頻繁に出されます　⇒読み取れるようにしておいてください

(問題例)CBT問題集p147,148,151,171,188

統計ソフトウェア:Rによる回帰分析結果

```
Residuals:
     Min       1Q   Median       3Q      Max
-0.31152 -0.14160 -0.01048  0.13554  0.32984
```

残差の
最小値・最大値、
四分位数・中央値
(関連)CBT問題集
p152問3[1]④

```
①Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   8.6000     0.3616  23.780 1.34e-15 ***
xx            1.8261     0.3590   5.086 6.56e-05 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

②Residual standard error: 0.1992 on 19 degrees of freedom
③Multiple R-squared:  0.5766, Adjusted R-squared:  0.5543
④F-statistic: 25.87 on 1 and 19 DF,  p-value: 6.558e-05
```

日本語版:
(ワンコピペエクセルシートでの回帰分析結果)

分散分析表			↓検定統計量			↓(④p-value)	↓(1:効果あり、0:無)	
要因	平方和	自由度	平均平方	F_0値	棄却域(下限)	P値	判定	効果
R(回帰)	1.027	1	1.027051962	25.8710	4.3807	6.55798E-05	1	有
E(残差)	0.754	19	0.039698956	―――――	―――――	―――――	―――――	―――――
計	1.781	20	―――――	―――――	―――――	―――――	―――――	―――――
	残差の標準誤差($\sqrt{V_e}$)=		0.199245969	↑④F-statistic		④回帰の自由度(DF)=	1	
						④残差の自由度(DF)=	19	

●ステップ4――――――:　検定・推定を行う

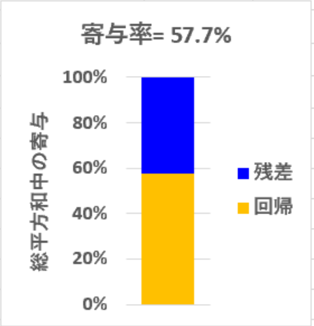
検定：　検定統計量F_0値は棄却域に　あるので

帰無仮説　H_0：　「回帰に意味はない」は　棄却できる

対立仮説　H_1：　「回帰に意味がある」　と言える

P値=　0.00007　⇔ α：0.05

(別解)P値はα(有意水準)より　小さいのでH0を棄却できる



●ステップ5――――――:　回帰分析の結果：

決定係数(寄与率)=	0.576564	(③Multiple R-squared)
自由度調整済決定係数=	0.5542779	(③Adjusted R-squared)
残差の標準誤差($\sqrt{Ve}$)=	0.199246	(②Residual standard error)

②残差の自由度(degrees of freedom)=　19

		①Estimate	①Std. Error	①t value	①Pr(>|t|)		
	①Coefficients			検定統計量		区間推定	区間推定
	回帰係数	点推定値	標準誤差	(t 値)	P値	(下限)	(上限)
(①Intercept)(切片)	β_0(^)=	8.6000	0.361639519	23.7805	1.33227E-15	7.8430	9.3569
(傾き)	β_1(^)=	1.8261	0.359016316	5.0864	6.55798E-05	1.0747	2.5775
	y = β_0(^) + β_1(^) x				(比較値：0)		

ひかり統計塾

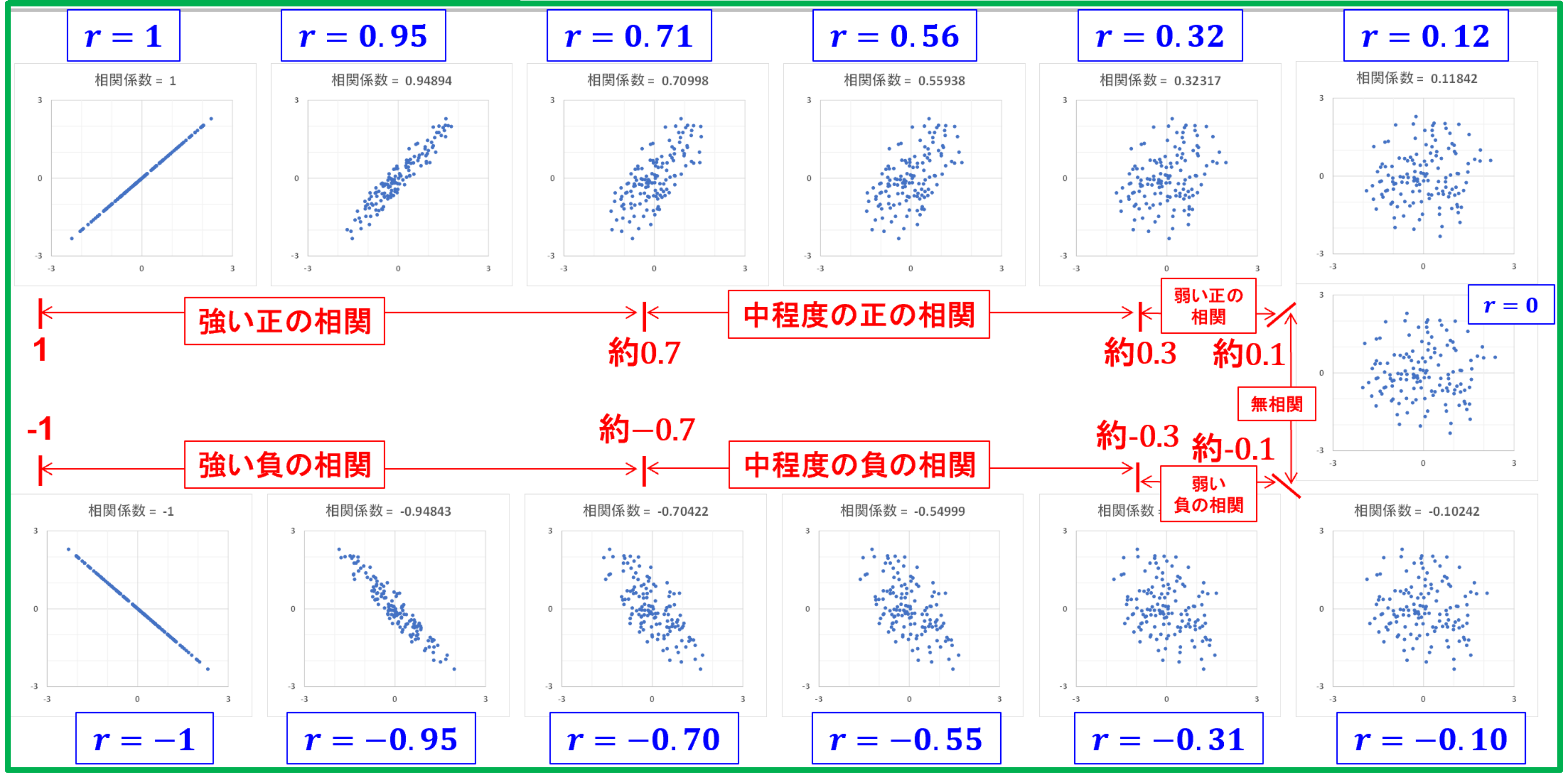
統計検定２級
公式問題集(CBT対応版)の解説
サブカテゴリー10-1．回帰分析の分野
問1 (p142-145) @お試しサイト

