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 (p126-141)

カイ二乗検定の分野

(9-1: 適合度検定の分野，9-2: 独立性検定の分野)

統計検定2級　CBT問題集　PART.2 目次

[C9]
[CATEGORY.9]
カイ二乗検定の分野

他に比べ、
比較的簡単です。

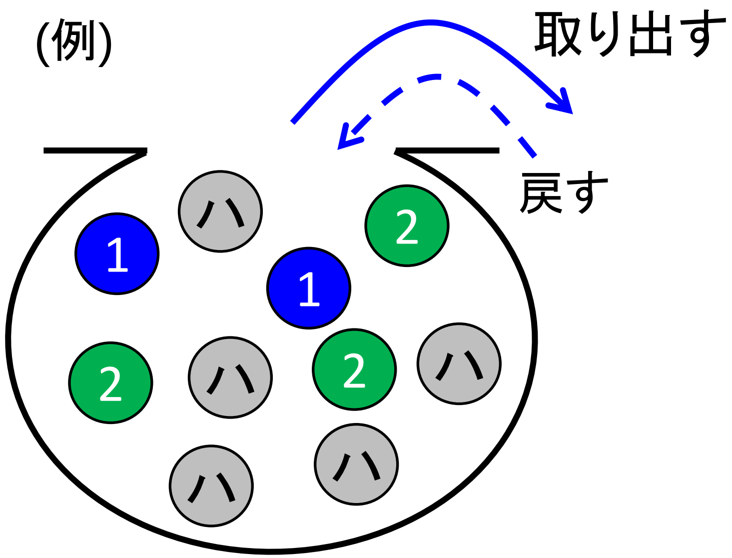
[C9-1] 適合度検定の分野
[C9-2] 独立性検定の分野

(p126.1)[C9-1]問1. 適合度検定の基本

(ABランク)

くじ引き

(例)

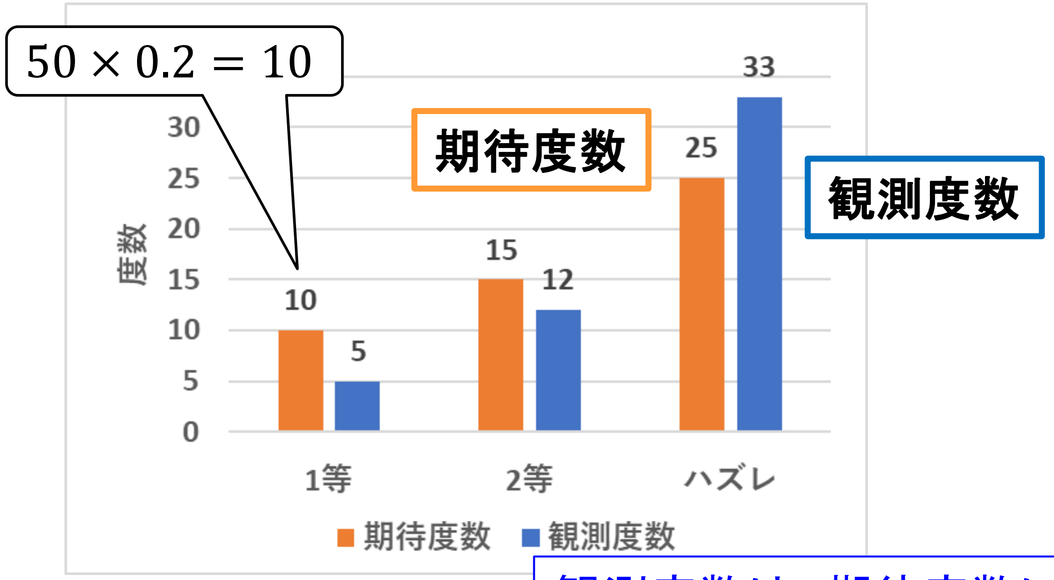


想定されている割合

1等：20%　2個
2等：30%　3個
ハズレ:50%　5個

袋中の玉が10個の場合

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ϕ=k—1
・検定統計量(χ^2統計量)は、以下で計算

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

・自由度ϕのχ^2分布を用いて検定する

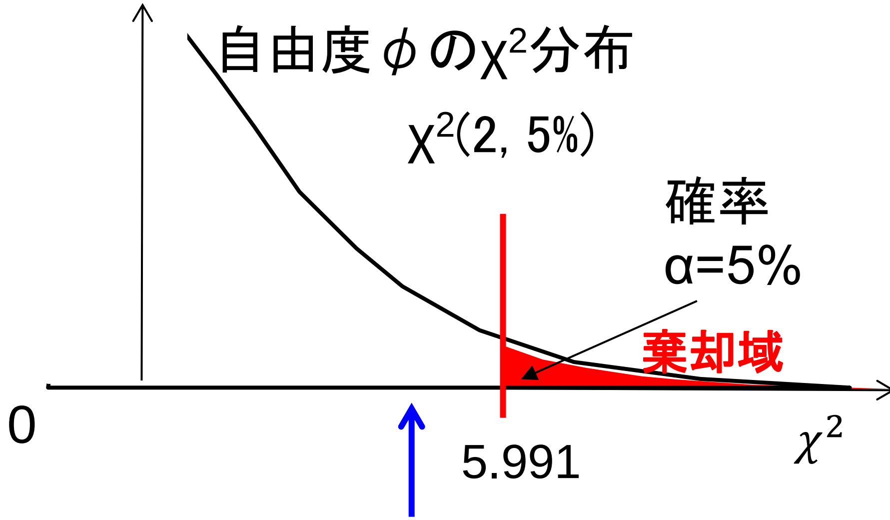
(ア) 自由度ϕ= **2**

(イ)検定統計量: χ_0^2 = **5.66**

(答) ④

χ²分布の上側P点

	P=	0.975	0.950	0.050	0.025
ϕ=	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



$\chi_0^2=5.66$

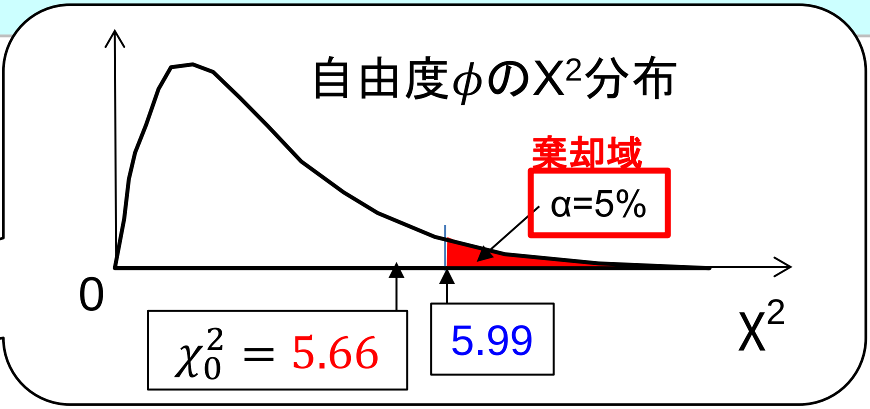
帰無仮説(H_0)は棄却できない ⇒(ウ)
対立仮説(H_1)が正しいかどうか
何も言えない

