

## 経理項目に基づく企業の構造化分析-構造の境界値検定-

統計センター 伊藤 孝之 統計センター 白川 清美

公的統計調査において、欠測値の存在が話題となっている。特に、経理項目における欠測値は、正確な統計を集計する上での影響が大きい。そのため、我々は「補定等のための最適な層の分け方」に着目した欠測値の補定方法を研究してきた。ここでは、産業分類、都道府県、従業者の規模などを用いて層化をした。しかしながら、企業の構造分析については着目していなかった。そこで本研究では、企業の構造の在り方に着目した層化を提案する。企業を取り巻く環境の変化に基づく企業構造の変化が「売上高」などの経理項目に及ぼす影響は大きい。このことにより、企業の戦略、市場の動向などのさまざまな要因から企業構造の分類ができる。具体的には、中小企業と大企業の境界付近のデータの分布の特性の分析を行う。これは、企業構造分析の一つの取組であり、今後の課題である「詳細な層の選定」に資することを目的とする。

中小企業であるか大企業であるかは、従業者数と資本金額の値によって決まり、従業者数でも資本金額でもどちらか一方が境界値以下であれば中小企業となる。つまり、中小企業には、中小企業 A (資本金のみ境界値越え)、中小企業 B (従業者数、資本金額とも境界値以下)、中小企業 C (従業者数のみ境界値越え) の 3 パターンが考えられる。3 パターンのそれぞれの中小企業と大企業の売上高に違いがあるか、試行したものである。

売上高に違いがあるかどうかの判定は、Chow の検定とした。これは、中小企業と大企業に分解して回帰分析を行った場合と、それらをまとめて回帰分析をした場合において、残差平方和の乖離に有意な差があるかを検証したものである。

なお、境界値検定に用いたデータは、平成 24 年経済センサス - 活動調査速報集計時の個票データである。それゆえ、産業分類は、産業大分類が検証の対象となる。

テストの結果では、概ね構造に違いがあると考えられる結果となった。この観点から、中小企業、大企業別で層を分けることを考慮に入れることができる。検証結果の詳細は、当日報告する。また、今後の予定では、経済センサス確報集計の個票データを用いることにより、産業細分類や他の要因を加えた企業構造の層化の分析を行うこととする。

### 参考文献

1. Ito, Takayuki, Yutaka Abe, and Tatsuo Noro. (2013). "The Best Stratification to Impute Missing Values of Turnover in Economic Surveys," *The 59th World Statistics Congress, Hong Kong, China, 25-30 August 2013*.
2. 白川清美. (2013). 「定性データに基づく企業構造の境界値分析」, 2013年度統計関連学会連合大会講演報告集, 大阪大学豊中キャンパス.
3. 中小企業庁ホームページ. (2013). 『中小企業・小規模企業者の定義』 . <http://www.chusho.meti.go.jp/faq/faq01.html>. (2013年6月27日アクセス).
4. 松田芳郎, 清水雅彦, 舟岡史雄. (2003). 『講座ミクロ統計分析第4巻 企業行動の変容 - ミクロデータによる接近』, 東京, 日本評論社.
5. 蓑谷千風彦, 縄田和満, 和合肇. (2007). 『計量経済学ハンドブック』, 東京, 朝倉書店.