(p142.2) [C10-1]問1[1]. 最小二乗法・傾きの検定

[1]:ABランク
[2]:ABランク
[3]:Bランク

小問[1]:　散布図⇒相関係数を選ぶ問題

散布図と相関係数の例



[1]
←左の例を参考にすると、

・p142の散布図：負の相関がある
⇒①②(正の相関)は×

③ほど($r=0.094$)弱くはない

⑤ほどの強い負の相関($r=-0.994$)はない

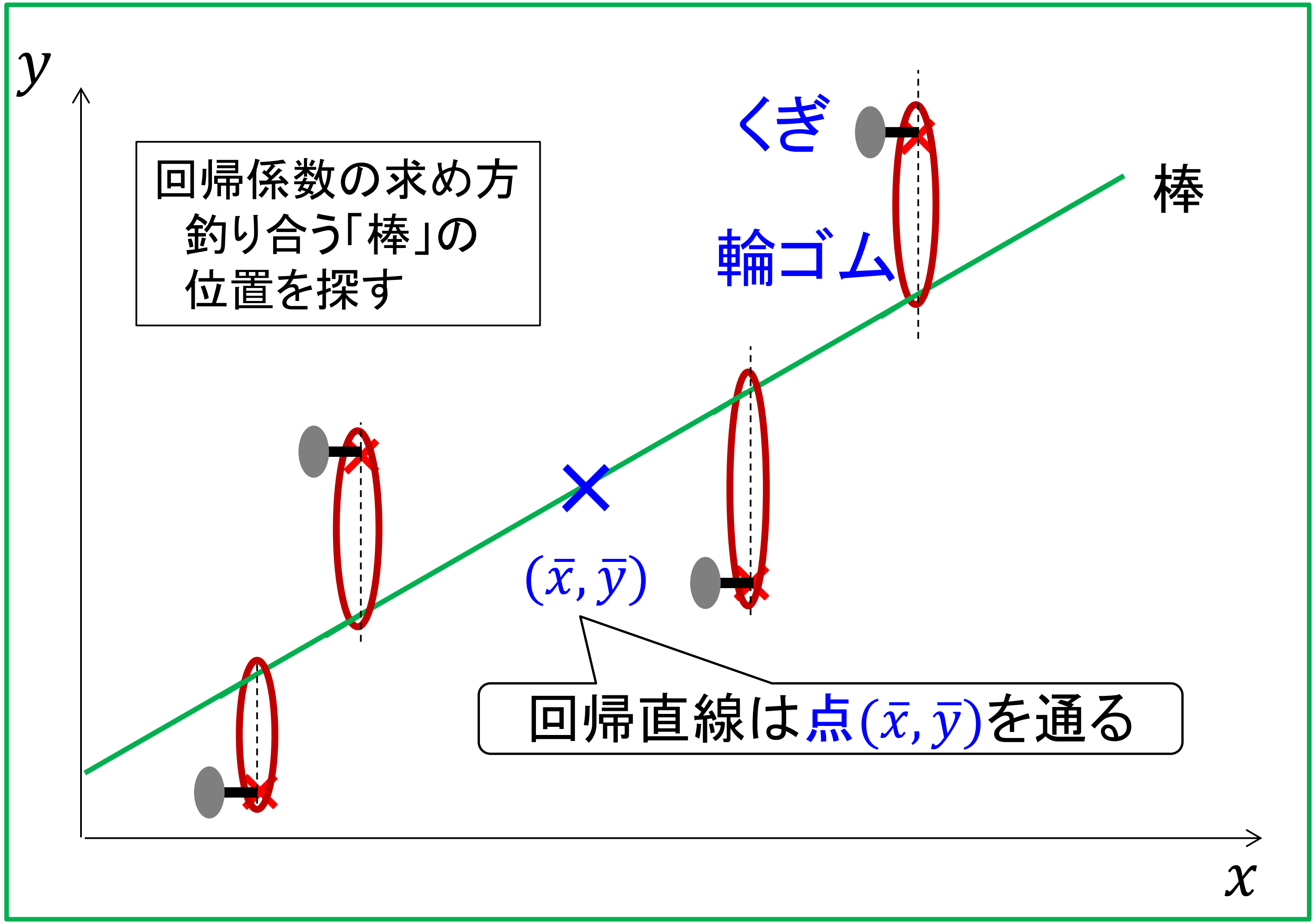
⇒④が適当

[1] (答)④

(p142.3) [C10-1]問1[2]. 最小二乗法・傾きの検定

[1]:ABランク
[2]:ABランク
[3]:Bランク

小問[2]:　回帰直線を選ぶ問題



①:下に寄り過ぎ。
(左上の点から離れすぎ)
③: 左下側に寄り過ぎ。
④: 散布図の右下付近の点からずれ過ぎ。
⑤: 直線の右上側には3点しかなく、
バランスが悪い。