(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)

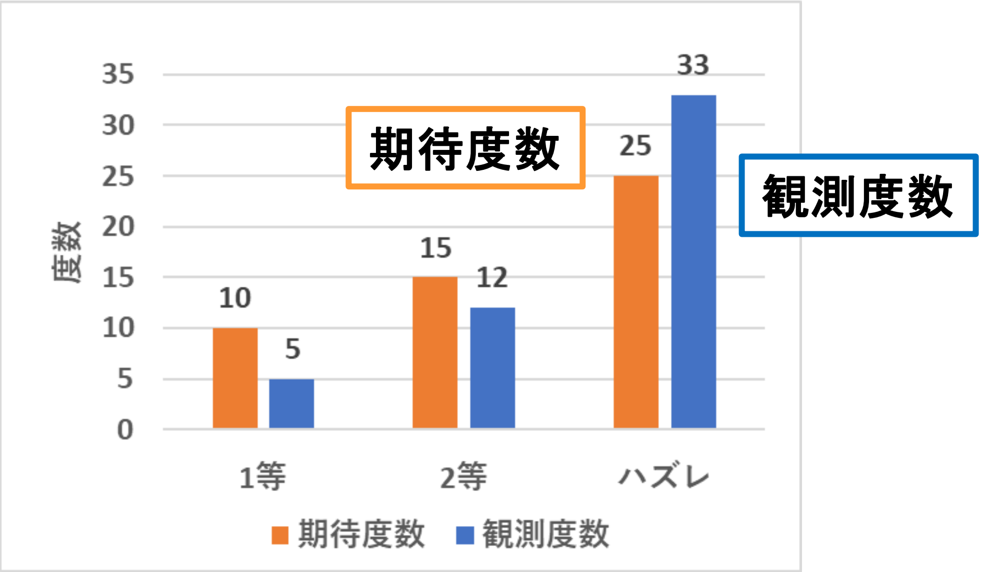
	1等	2等	ハズレ	合計
観測度数	**5**	**12**	**33**	**50**
期待されている割合	**0.2**	**0.3**	**0.5**	
期待度数	(50×0.2=)**10**	(50×0.3=)**15**	(50×0.5=)**25**	**50**
$\frac{(観測度数-期待度数)^2}{期待度数}$	$\frac{(5-10)^2}{10}$	$\frac{(12-15)^2}{15}$	$\frac{(33-25)^2}{25}$	**5.66**

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

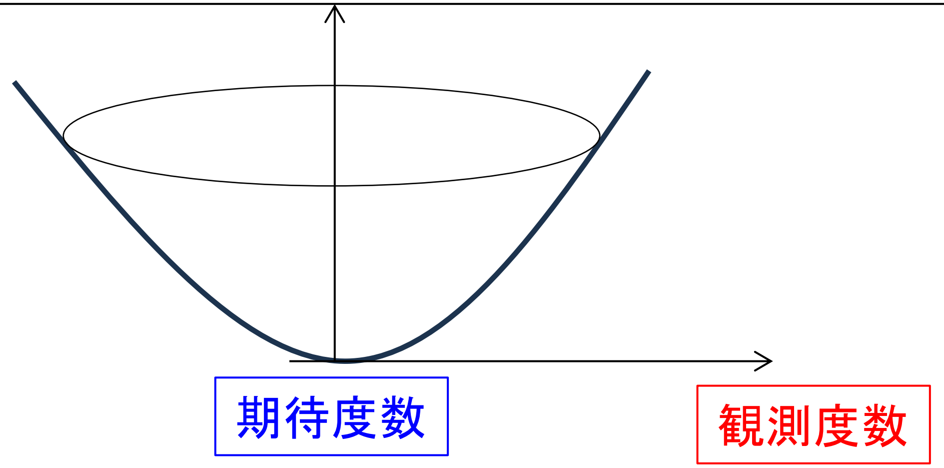
	1等	2等	ハズレ	合計
人数	★	★		150

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)

(p128.1)[C9-1]問2. 一様性の適合度検定

(ABランク)

帰無仮説(H_0):交通事故の発生率は曜日に依存しない
対立仮説(H_1):交通事故の発生率は曜日に依存する

を調べる

χ^2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

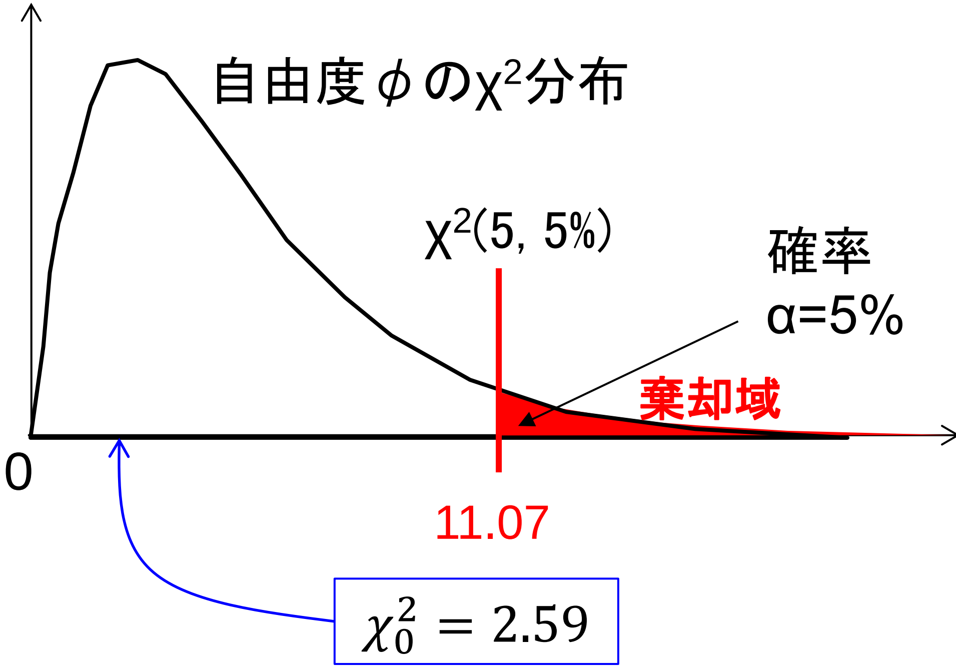
	月	火	水	木	金	土	合計
観測度数	14	19	15	22	16	16	102
期待される割合	1/6	1/6	1/6	1/6	1/6	1/6	
期待度数	17	17	17	17	17	17	102
$\frac{(\text{観測度数}-\text{期待度数})^2}{\text{期待度数}}$			計算不要				2.59