# 経理項目に基づく企業の構造化分析 -構造の境界値検定-

2013年度統計関連学会連合大会  
大会日程：2013年9月9日（月）  
場 所：大阪大学 豊中キャンパス

独立行政法人統計センター ○伊藤孝之  
白川清美



独立行政法人

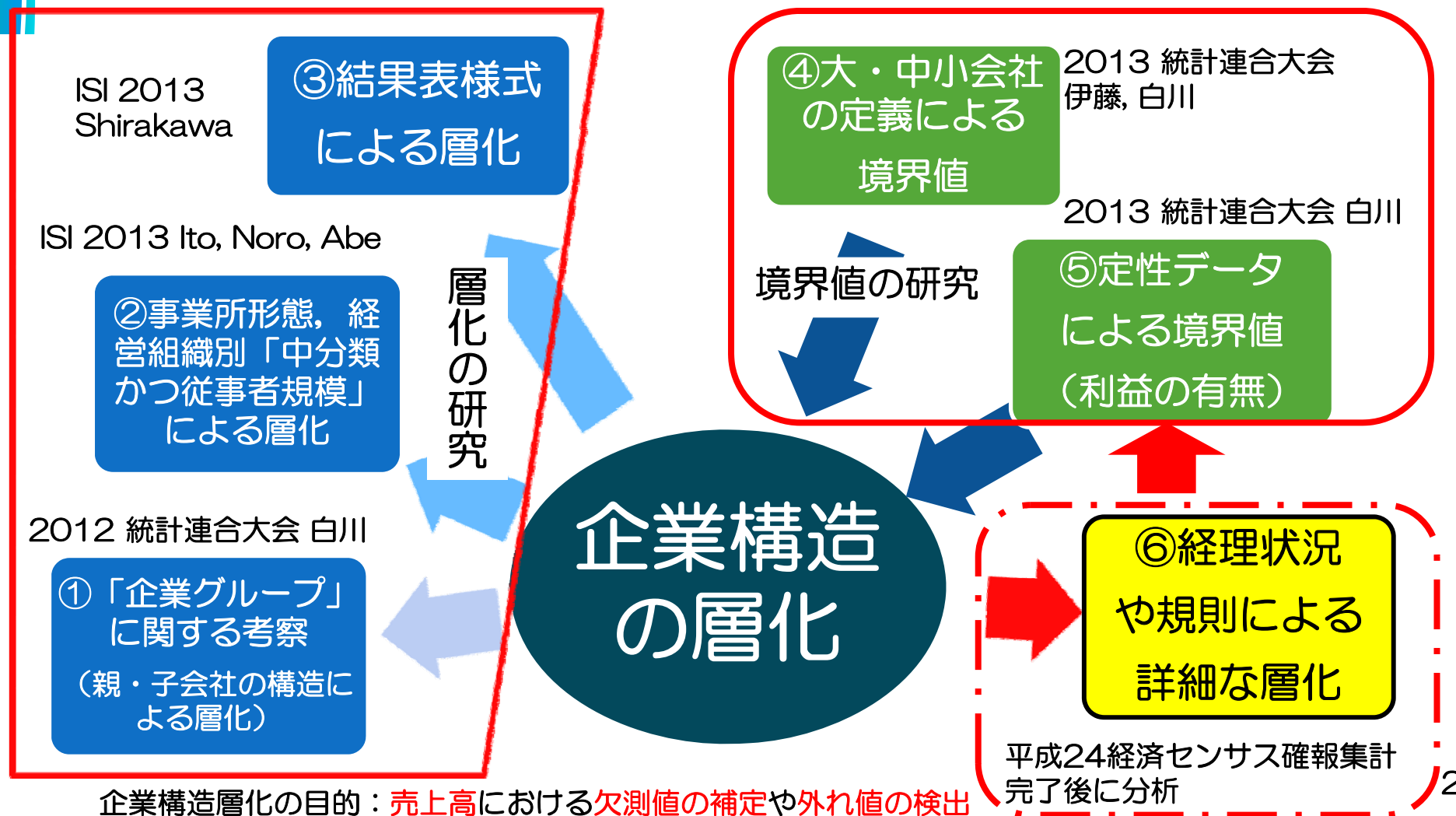
統計センター

この発表は個人の見解であり，所属する組織の公式見解ではありません。

# 企業構造層化に関する研究の関連図

## -層化の効果と境界値の設定-

### 企業構造層化に関する研究関連図



企業構造層化の目的：売上高における欠測値の補定や外れ値の検出

# 【背景 1】

---

## 【経理項目における従来の層化の種類】

- 経理項目の分布の特性を検証することは、正確な統計を集計するうえで重要である。
- そこで、分布の特性について研究を行ってきた。
- 研究過程で、売上高の分布がいくつかの層に分けられることを観測した。
- 過去、事業所形態、経営組織別、産業別、都道府県別、従業者情報の規模別で層を区切って欠測値補定の研究をすすめてきた。

(参考：ISI 2013 Ito, Abe, Noro)

# 【背景2】

## 【中小企業と大企業の境界値付近の検証】

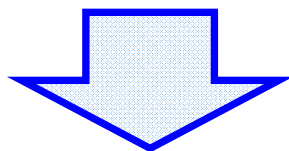
- 企業の構造分析については、あまり着目されていなかった。
- 本研究では、中小企業と大企業で経理項目（本件は売上高）に違いが存在するか分析する。
- しかしながら、公的統計調査において詳細な経理状況を把握することには困難が多い。
- さらに、企業構造を推測できる調査項目には限りがある。そこで、本研究では、売上高を説明する変数として、公的統計調査全般で調査対象項目となっている、**従業者情報と資本金**を用いて検証する。

# 目的

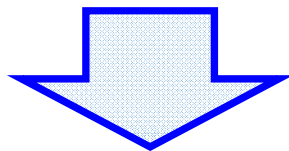
---

## 【経済センサス - 活動調査への期待】

- 中小企業、大企業で売上高に違いが存在する



- 定性データによる境界値（利益の有無）分布の検証



- さらに詳細な層化を実現する

# 1. 経理項目における従来の層化の種類

---

## 経理項目の層化の種類（従来）

- 事業所形態（単独事業所、複数事業所企業）
- 経営組織（個人経営、会社、会社以外の法人 など）
- 産業（大分類、中分類 など）
- 都道府県別（47県、または大都市圏）
- 従業者の規模（1-4人、5-9人 等）
  