⇒②がバランス面でも、
$(\bar{x}, \bar{y})$を通りそうという観点からも
最もいい。

[2](答)②

(p142.4) [C10-1]問1[3]. 最小二乗法・傾きの検定

[1]:ABランク
[2]:ABランク
[3]:Bランク

小問[3]:　　回帰式:　$y = \beta_1 x + \beta_0$

回帰係数: β_1
(a)推定値(Estimate)= -0.1450
(b)標準誤差(Std. Error)= 0.02316

(c)残差の自由度ϕ_e = $n-2=23$

但し、データ組数 $n=$ $2015-1991+1=25$

(1) 帰無仮説(H_0): $\beta_1 = 0$

帰無仮説(H_0)が正しい時の
検定統計量の値は、

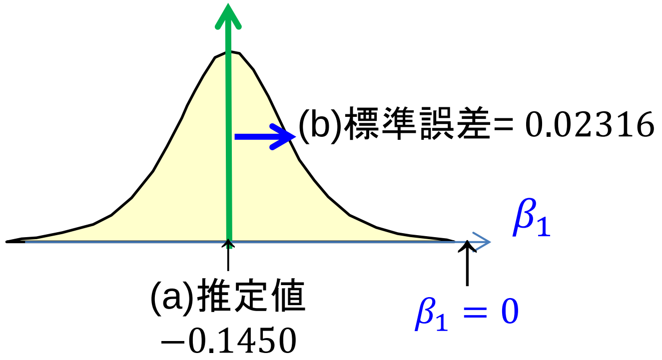
★ $t = \dfrac{(a)\text{推定値} - \beta_1}{(b)\text{標準誤差}} = \dfrac{-0.1450 - \beta_1}{0.02316} = \dfrac{-0.1450 - 0}{0.02316} = -6.2651$

★ 検定統計量が従う分布は・・・ 自由度ϕ_e =23のt分布

(公式) $t = \dfrac{(a)\text{推定値} - \beta_1}{(b)\text{標準誤差}}$ ～自由度ϕ_eのt分布

単回帰分析での残差の自由度: $\phi_e = n - 2$

(n: データ組数) $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$)



仮に、有意水準5%で、
「対立仮説$\beta_1 \neq 0$」として検定すると、
棄却域|t|≧2.069 ⇒ t値は棄却域内にある
⇒H_0を棄却⇒H_1:正しい、 となります。

済

[3]検定統計量(t値)= -6.27
t～自由度23のt分布に従う
[3](答)⑤

(p142.4) [C10-1]問1[3]. 最小二乗法・傾きの検定

[1]:ABランク
[2]:ABランク
[3]:Bランク

小問[3]:　　回帰式:　$y=\beta_1 x+\beta_0$

回帰係数: β_1

(a)推定値(Estimate)= -0.1450

(b)標準誤差(Std. Error)= 0.02316

(c)残差の自由度$\phi_e=$　$n-2=23$

但し、データ組数 $n=$　$2015-1991+1=25$

(1) 帰無仮説(H_0): $\beta_1=0$

帰無仮説(H_0)が正しい時の
検定統計量の値は、

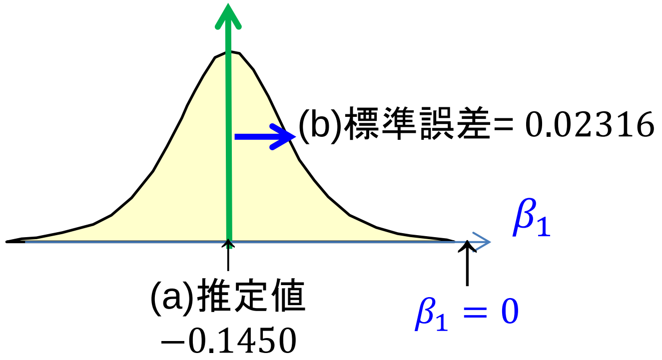
★ $t=\dfrac{(a)\text{推定値}-\beta_1}{(b)\text{標準誤差}}=\dfrac{-0.1450-\beta_1}{0.02316}=\dfrac{-0.1450-0}{0.02316}=-6.2651$

★ 検定統計量が従う分布は･･･ 自由度ϕ_e =23のt分布

(公式)　$t=\dfrac{(a)\text{推定値}-\beta_1}{(b)\text{標準誤差}}$ ～自由度ϕ_eのt分布

単回帰分析での残差の自由度: $\phi_e=n-2$

(n: データ組数) $(x_1,y_1),(x_2,y_2),\dots,(x_n,y_n)$)



仮に、有意水準5%で、
「対立仮説$\beta_1\neq 0$」として検定すると、
棄却域|t|≧2.069 ⇒ t値は棄却域内にある
⇒H_0を棄却⇒H_1:正しい、 となります。