$102 \times \frac{1}{6} = 17$

カテゴリー数:k = 6

⇒自由度ϕ=k-1= 5

結果が与えられている場合も多いです。
問題をよく読んでください



(ア)自由度ϕ=5のχ^2分布
⇒②or③

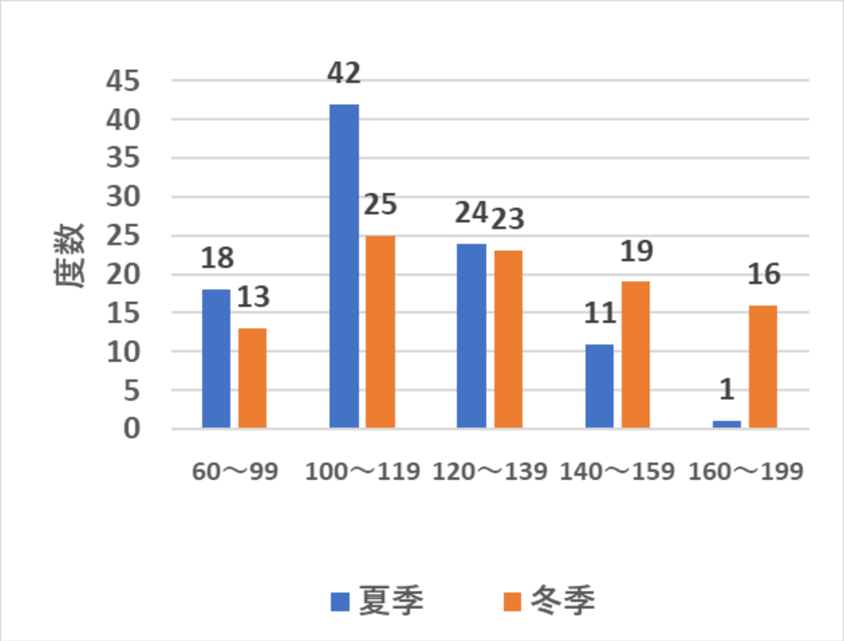
(イ)帰無仮説を棄却しない

(答) ③

済

(p130.1)[C9-1]問3. 同等性の検定

(ABランク)



夏季と冬季で、月毎のドライバーの死者数の分布に
違いがあるか？の検定

	60～99	100～119	120～139	140～159	160～199	合計
夏季	18	42	24	11	1	96
冬季	13	25	23	19	16	96
合計	31	67	47	30	17	192

サブカテゴリー9-1:「適合度の検定」 を使う場合、
既知の分布(一様分布、くじの想定比率、ポアソン分布、・・・)の
情報があるべき。 しかし、情報はない。

サブカテゴリー9-2:「独立性の検定」
季節(「夏季」「冬季」)で発生数の分布が違うと言えるかは
調べられる
⇒今回、この手法で検定を実施する

本問題に取り組まれる際は、事前に、
サブカテゴリー9-2:「独立性の検定」の学習をなさってください。

(p130.1)[C9-1]問3. 同等性の検定

(ABランク)

帰無仮説(H_0): 2つの分布が同等である
対立仮説(H_1): 2つの分布が同等ではない

$m = 5$, $k = 2$

	60～99	100～119	120～139	140～159	160～199	合計
夏季	18	42	24	11	1	96
冬季	13	25	23	19	16	96
合計	31	67	47	30	17	192

Q1. この□の期待度数は？ $67 \times \frac{96}{192} = 33.5$ (ア) ⇒ 候補:③④

Q2. 自由度は？ $\phi = (2-1) \times (5-1) = 4$

Q3. 棄却域の境界値は？ 9.488

Q4. χ^2値は棄却域にある？ ある

⇒帰無仮説を棄却する
⇒対立仮説は正しいと言える
⇒季節により分布は異なると言える

⇒(答)④ 済

公式

m, k

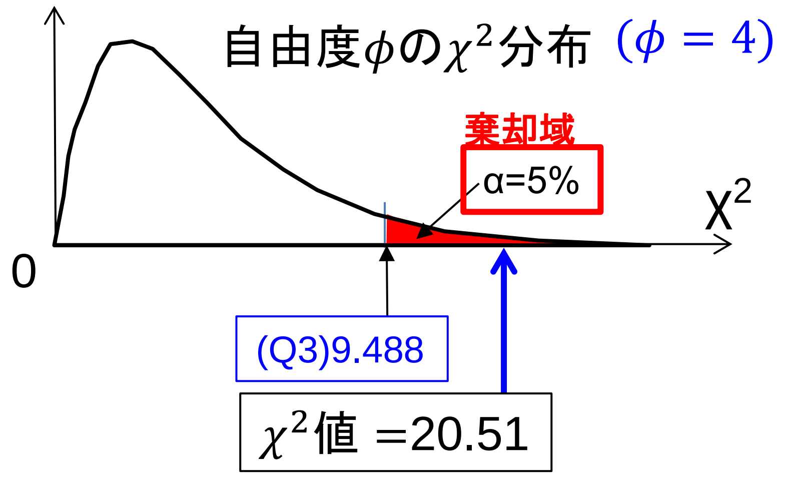
	カテゴリー1	･･･	カテゴリーm	計
グループ1	(d)			(c)
･･･				･･･
グループk				
計	(a)	･･･		(b)

期待度数の求め方: $(d) = (c) \times \frac{(a)}{(b)}$

自由度: $\phi = (k-1) \times (m-1)$

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



(p132.1)[C9-1]問4. ポアソン分布の当てはめ (ABランク)

帰無仮説(H_0): 1日当たりの交通事故死者数が、λ=2のポアソン分布に従っている
対立仮説(H_1): 1日当たりの交通事故死者数は、λ=2のポアソン分布に従っていない