- 中小企業、大企業の別

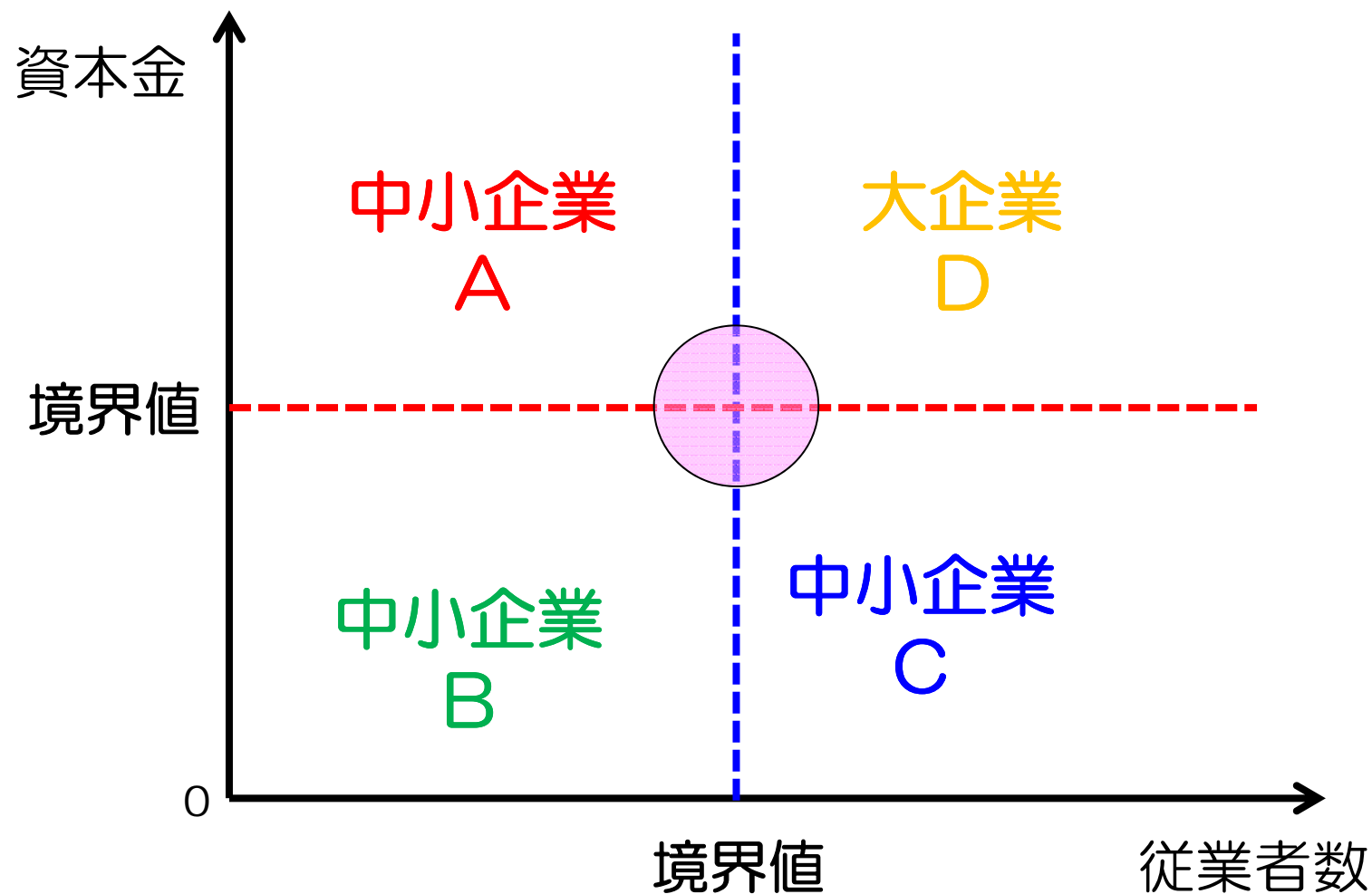
## 2. 中小企業の定義

業種分類	中小企業基本法の定義
製造業その他	資本金の額又は出資の総額が3億円以下の会社又は常時使用する従業員の数が300人以下の会社及び個人
卸売業	資本金の額又は出資の総額が1億円以下の会社又は常時使用する従業員の数が100人以下の会社及び個人
小売業	資本金の額又は出資の総額が5千万円以下の会社又は常時使用する従業員の数が50人以下の会社及び個人
サービス業	資本金の額又は出資の総額が5千万円以下の会社又は常時使用する従業員の数が100人以下の会社及び個人

参考：中小企業庁：「中小企業・小規模企業者の定義」より



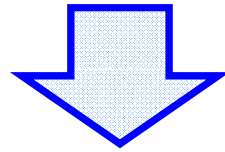
### 3. 境界値付近のデータの特定



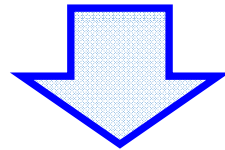
### 3. 境界値付近のデータの特定（続き）

---

① 従業者数と資本金をZスコア化



② 中心（従業者と資本金の境界値）からの距離（ユークリッド距離）を測定



③ 距離が境界値から近いデータを優先して選定

## 4. 検証の方法と判定方法

---

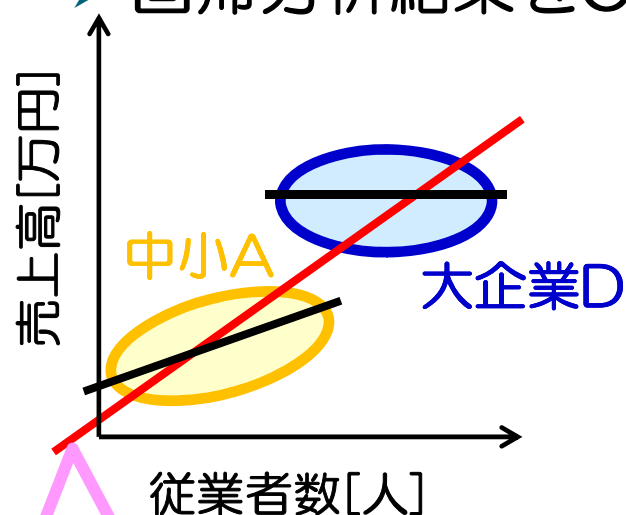
### 【検証の方法】

- 回帰分析（目的変数：売上高、説明変数：従業者、又は資本金）
- 比較対象データは、
  - ・ 中小企業A（資本金のみ境界値超え）と大企業D
  - ・ 中小企業B（従業者、資本とも境界値未満）と大企業D
  - ・ 中小企業C（従業者のみ境界値超え）と大企業D
  - ・ 中小企業Aと中小企業B（参考）
  - ・ 中小企業Bと中小企業C（参考）