済

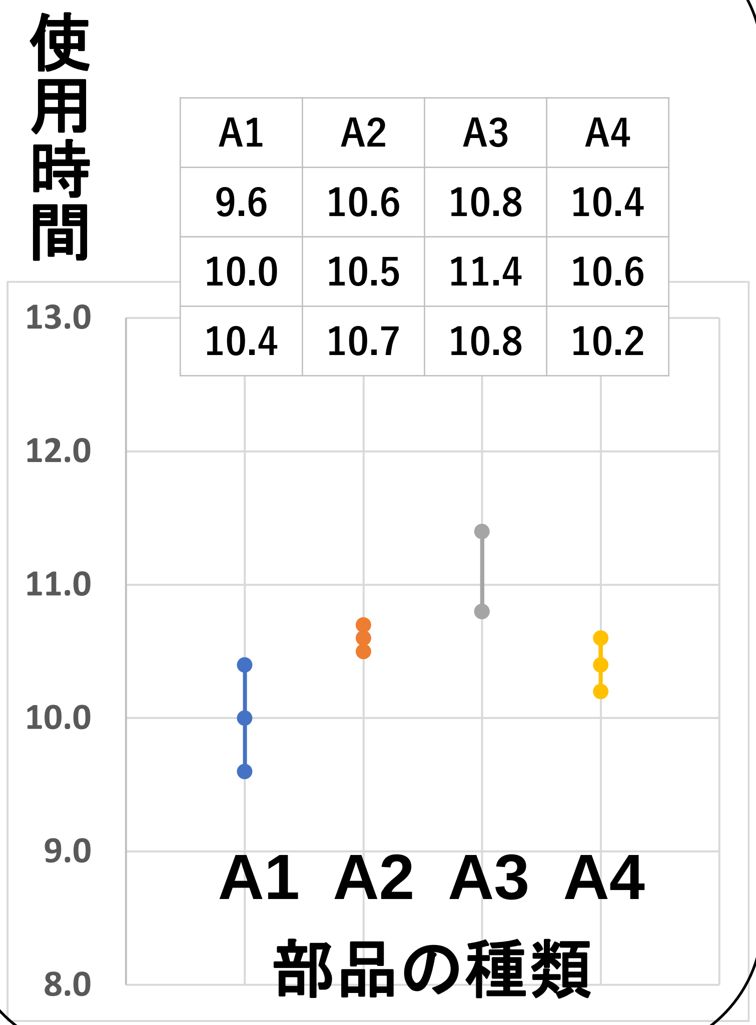
[3]検定統計量(t値)= -6.27
t～自由度23のt分布に従う
[3](答)⑤

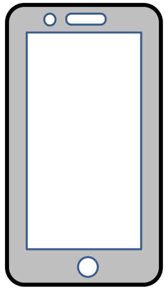
ひかり統計塾

統計検定２級
公式問題集(CBT対応版)の解説
サブカテゴリー10-2：分散分析の分野
問1 (p160-162) @お試しサイト

(p160.1a) [C10-2]分散分析の基礎: 対象とする問題

A1	A2	A3	A4
9.6	10.6	10.8	10.4
10.0	10.5	11.4	10.6
10.4	10.7	10.8	10.2





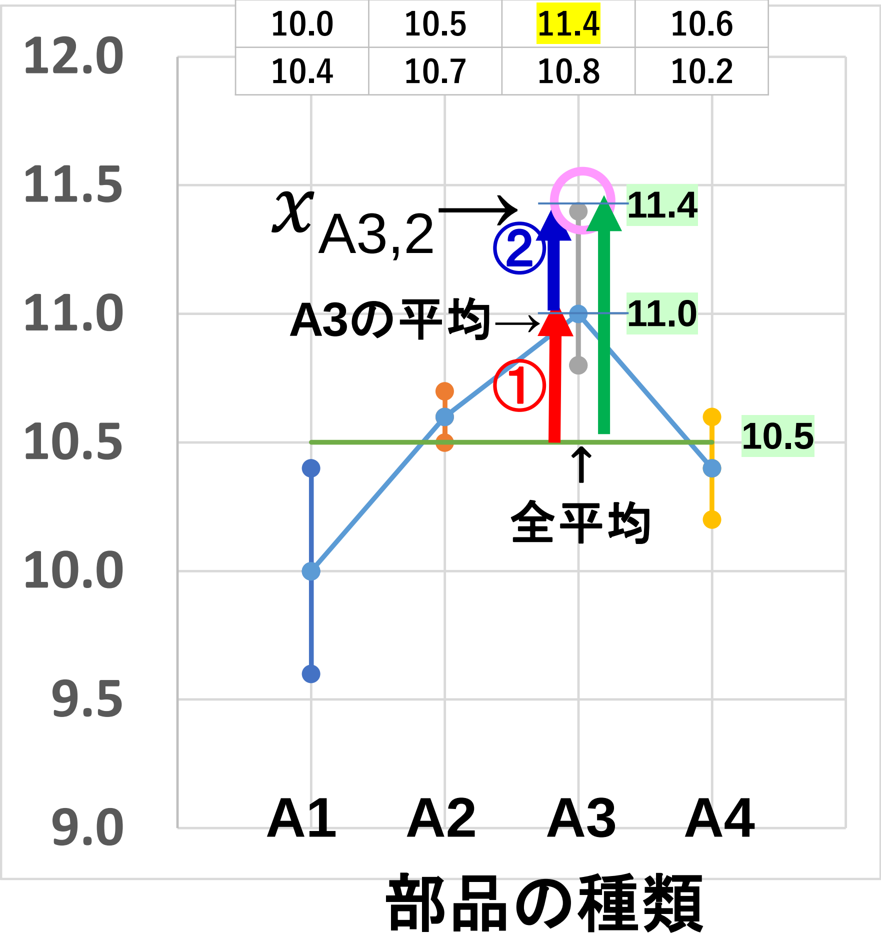
例題:スマホの1充電当たりの使用時間を長くする検討を行っています。ある部品に4種(*)を使い、使用時間を調べました。

(*)部品A1: 旧部品、 部品A2,A3,A4:新部品の候補

それぞれ、3個のサンプルを作成し、使用時間を調べました。
この部品の種類により、使用時間が異なると言えますか？

(p160.1b) [C10-2]分散分析の基礎: ばらつきの分解

A1	A2	A3	A4
9.6	10.6	10.8	10.4
10.0	10.5	11.4	10.6
10.4	10.7	10.8	10.2



ばらつきの総和(偏差平方和, 変動)は、

$$S = (11.4 - 10.5)^2 + \cdots \text{(残り11個分)}$$

$$= \{(11.4 - 11.0) + (11.0 - 10.5)\}^2 + \cdots \text{(残り11個分)}$$

②：ばらつき　①：部品の効果

総平方和

$$S = S_T = S_E + S_A$$

$$\begin{cases} S_E = (11.4 - 11.0)^2 + \cdots \text{(残り11個分)} \\ S_A = (11.0 - 10.5)^2 + \cdots \text{(残り11個分)} \end{cases}$$