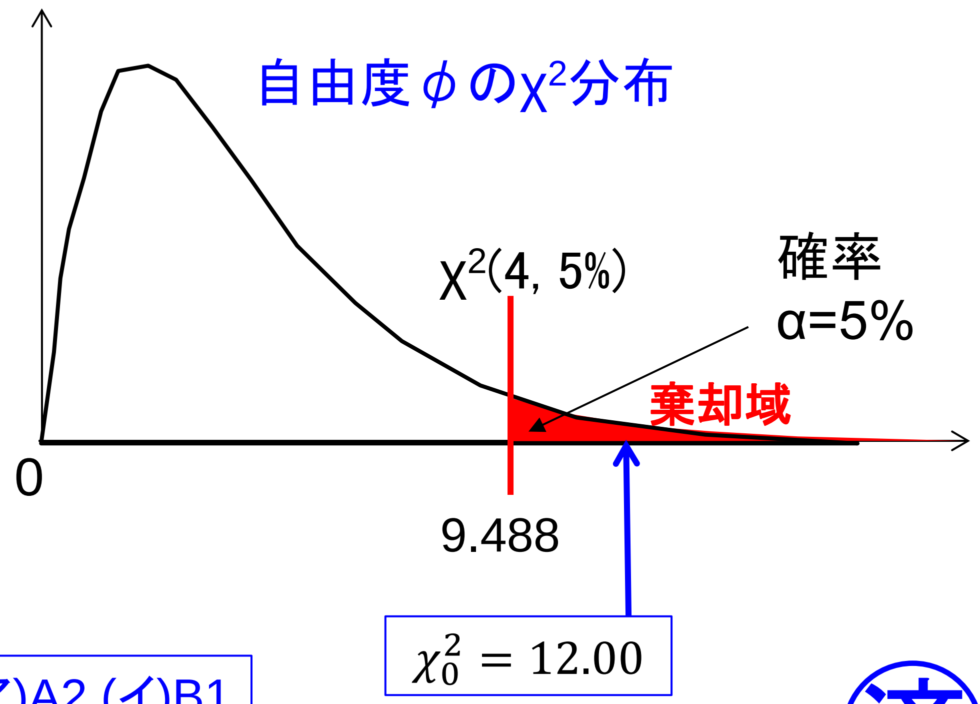
死者数(人)	0	1	2	3	4以上	合計
観測度数						
期待度数						
$\frac{(\text{観測度数}-\text{期待度数})^2}{\text{期待度数}}$						12.00

カテゴリー数: $k = 5$

自由度= $k-1=5-1=4$ ⇒(ア) A2: 4

χ^2分布の上側P点

	P=	0.975	0.950	0.050	0.025
$\phi=$	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



χ_0^2値は、棄却域にあるので、H_0を棄却できる。
H_1が正しいと言える

⇒(イ)B1

(ア)A2,(イ)B1
(答)③

済

[C9]
[CATEGORY.9]
カイ二乗検定の分野

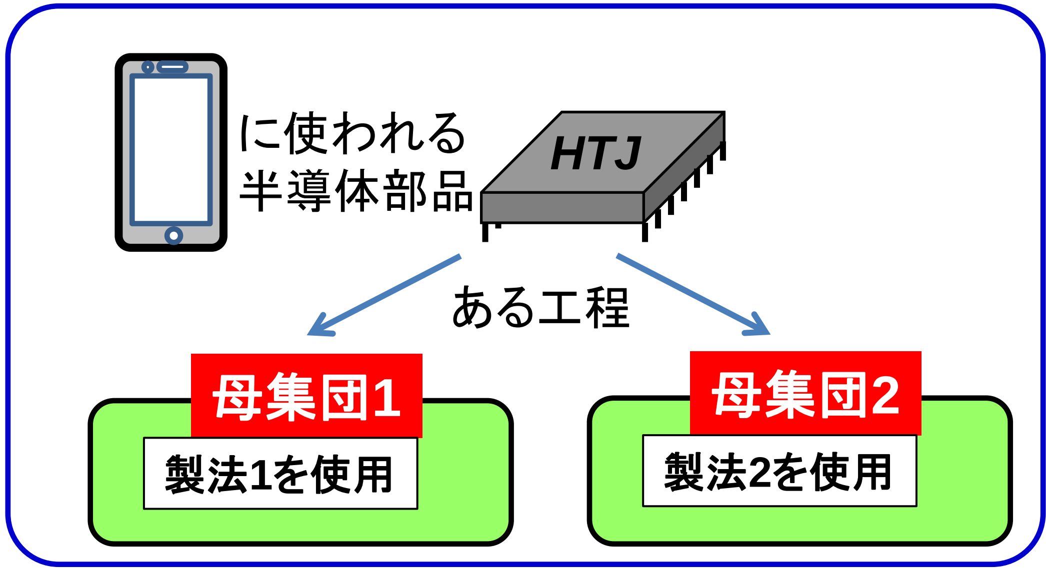
[C9-1] 適合度検定の分野

[C9-2] 独立性検定の分野

一般に、
「分割表」
「クロス集計表」
による検定と
呼ばれます

(p134.0.2)分割表の基本：例題

Q:　以下の3つのケースで、
不良品の発生は、
製法により違いがあると言えますか？



(ケース1)

	良品	不良品
製法1	10	10
製法2	10	10

違いはないですね。

(ケース2)

	良品	不良品
製法1	12	8
製法2	8	12

たまたま差が出たのか、
本質的に違うのかは･･･
どうでしょうか？

(ケース3)

	良品	不良品
製法1	19	1
製法2	1	19

違いがありますね。

(p134.0.3)分割表の基本：考え方

(ケース1)

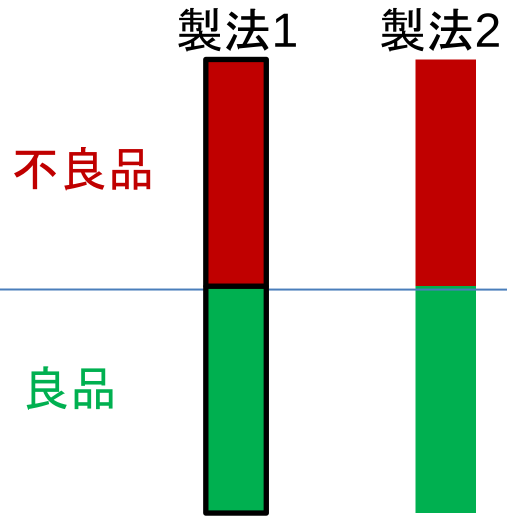
	良品	不良品
製法1	10	10
製法2	10	10

(ケース2)

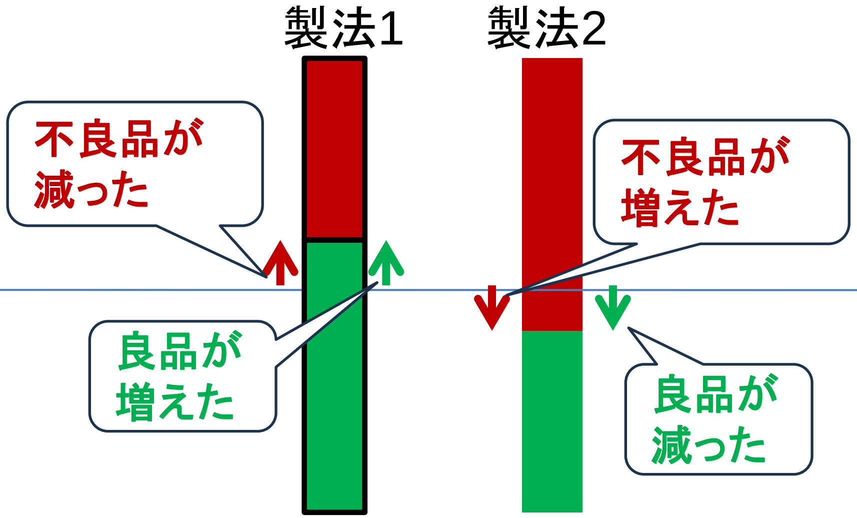
	良品	不良品
製法1	12	8
製法2	8	12

(ケース3)