## 4. 検証の方法と判定方法（続き）

### 【判定方法】

➤ 回帰分析結果をChowの検定（F検定）により判定



$$F = \frac{(a - b) \times f_2}{b \times f_1}$$

中小Aと大企業D  
で層を分けずに回  
帰分析した場合の  
回帰直線

a : 中小Aと大企業Dを混合した状態での残差平方和

b : 中小Aの残差平方和 + 大企業Dの残差平方和

$f_1$  : 目的変数と説明変数の個数

$f_2$  : 中小Aのデータ数 + 大企業Dのデータ数 -  $2 \times f_1$

# 5. 検証に用いたデータと産業情報

## 検証に用いたデータ

- ▶ 平成24年経済センサス - 活動調査（速報版）データ
  - ▶ うち、複数事業所企業で個人経営以外
  - ▶ かつ、東京都のデータ
  - ▶ データを対数変換する

## ▶ 産業情報

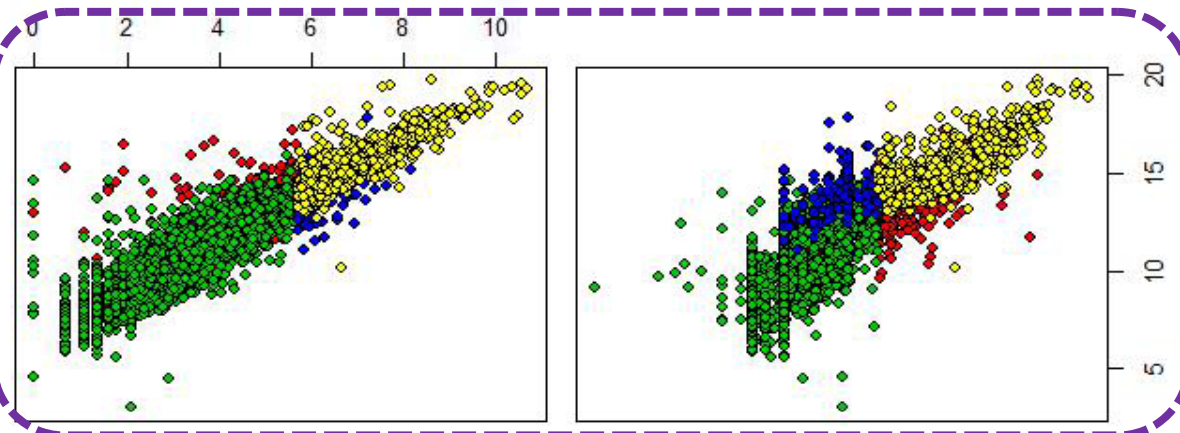
産業大分類	全データ数
E（製造業）	5253
G（情報通信業）	1315
L（学術研究, 専門・技術サービス業）	1035
N（生活関連サービス業, 娯楽業）	1177
R（サービス業（他に分類されないもの））	1378

# 6. 散布図行列 ～産業E（製造業）対数変換～

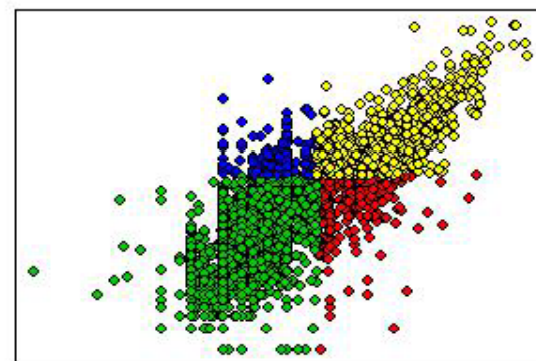
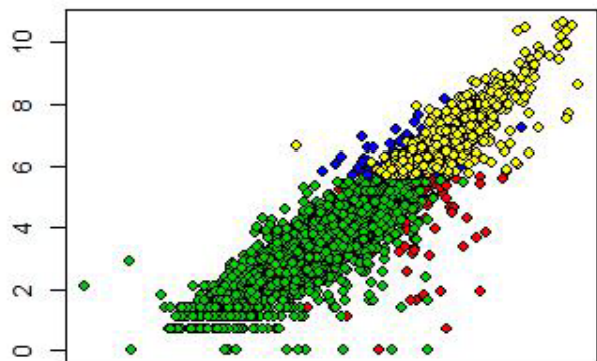
データ数：5253

- は中小A
- は中小B
- は中小C
- は大企業D

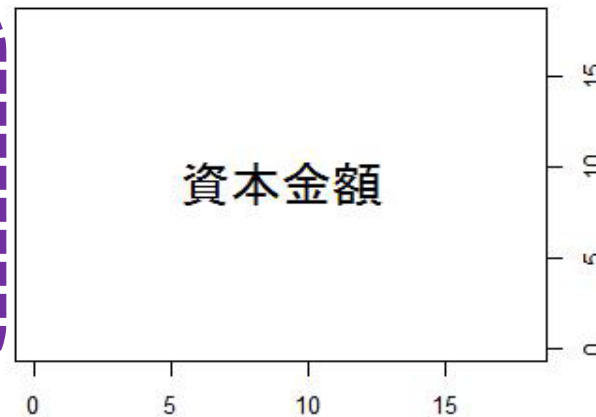
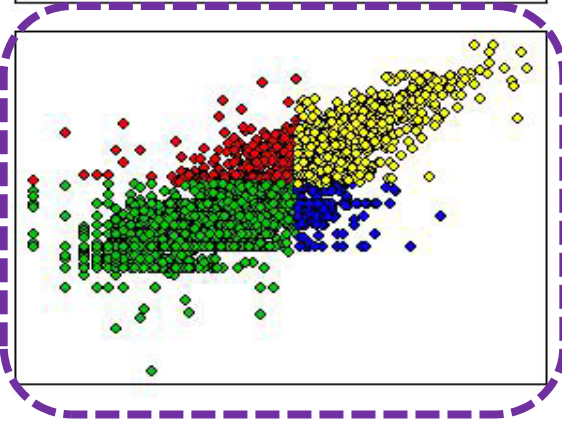
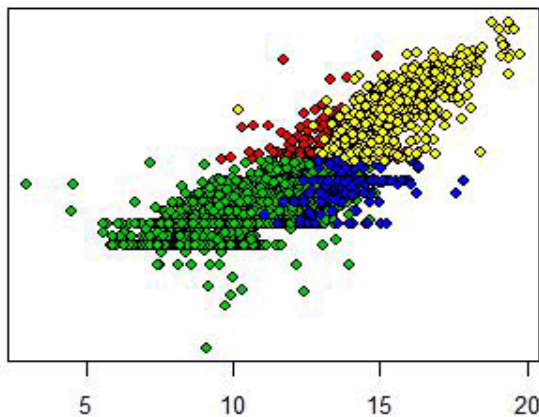
売上高