②：ばらつき
誤差平方和

①：部品の効果
A間平方和

のように、総平方和を①と②に分けることができました。

$$(x + y)^2 = x^2 + y^2 + 2xy$$

(注) $2 \times (11.4 - 11.0) \times (11.0 - 10.5) + \cdots$ (残り11個分)
が出てきますが、この総和は0になります。(「入門統計解析法」p119[注6.1])

(p160.1c) [C10-2]分散分析の基礎: 自由度

母平均の検定・推定
(母分散:未知)

$$\text{不偏分散} = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2 = \frac{\text{偏差平方和}}{\text{自由度}}$$

今の例では、

データ総数: $4 \times 3 = 12$
全自由度= データ総数 − 1
$\phi_T = n - 1 = 12 - 1 = 11$

A1	A2	A3	A4
9.6	10.6	10.8	10.4
10.0	10.5	11.4	10.6
10.4	10.7	10.8	10.2

4水準
⇒因子Aの自由度= 水準数 − 1
$\phi_A = 4 - 1 = 3$

⇒誤差の自由度= 全自由度 − 因子Aの自由度
$\phi_E = \phi_T - \phi_A$
$= 11 - 3 = 8$

公式: 全自由度(ϕ_T) = データ総数 − 1
因子Aの自由度(ϕ_A) = 水準数 − 1
誤差の自由度(ϕ_E)= $\phi_T - \phi_A$

のように、3つの自由度(ϕ_T, ϕ_A, ϕ_E)を求めます。

(p160.1d) [C10-2]分散分析の基礎: 自由度(練習)

灰色部分には値が有り、白色部分には値が無いとします

3種の自由度を求めてください

繰返し↓

A1	A2	A3	A4	A5
値:有り				

全自由度: $\phi_T = 5 \times 4 - 1 = 19$

因子Aの自由度: $\phi_A = 5 - 1 = 4$　$(A$: 5水準だから$)$

誤差の自由度: $\phi_E = 19 - 4 = 15$

繰返し↓

A1	A2	A3	A4	A5
値:有り				
	値:無し		値:無し	

全自由度: $\phi_T = (5 \times 4 - 2) - 1 = 17$

因子Aの自由度: $\phi_A = 5 - 1 = 4$

誤差の自由度: $\phi_E = 17 - 4 = 13$

公式:
全自由度(ϕ_T) = データ総数 − 1
因子Aの自由度(ϕ_A) = 水準数 − 1
誤差の自由度(ϕ_E)= 全自由度 − 因子Aの自由度 = $\phi_T - \phi_A$

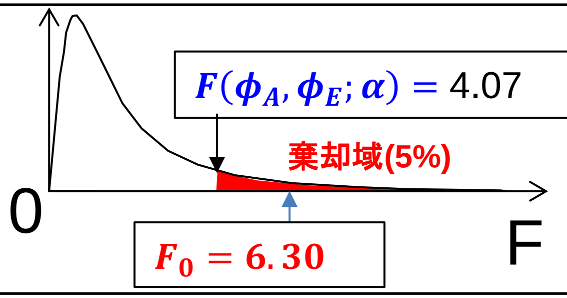
(p160.1e) [C10-2]分散分析の基礎: 検定の手順

(1): 2つの仮説を立てます
値の説明:(スマホの使用可能時間)要因A:(部品の種類)
帰無仮説 H_0: 要因Aの効果はない
対立仮説 H_1: 要因Aの効果はある

理論:帰無仮説が正しい時、検定統計量: $F_0 = \frac{V_A}{V_E}$は自由度(ϕ_A, ϕ_E)のF分布に従う

(2): 有意水準αを定めます: α= 0.05

A1	A2	A3	A4
9.6	10.6	10.8	10.4
10.0	10.5	11.4	10.6
10.4	10.7	10.8	10.2



(3):棄却域を定めます
データ総数(n)=(12),要因Aの水準数=(4)
全自由度: $\phi_T = 12 - 1 = 11$
因子Aの自由度: $\phi_A = 4 - 1 = 3$
誤差の自由度: $\phi_E = 11 - 3 = 8$
棄却域: $F > F(\phi_A, \phi_E; \alpha) = F(3, 8; 0.05) = 4.07$

(5): H_0を棄却する・しないの判定
F_0 =(6.30) は棄却域に(ある ない)

どっちですか?

(●) $F = F_0$が棄却域にあるので、H_0を棄却する ⇒ H_1は正しいと言える
( ) $F = F_0$が棄却域にないので、H_0は棄却できない ⇒ H_1は正しいかどうか何も言えない

(4)分散分析表を作成し、検定統計量を求めます。

★超重要

要因	(a)平方和 S		(b)自由度 φ		(c)平均平方(分散) V=S/φ	(d)F_0 =V_A/V_E	(d),(e)の大小比較	(e)棄却域の下限
A(因子)	1.56	÷	3	=	0.52	6.30	>	4.07
E(誤差)	0.66	÷	8	=	0.0825			
計	2.22		11					