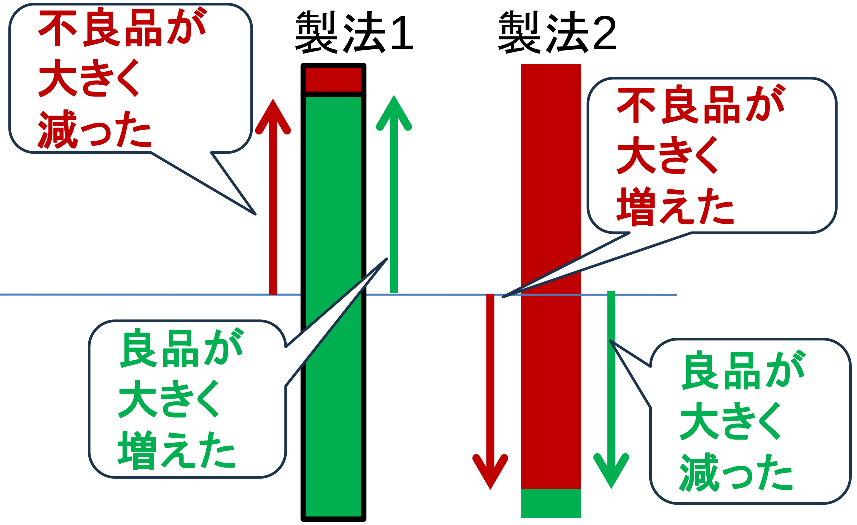
	良品	不良品
製法1	**19**	1
製法2	1	**19**



製法1,2で
良品・不良品の発生に
差がない時に
期待される状態



製法1,2で
良品・不良品の発生に
差がある状態



製法1,2で
良品・不良品の発生に
大きな差がある状態

「期待される状態」と「その状態からの差の大きさ」を考えると良さそう

(p134.0.4)分割表の基本：手順：(1)期待度数

①製法1では、良品が多い

①製法2では、不良品が多い

	良品	不良品	計
製法1	12	8	20
製法2	8	12	20
計	20	20	40

②製法1,2の全体では、差が無い

公式

期待度数の求め方 : $(d) = (c) \times \frac{(a)}{(b)}$

	カテゴリー1	・・・	計
グループ1	(d)		(c)
・・・			・・・
計	(a)	・・・	(b)

比でも計算可 : $(d):(c) = (a):(b)$

良品・不良品の発生が製法に関係ない場合に期待される発生個数(**期待度数**)を求めましょう

	良品	不良品	計
製法1	10	10	20
製法2	10	10	20
計	20	20	40

③全体では、20/40=50%が良品

③全体では、20/40=50%が不良品

④製法1で作られた20個中、

$20 \times \frac{20}{40} = 10$個が良品に、

$20 \times \frac{20}{40} = 10$個が不良品となるはず

⑤同様に、製法2で作られた20個中、

$20 \times \frac{20}{40} = 10$個が良品に、

$20 \times \frac{20}{40} = 10$個が不良品となるはず

(p134.0.5)分割表の基本：手順：(2)χ^2統計量

実度数
(観測度数)

	良品	不良品	計
製法1	12	8	20
製法2	8	12	20
計	20	20	40

期待度数

製法1,2間で
不良品の発生に
差がない時に
期待される度数

	良品	不良品	計
製法1	10	10	20
製法2	10	10	20
計	20	20	40

製法が2種類、良品・不良品で2種類、
計2×2=4個分の和を求めます

$$\chi_0^2=\frac{(12-10)^2}{10}+\frac{(8-10)^2}{10}+\frac{(8-10)^2}{10}+\frac{(12-10)^2}{10}=1.6$$

$$\chi_0^2=\sum\frac{(\text{実度数}-\text{期待度数})^2}{\text{期待度数}}$$

ピアソンのχ^2統計量を計算します

(p134.0.6)分割表の基本：手順：(3) 自由度

カテゴリー数: b

グループ数: a

	超良品	良品	・・・	計
製法1	12	8		30
製法2	8	12		30
・・・				
計	30	30		90

統計理論によると、
ピアソンのχ^2統計量：

$$\chi_0^2 = \sum \frac{(\text{実度数} - \text{期待度数})^2}{\text{期待度数}}$$

は自由度$\phi = (a-1)(b-1)$の
χ^2分布に近似的に従うことが
知られています。

合計値：固定の時、
★を決めると
すべて決まります。
★の数=自由度

$b = 3$, $a = 3$

	超良品	良品	不良品	計
製法1	★	★		39
製法2	★	★		10
製法3				11
計	29	20	11	60

自由度：　$\phi = (a-1)(b-1) = (3-1) \times (3-1) = 4$

$b = 2$, $a = 2$

	良品	不良品	計
製法1	★		20
製法2			20
計	20	20	40

Q. この表での自由度はいくらでしょう？

$\phi = (a-1)(b-1) = (2-1) \times (2-1) = 1$

(p134.0.7)分割表の基本：手順：(4) χ^2検定

実度数

	良品	不良品	計
製法1	12	8	20
製法2	8	12	20
計	20	20	40

期待度数

	良品	不良品	計
製法1	10	10	20
製法2	10	10	20
計	20	20	40

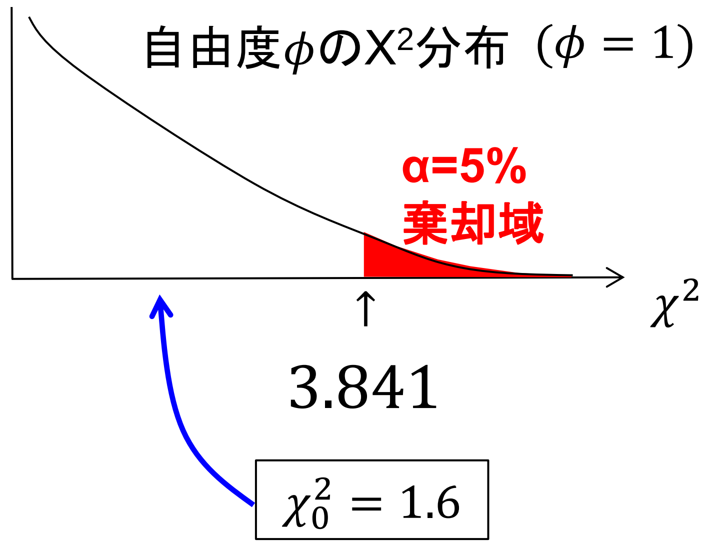
χ²分布の上側P点

	P=	0.975	0.950	0.050	0.025
ϕ=	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