従業者数



資本金額



# 7-1. Chowの検定結果 産業大分類：E 対数変換 ～中小A（資本金のみ境界値超え）と大企業D～

○境界値付近のデータ 目的変数：売上高

説明変数	中小、大企業 業の別	切片		従業者		資本金	
		係数	標準誤差	係数	標準誤差	係数	標準誤差
従業者	中小A	6.8	1.9	1.3	0.4	-	-
	大企業D	9.0	2.5	0.9	0.4	-	-
資本金	中小A	7.6	1.5	-	-	0.6	0.1
	大企業D	14.3	1.2	-	-	0.0	0.1

±2標準誤差での範囲

説明変数	中小、大企業 業の別	切片		従業者		資本金	
		下限値	上限値	下限値	上限値	下限値	上限値
従業者	中小A	2.88	10.63	0.58	2.01	-	-
	大企業D	3.97	14.07	0.02	1.72	-	-
資本金	中小A	4.65	10.54	-	-	0.29	0.83
	大企業D	11.81	16.71	-	-	-0.22	0.21

【説明変数が従業者】⇒切片、傾きとも分布の傾向が同じ

【説明変数が資本金】⇒切片、傾きとも分布の傾向が違う

# 7-1. Chowの検定結果（続き）

～中小A（資本金のみ境界値超え）と大企業D～

境界値付近のデータ

産業大分類：E 対数変換

説明変数	中小、大企業の別	自由度調整済R <sup>2</sup>	Chowの検定			データ数	
			F値	上側5%	検定結果	中小A	大企業
従業者	中小A	0.11	0.62	3.04	容認	100	-
	大企業D	0.03				-	100
資本金	中小A	0.14	13.35	3.04	棄却	100	-
	大企業D	-0.01				-	100