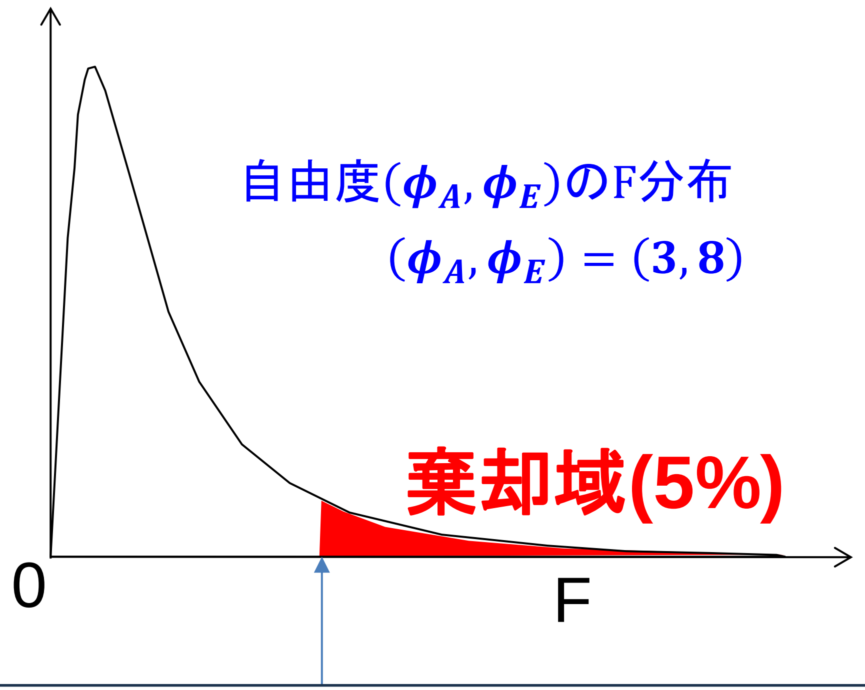
検定統計量

(p160.1f) [C10-2]分散分析の基礎: F分布表

F分布:

P= 0.05

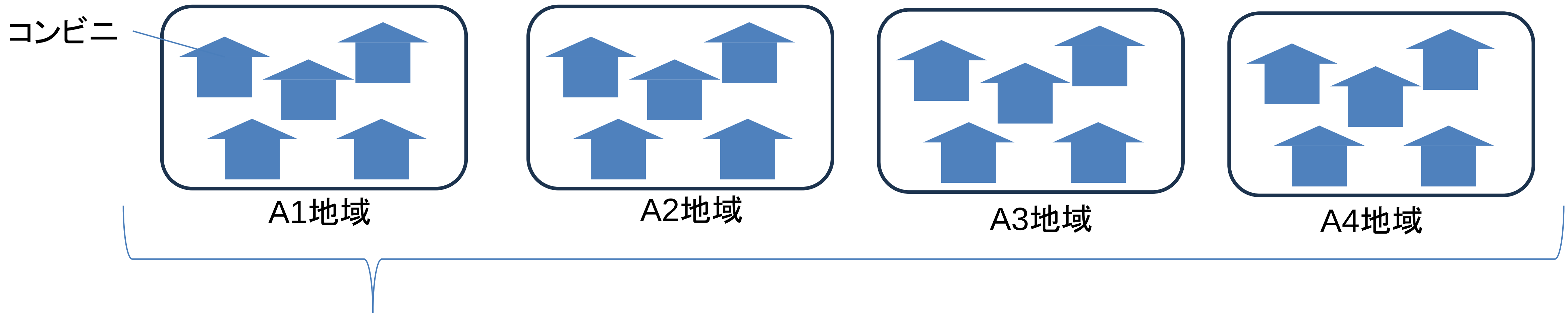
		ϕ1							
		1	2	3	4	5	6	7	8
ϕ2=	1	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88
	2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37
	3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85
	4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04
	5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82
	6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15
	7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73
	8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44
	9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23
	10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07
	11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95
	12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85



$F(\phi_A, \phi_E; \alpha) = F(3, 8; 5\%)$ $= 4.07$

(p160.2) [C10-2]問1. 一元配置分散分析の基本

[1]:ABランク
[2]:ABランク
[3]:ABランク



地域ごとの売上げに違いがあるか、
を1元配置分散分析で調べる

売上げ

地域	A1	A2	A3	A4
5店舗 (繰り返し=5)				

Q: 以下の自由度はいくらでしょう？

全自由度: $\boldsymbol{\phi_T = 4 \times 5 - 1 = 19}$

因子Aの自由度: $\boldsymbol{\phi_A = 4 - 1 = 3}$ (ア)

誤差の自由度: $\boldsymbol{\phi_E = 19 - 3 = 16}$ (イ)

公式: 全自由度($\boldsymbol{\phi_T}$) = データ総数 − 1
因子Aの自由度($\boldsymbol{\phi_A}$) = 水準数 − 1
誤差の自由度($\boldsymbol{\phi_E}$)= $\boldsymbol{\phi_T - \phi_A}$

(p160.3) [C10-2]問1[1][2]. 一元配置分散分析の基本

[1]:ABランク
[2]:ABランク
[3]:ABランク

分散分析表　⇒(ウ)、(エ)、(オ)を求めてください

要因	平方和S	自由度φ	平均平方V	F_0
A(因子)	0.2204	**(ア)　3**	**(ウ)　0.0735**	**(オ)　3.49**
E(誤差)	0.3370	**(イ)　16**	**(エ)　0.0211**	
計	0.5574	19	**0.0293**	

[2](答)は？　・・・　⑤

(公式)分散分析表@一元配置分散分析

要因	平方和S		自由度φ		平均平方V		F_0
A(因子)	S_A	÷	ϕ_A	=	$V_A=S_A/\phi_A$	≠	V_A/V_E
E(誤差)	S_E	÷	ϕ_E	=	$V_E=S_E/\phi_E$		
計	S_T		ϕ_T				