ピアソンのχ^2統計量 (検定統計量):

$$\chi_0^2 = \frac{(12-10)^2}{10} + \frac{(8-\ 10)^2}{10} + \frac{(8-\ 10)^2}{10} + \frac{(12-\ 10)^2}{10} = 1.6$$

自由度: $\phi = (a-1)(b-1) = (2-1)\times(2-1) = 1$



検定統計量$\chi_0^2 = 1.6$は棄却域に入っていない

⇒不良品の発生率が製法により異なるとは言えない。

(p134.0.8)分割表の基本：手順：まとめ

①:　2つの仮説を立てます

帰無仮説(H_0):製法1,2により良品、不良品の出方に違いはない

対立仮説(H_1):製法1,2により良品、不良品の出方に違いがある

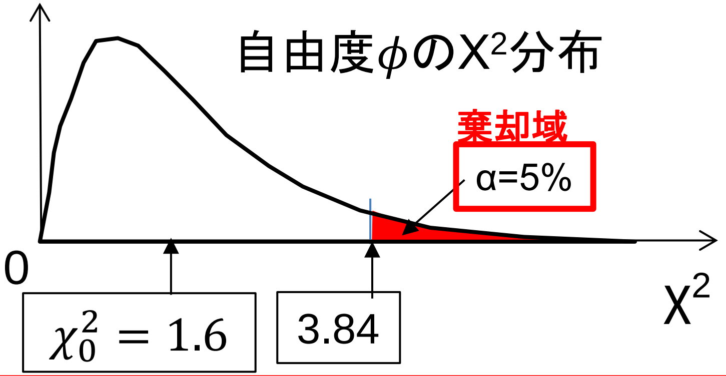
②:　有意水準αを定めます

α=0.05=5%

③:棄却域を定めます

棄却域: $\chi^2 > \chi^2(\phi, \alpha) = 3.84$

但し、自由度: $\phi = 1$



(補足)自由度の求め方

	良品	不良品
装置1	12	8
装置2	8	12

a個 $(a = 2)$

b個 $(b = 2)$

自由度:

$\phi = (a-1)(b-1) = 1$

④:

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

を計算します

$\Rightarrow \chi_0^2 = 1.6$

⑤: H_0を棄却する・しないの判定

$\boldsymbol{\chi^2 = \chi_0^2}$が棄却域にあれば、 H_0を棄却する

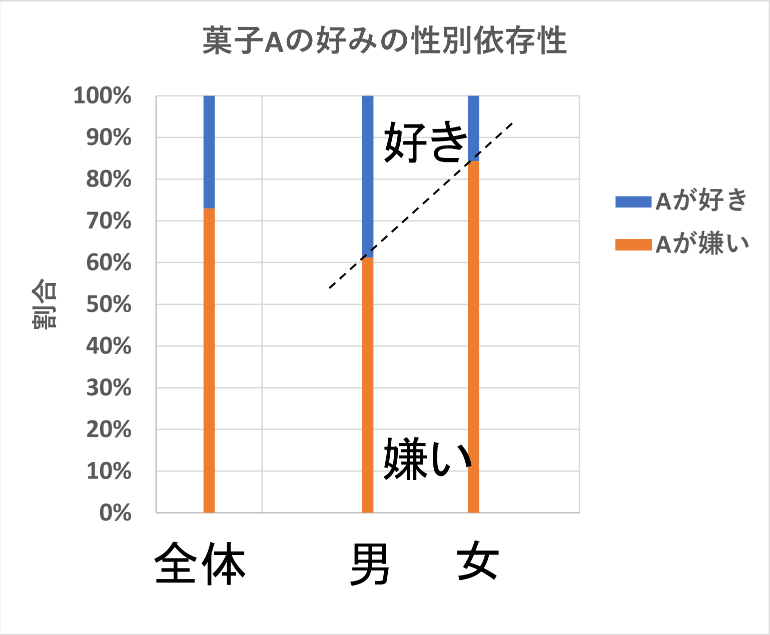
$\Rightarrow$ H_1は正しいと言える

$\boldsymbol{\chi^2 = \chi_0^2}$が棄却域になければ、$H_0$は棄却できない

$\Rightarrow$ H_1は正しいかどうか何も言えない

(p134.1)[C9-2]問1. 期待度数・自由度

(ABランク)



実データ（観測度数）

	Aが好き	Aが嫌い	合計
男	19	30	49
女	8	43	51
合計	27	73	100

期待度数

	Aが好き	Aが嫌い	合計
男	★		49
女			51
合計	27	73	100

公式

m (列数), k (行数)

	カテゴリー1	・・・	カテゴリーm	計
グループ1	(d)			(c)
・・・				・・・
グループk				
計	(a)	・・・		(b)

期待度数の求め方 : $(d) = (c) \times \frac{(a)}{(b)}$

自由度 : $\phi = (k-1) \times (m-1)$

★ 期待度数(男、Aが好き) = $49 \times \frac{27}{100} = 13.23$ (ア)

自由度= (2-1)×(2-1)=1 (イ)

(答)①

済

(p136.1)[C9-2]問2. 期待度数・独立性検定 (ABランク)