有意水準：5%  
 容認：売上高に違いがない  
 棄却：売上高に違いがある

【説明変数が従業者】⇒切片、傾きとも分布の傾向が同じ

【説明変数が資本金】⇒切片、傾きとも分布の傾向が違う



# 7-1. Chowの検定結果（続き）

～中小A（資本金のみ境界値超え）と大企業D～

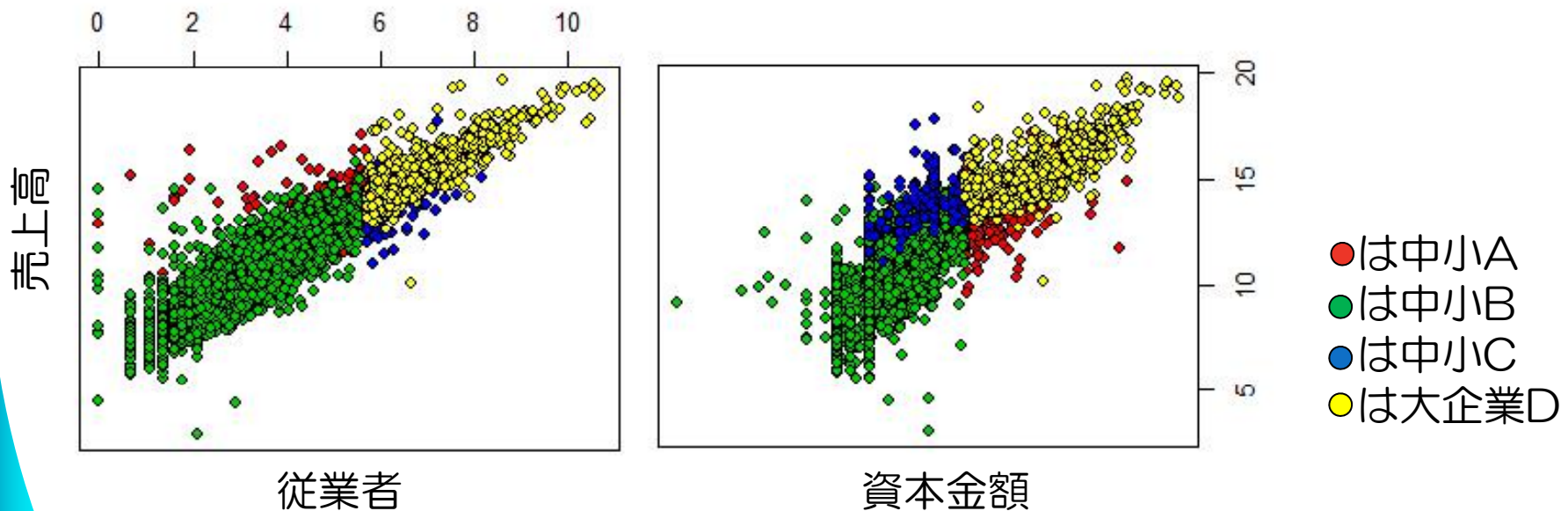
全データ

産業大分類：E 対数変換

説明変数	中小、大企業の別	自由度調整済R <sup>2</sup>	Chowの検定			データ数	
			検証F値	上側5%	検定結果	中小A	大企業
従業者	中小A	0.08	60.56	3.01	棄却	296	-
	大企業D	0.66				-	584
資本金	中小A	0.11	95.31	3.01	棄却	296	-
	大企業D	0.51				-	584

【説明変数が従業者】⇒切片、傾きとも分布の傾向が違う

【説明変数が資本金】⇒切片は分布の傾向が同じであるが、傾きは違う



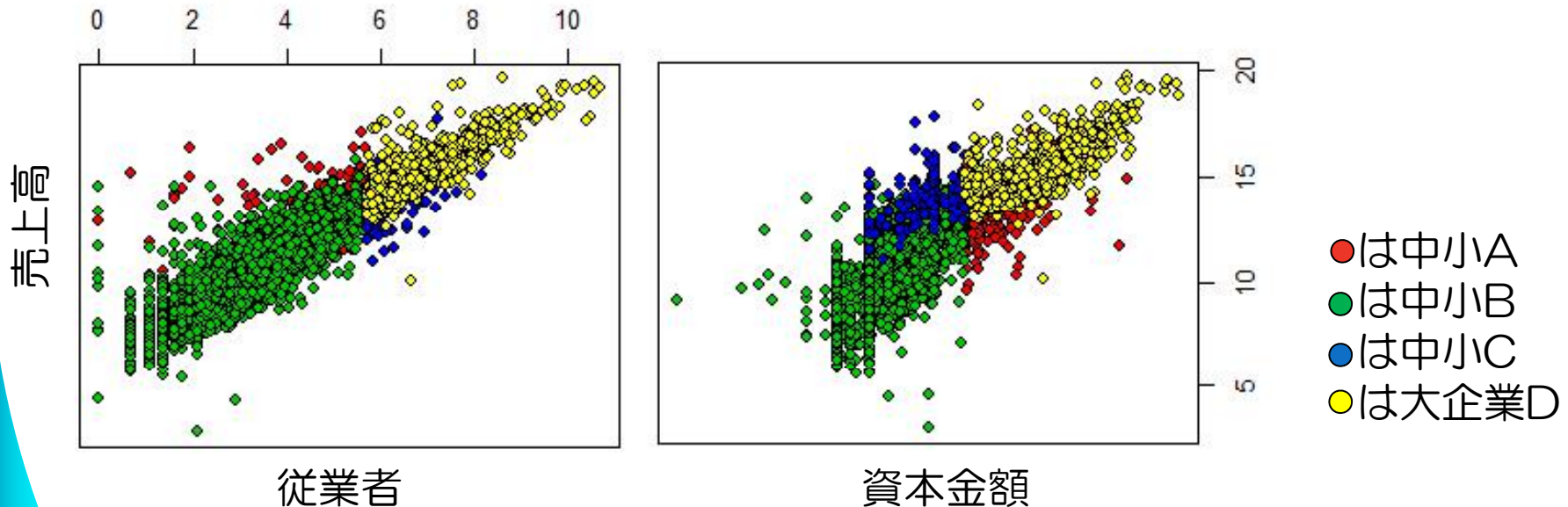
## 7-2. Chowの検定結果 産業大分類：E 対数変換 ~中小B(従業者、資本とも境界値未満)と大企業D~

境界値付近の別	説明変数	中小、大企業の別	自由度調整済R <sup>2</sup>	Chowの検定			データ数	
				F値	上側5%	検定結果	中小B	大企業
境界値付近データ	従業者	中小B	0.050	4.18	3.04	棄却	100	-
		大企業D	0.032				-	100
	資本金	中小B	0.059	7.28	3.04	棄却	100	-
		大企業D	-0.010				-	100
全データ	従業者	中小B	0.729	43.13	3.00	棄却	4202	-
		大企業D	0.658				-	584
	資本金	中小B	0.454	62.07	3.00	棄却	4202	-
		大企業D	0.513				-	584