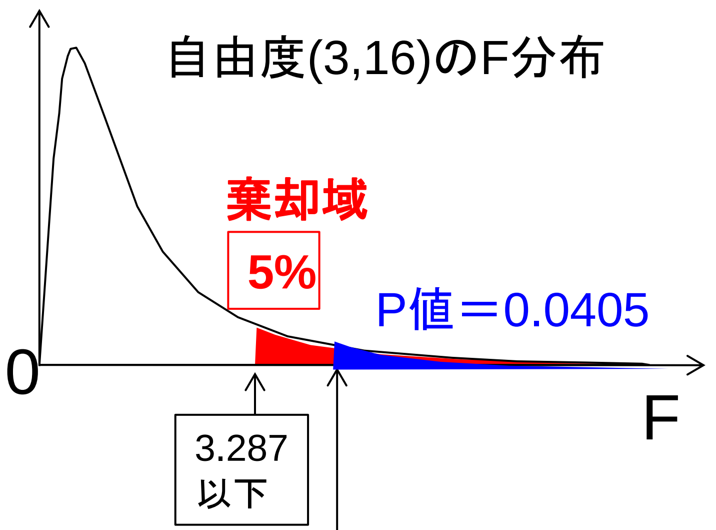
[1]不偏分散は？

$$\text{不偏分散} = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2 = \frac{\text{偏差平方和}}{\text{自由度}} = \frac{S_T}{\phi_T} = \frac{0.5574}{19} = 0.0293$$

[1](答)①

(p160.4) [C10-2]問1[3]. 一元配置分散分析の基本

[1]:ABランク
[2]:ABランク
[3]:ABランク



F_0=3.49

帰無仮説　H_0:　売上は地域に依存しない ($\mu_1=\mu_2=\mu_3=$･･･)
対立仮説　H_1:　売上は地域に依存する

Q: どちら？
- ($\mu_1,\mu_2,\ldots$の少なくとも1つが異なる)⇒①②
- ~~($\mu_1,\mu_2,\ldots$の全てが異なる)⇒③④~~

検定統計量：F_0=3.49
有意水準：5%（[3]の問題文に記載）

Q: 帰無仮説H_0は棄却されますか？
(問題をよく見てください)

⇒棄却される
[3]の候補：①,③

P=　0.05

	φ1 1	2	3	4
φ2= 5	6.608	5.786	5.409	5.192
10	4.965	4.103	3.708	3.478
15	4.543	3.682	3.287	3.056
20	4.351	3.493	3.098	2.866
25	4.242	3.385	2.991	2.759
30	4.171	3.316	2.922	2.690

（解法1）p160の分散分析表の右端のP値=0.0405 ＜0.05 (有意水準)

（解法2）自由度(3,16)の上側5%点･･･p203の表にはない
(3,15):3.287，(3,20):3.098 ⇒　この間にあるはず
F_0=3.49は、棄却域に入っている

⇒売上げは地域により違いがあると言える

[3](答)①

済

ひかり統計塾

統計検定２級
公式問題集(CBT対応版)の解説
PART. ３　模擬テスト
問1,2 (p176-177) @お試しサイト

(p176) 問1. 箱ひげ図からの読み取り

(Aランク)

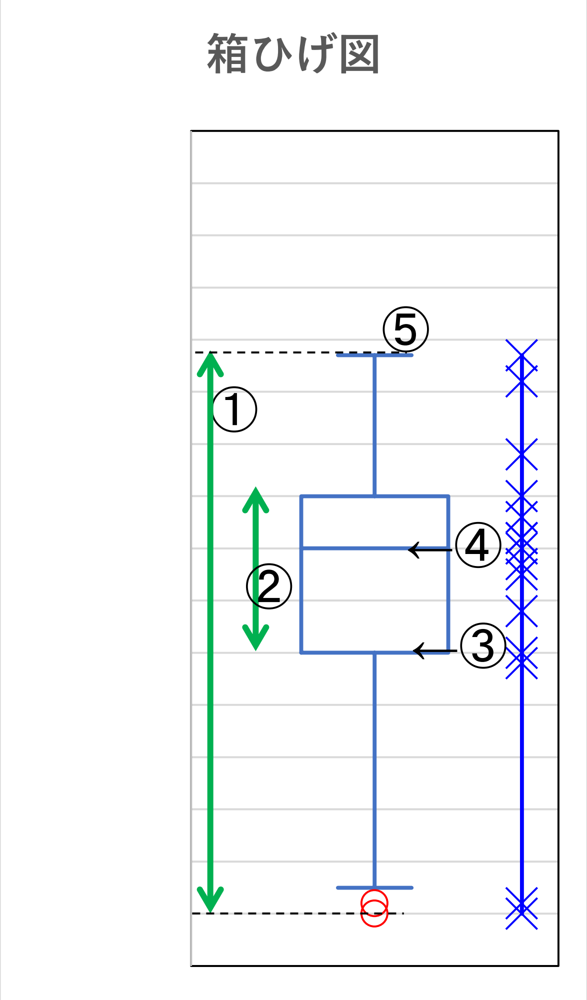
①(少なくとも)広島よりも福岡の方が範囲が広いので、×

② (少なくとも)名古屋よりも東京が
四分位範囲が小さいので、×

③正しい、〇

④中央値は、大阪より名古屋が小さいので、×

⑤ 最大値:東京ではなく名古屋が最小なので、×



箱ひげ図から、「範囲」「四分位範囲」「第1,3四分位数」「中央値」
「最大値」「最小値」を読み取れるようにしておいてください。

(答) ③

済

(p177) 問2(1). 標準化得点の性質

(Bランク)