$b=3$

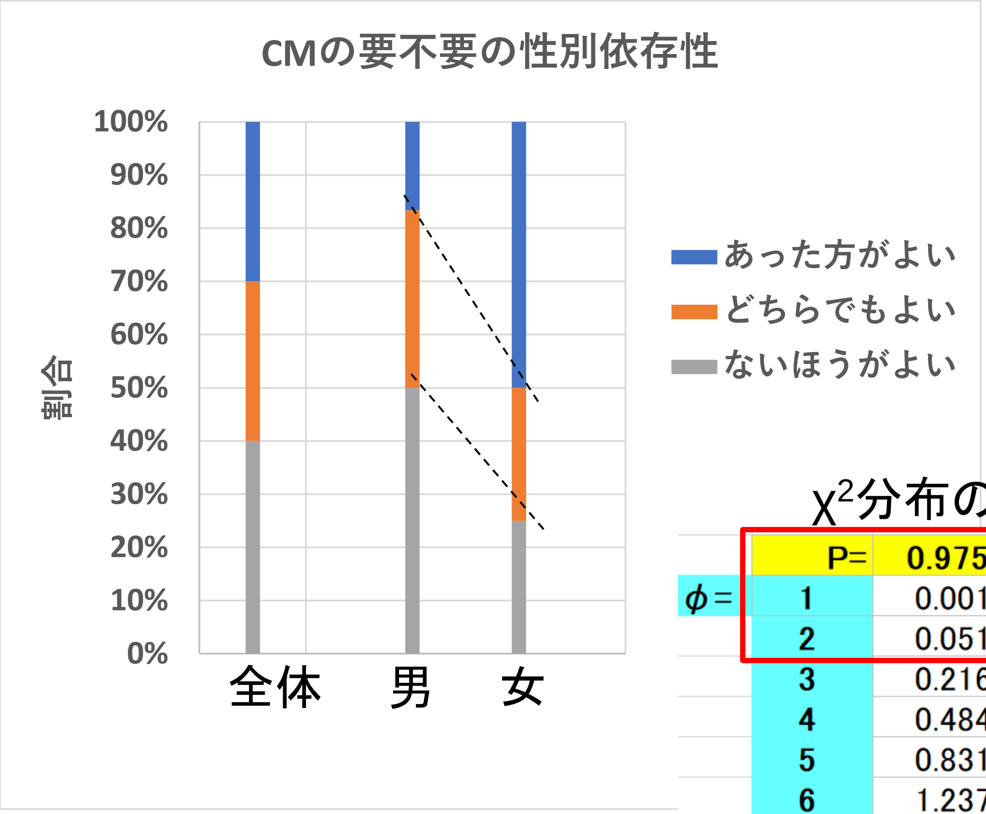
$a=2$

	あった方がよい	どちらでもよい	ないほうがよい	合計
男	5	10	15	30
女	10	5	5	20
合計	15	15	20	50

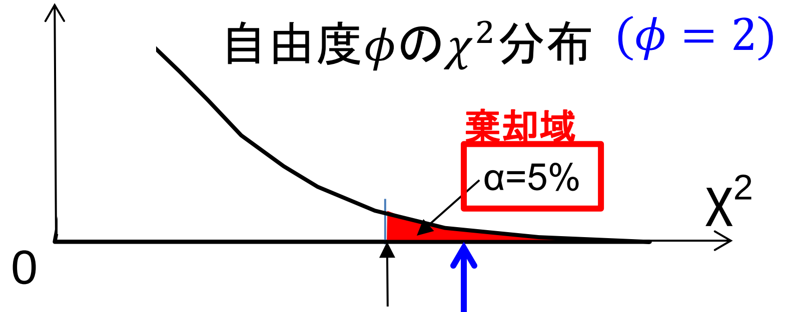
統計理論によると、ピアソンのχ^2統計量：

$$\chi_0^2=\sum\frac{(\text{実度数}-\text{期待度数})^2}{\text{期待度数}}$$

は自由度$\phi=(a-1)(b-1)$の
χ^2分布に近似的に従うことが知られています。



(1)χ^2分布の自由度ϕ = (2-1)×(3-1)=2



(3) 5.991

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

(2) χ^2値 = 6.60

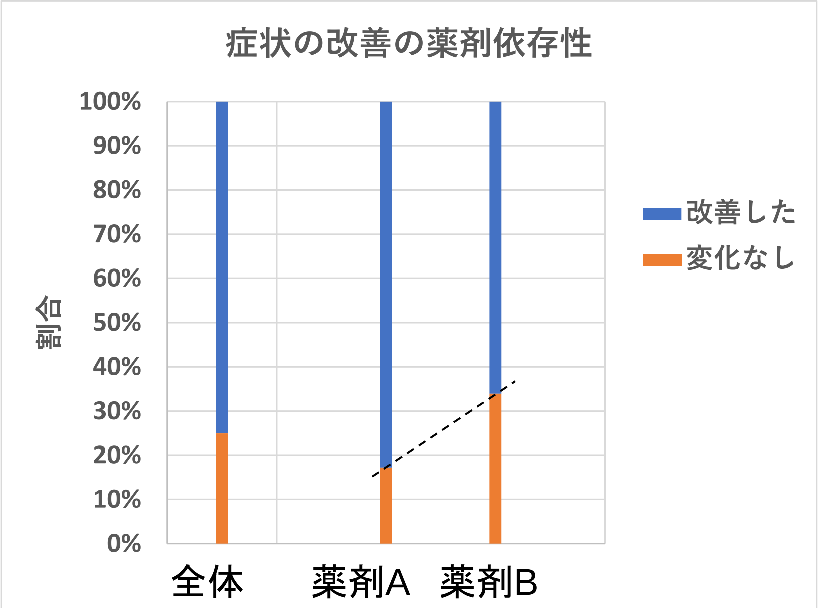
(4)棄却域にある

⇒帰無仮説を棄却する

済

(答)②

(p138.1)[C9-2]問3. 期待度数・独立性検定 (ABランク)



流れ@χ^2検定

- 検定統計量(χ^2統計量)
  - 観測度数(実度数)
  - 期待度数
- 棄却域←自由度、α=5%

⇒検定

観測度数

$m=2$, $k=2$

	改善した	変化なし	合計
薬剤A	53	11	64
薬剤B	37	19	56
合計	90	30	120

期待度数

	改善した	変化なし	合計
薬剤A	★48		64
薬剤B			56
合計	90	30	120

公式

m, k

	カテゴリー1	・・・	カテゴリーm	計
グループ1	(d)			(c)
・・・				・・・
グループk				
計	(a)	・・・		(b)

期待度数の求め方:$(d)=(c)\times\dfrac{(a)}{(b)}$

自由度:$\phi=(k-1)\times(m-1)$

Q1: 自由度$\phi=$ $(k-1)(m-1)=1$

Q2: 期待度数(★) = $64\times\dfrac{90}{120}=48$

(答)(ア):A2 ⇒③④⑤

Q3: 他の期待度数は?