全データに対して、

【説明変数が従業者】⇒切片、傾きとも分布の傾向が違う

【説明変数が資本金】⇒切片、傾きとも分布の傾向が違う



# 7-3. Chowの検定結果

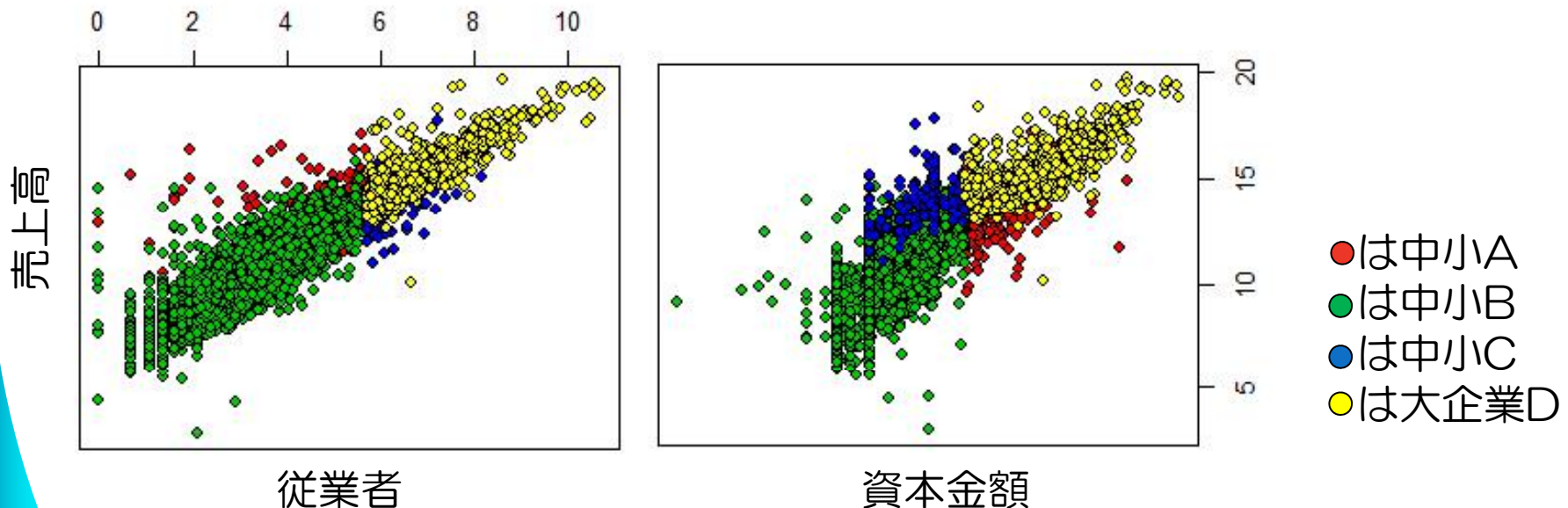
～中小C(従業者のみ境界値超え)と大企業D～

境界値付近の別	説明変数	中小、大企業の別	自由度調整済R <sup>2</sup>	Chowの検定			データ数	
				F値	上側5%	検定結果	中小C	大企業
境界値付近データ	従業者	中小C	0.004	11.92	3.04	棄却	100	-
		大企業D	0.032				-	100
	資本金	中小C	0.181	4.55	3.04	棄却	100	-
		大企業D	-0.010				-	100
全データ	従業者	中小C	0.299	50.28	3.01	棄却	158	-
		大企業D	0.658				-	584
	資本金	中小C	0.090	26.64	3.01	棄却	158	-
		大企業D	0.513				-	584