参考:(p34)[カテゴリー1]問7. 線形変換による平均・標準偏差

$x_1, \dots, x_n \xrightarrow{\text{標準化}} z_1, \dots, z_n$

$\downarrow y_i = ax_i + b \ (a \neq 0)$

$y_1, \dots, y_n \xrightarrow[(a>0)]{\text{標準化}} z_1, \dots, z_n$

$y_1, \dots, y_n \xrightarrow{(a<0)} -z_1, \dots, -z_n$

$\bar{C} = \frac{1}{n}\sum_{i=1}^{n} C_i$

$V_C = \frac{1}{n-1}\sum_{i=1}^{n}(C_i - \bar{C})^2$

$n = 17$

$\bar{C} = 2.4, V_C = 7.0^2$

平均:$\bar{C}$,

不偏分散:V_C、対応する標準偏差:$\sqrt{V_C}$

・摂氏での温度:C_i

$F_i = 1.8C_i + 32$

$\bar{F} = 1.8\bar{C} + 32$

$V_F = 1.8^2 V_C$

華氏での温度:F_i

平均:$\bar{F}$,

不偏分散:V_F、対応する標準偏差:$\sqrt{V_F}$

$\bar{F} = \frac{1}{n}\sum_{i=1}^{n} F_i$

$V_F = \frac{1}{n-1}\sum_{i=1}^{n}(F_i - \bar{F})^2$

$\bar{z} = \frac{1}{n}\sum_{i=1}^{n} z_i = 0$

$V_z = \frac{1}{n-1}\sum_{i=1}^{n}(z_i - \bar{z})^2 = 1$

$n = 17$

$\Rightarrow$ I:○

z_iの平均 $= 0$

z_iの不偏分散 $= 1$

標準化得点:$z_i = \dfrac{C_i - \bar{C}}{\sqrt{V_C}}$

標準化得点:$w_i = \dfrac{F_i - \bar{F}}{\sqrt{V_F}}$

w_iの平均 $= 0$

w_iの不偏分散 $= 1$

$\bar{w} = \frac{1}{n}\sum_{i=1}^{n} w_i = 0$

$V_w = \frac{1}{n-1}\sum_{i=1}^{n}(w_i - \bar{w})^2 = 1$

Q(II): z_iの最大値は？2.5より小さい？

摂氏温度での最大値: C_i =22℃@マイアミ

$C_i = 22 \Rightarrow z_i = \dfrac{C_i - \bar{C}}{\sqrt{V_C}} = \dfrac{22 - 2.4}{7} = 2.8 > 2.5$

$\Rightarrow$ II: ×

$\bar{C} = 2.4, V_C = 7.0^2$

Q(III): $z_i = w_i$ ？

$w_i = \dfrac{F_i - \bar{F}}{\sqrt{V_F}} = \dfrac{(1.8C_i + 32) - (1.8\bar{C} + 32)}{\sqrt{1.8^2 V_C}}$

$= \dfrac{1.8(C_i - \bar{C})}{1.8\sqrt{V_C}} = \dfrac{C_i - \bar{C}}{\sqrt{V_C}} = z_i$

$\Rightarrow$ III:○

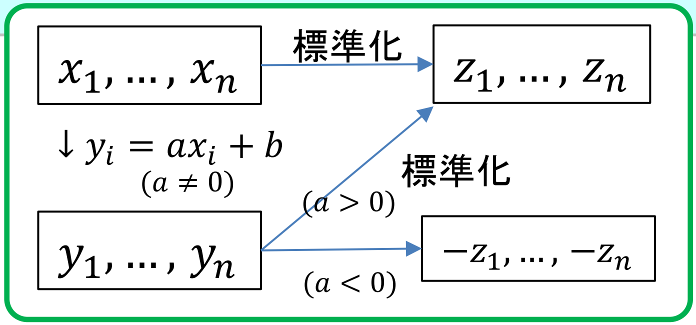
済

I:○、II: ×、III:○

(答) ④

(p177) 問2(2).標準化得点の性質

(Bランク)

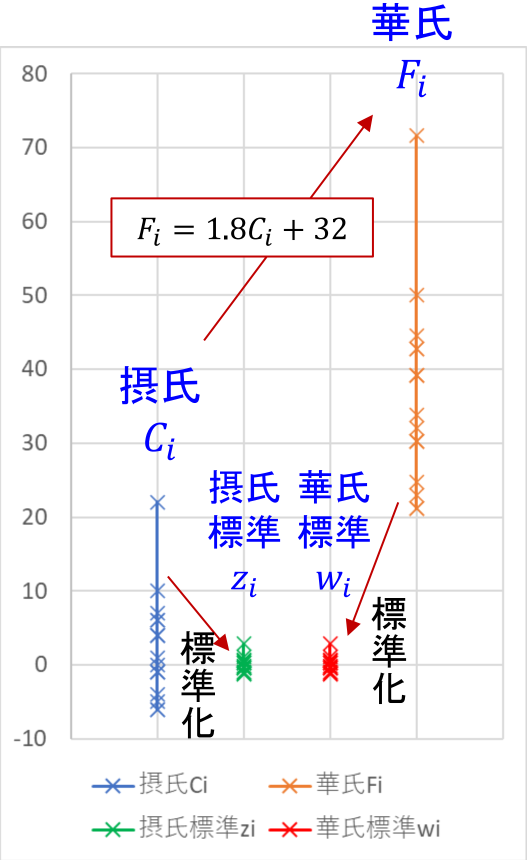


I: z_iの平均=0, 不偏分散=1

平均	2.3529	36.2353	0.0000	0.0000
不偏分散	48.4926	157.1162	1.0000	1.0000
標準偏差	6.9637	12.5346	1.0000	1.0000
No.	**摂氏(C_i)**	**華氏(F_i)**	**摂氏標準z_i**	**華氏標準w_i**
1	1	33.8	-0.194286	-0.194286
2	-6	21.2	-1.199503	-1.199503
3	6	42.8	0.523727	0.523727
4	4	39.2	0.236522	0.236522
5	-6	21.2	-1.199503	-1.199503
6	-4	24.8	-0.912298	-0.912298
7	-1	30.2	-0.481491	-0.481491
8	4	39.2	0.236522	0.236522
9	-1	30.2	-0.481491	-0.481491
10	-1	30.2	-0.481491	-0.481491
11	4	39.2	0.236522	0.236522
12	-5	23	-1.055901	-1.055901
13	6	42.8	0.523727	0.523727
14	22	71.6	2.821367	2.821367
15	7	44.6	0.667329	0.667329
16	10	50	1.098137	1.098137
17	0	32	-0.337888	-0.337888

II: z_iの最大値は2.5より大きい

III: $z_i = w_i$



標準化得点を拡大