	改善した	変化なし	合計
薬剤A	48	64-48= 16	64
薬剤B	90-48= 42	30-16= 14	56
合計	90	30	120

(p138.2)[C9-2]問3. 期待度数・独立性検定 (ABランク)

実度数

	改善した	変化なし	合計
薬剤A	53	11	64
薬剤B	37	19	56
合計	90	30	120

期待度数

	改善した	変化なし	合計
薬剤A	48	16	64
薬剤B	42	14	56
合計	90	30	120

$$\frac{(\text{実度数}-\text{期待度数})^2}{\text{期待度数}}$$

	改善した	変化なし
薬剤A	25/48	25/16
薬剤B	25/42	25/14

検定統計量(χ^2統計量):

$$\text{Q1: } \chi_0^2 = \sum \frac{(\text{実度数}-\text{期待度数})^2}{\text{期待度数}}$$

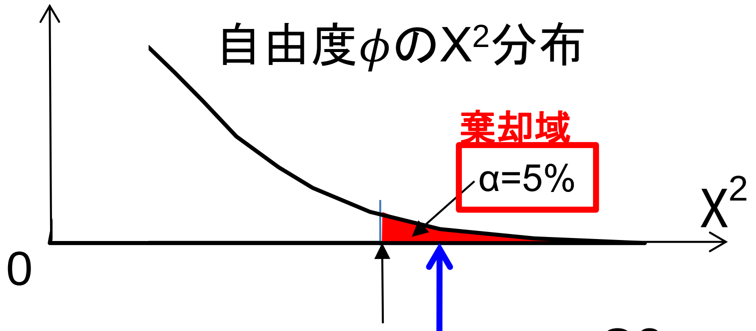
$$= 25 \times \left(\frac{1}{48}+\frac{1}{16}+\frac{1}{42}+\frac{1}{14}\right) = 4.46$$

統計理論によると、ピアソンのχ^2統計量:

$$\chi_0^2 = \sum \frac{(\text{実度数}-\text{期待度数})^2}{\text{期待度数}}$$

は自由度$\phi=(a-1)(b-1)$の
χ^2分布に近似的に従うことが知られています。

自由度$\phi = 1$



Q2: 3.841

Q1: χ_0^2 = 4.46

Q3:
χ^2値は
棄却域
にある

χ²分布の上側P点

	P=	0.975	0.950	0.050	0.025
φ=	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

χ^2値は棄却域にある
帰無仮説は棄却される
薬により効果に違いがあると言える

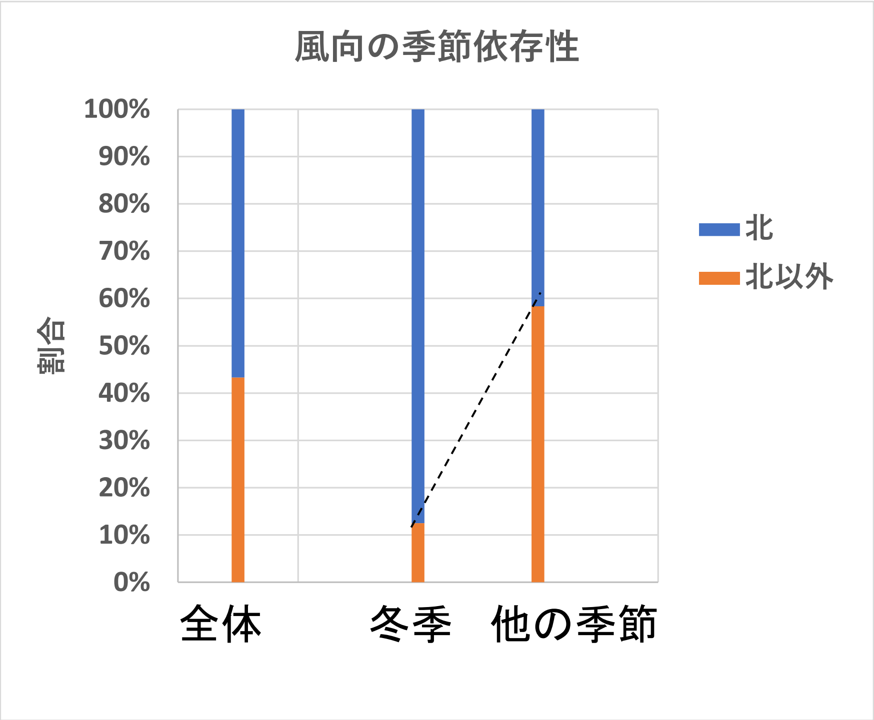
(答)(イ):B1 ⇒③

(答)(ア):A2 ⇒③④⑤

(答)③

済

(p140.1)[C9-2]問4. 検定結果に関する正誤 (ABランク)



「季節」「風向」間に関連があるか？

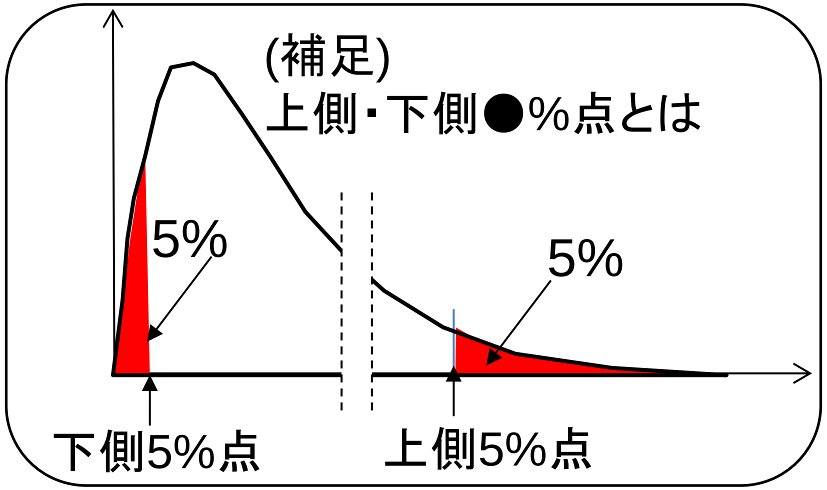
検定統計量 :$\chi_0^2 = 69.04$

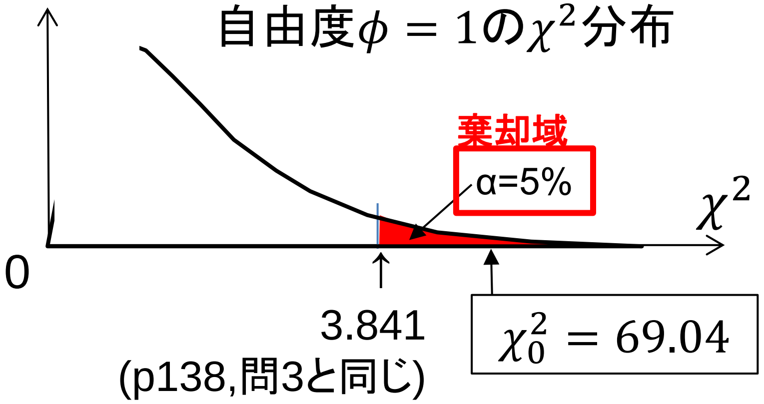
自由度$\phi =$ 1

- 上側5%点
- 下側5%点
- 両側5%点

今回、どれを使う？

(答)の候補 ④、⑤





④「季節」「風向」間に関連はない ⇔帰無仮説(H_0)

⑤「季節」「風向」間に関連がある ⇔対立仮説(H_1)

χ^2値は(H_0の)棄却域にある

⇒帰無仮説(H_0)は棄却される

⇒対立仮説(H_1)は正しいと言える

(答)⑤

済