全データに対して、

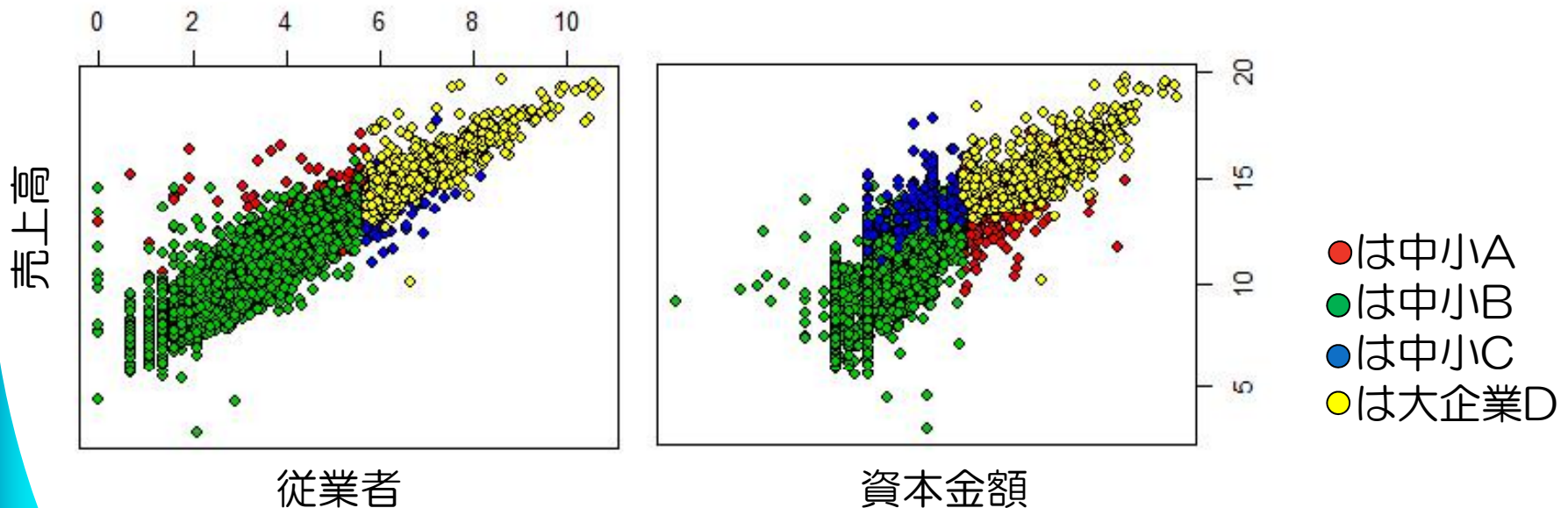
【説明変数が従業者】⇒切片、傾きとも分布の傾向が違う

【説明変数が資本金】⇒切片、傾きとも分布の傾向が違う



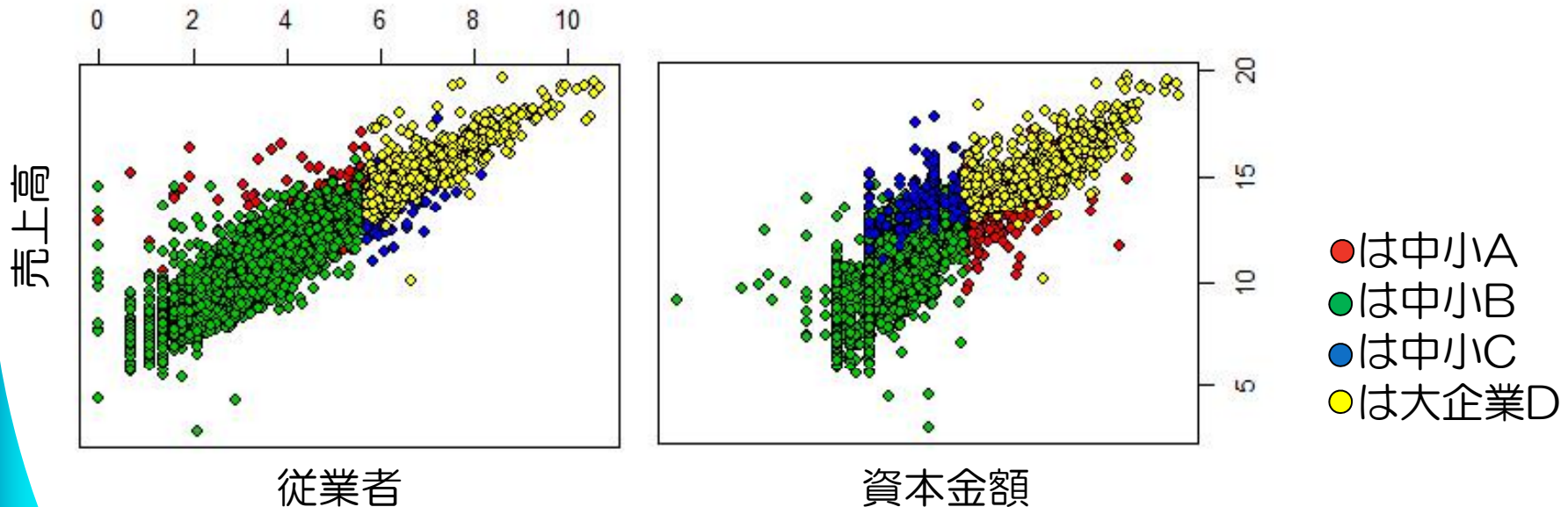
# 7-4. Chowの検定結果 産業大分類：E 対数変換 ~中小Aと中小B~ (参考)

境界値付近の別	説明変数	中小、大企業の別	自由度調整済R <sup>2</sup>	Chowの検定			データ数	
				F値	上側5%	検定結果	中小A	中小B
境界値付近データ	従業者	中小A	0.108	23.30	3.04	棄却	100	-
		中小B	0.050				-	100
	資本金	中小A	0.142	2.46	3.04	容認	100	-
		中小B	0.059				-	100
全データ	従業者	中小A	0.076	360.16	3.00	棄却	296	-
		中小B	0.729				-	4202
	資本金	中小A	0.114	71.22	3.00	棄却	296	-
		中小B	0.454				-	4202



# 7-5. Chowの検定結果 産業大分類：E 対数変換 ～中小Bと中小C～（参考）

境界値付近の別	分析の別	中小、大企業の別	自由度調整済R <sup>2</sup>	Chowの検定			データ数	
				F値	上側5%	検定結果	中小B	大企業
境界値付近データ	従業者	中小B	0.050	0.643	3.042	容認	100	-
		中小C	0.004				-	100
	資本金	中小B	0.059	9.263	3.042	棄却	100	-
		中小C	0.181				-	100
全データ	従業者	中小B	0.729	5.916	2.998	棄却	4202	-
		中小C	0.299				-	158
	資本金	中小B	0.454	232	2.998	棄却	4202	-
		中小C	0.090				-	158



# 7-6. Chowの検定結果 ～中小企業と大企業～

産業大分類	比較対象	説明変数			
		従業者		資本金	
		境界値付近	全データ	境界値付近	全データ
E	AとD	容認	棄却	棄却	棄却
	BとD	棄却	棄却	棄却	棄却
	CとD	棄却	棄却	棄却	棄却
G	AとD	容認	棄却	棄却	棄却
	BとD	棄却	棄却	棄却	棄却
	CとD	棄却	棄却	容認	棄却
L	AとD	容認	棄却	棄却	棄却
	BとD	容認	棄却	棄却	棄却
	CとD	棄却	棄却	容認	容認
N	AとD	容認	容認	棄却	棄却
	BとD	棄却	棄却	棄却	棄却
	CとD	棄却	棄却	容認	容認
R	AとD	容認	容認	棄却	棄却
	BとD	棄却	棄却	棄却	棄却
	CとD	棄却	棄却	容認	棄却

## 5産業分の結果

比較対象	容認数	棄却数	棄却の割合
AとD	7	13	65%
BとD	1	19	95%
CとD	6	14	70%

境界値の別	容認数	棄却数	棄却の割合
境界値付近	10	20	67%
全データ	4	26	87%

	容認数	棄却数	棄却の割合
全体	14	46	77%

# 7-7. Chowの検定結果 ～中小企業～（参考）

産業大分類	比較対象	説明変数			
		従業者		資本金	
		境界値付近	全データ	境界値付近	全データ
E	AとB	棄却	棄却	容認	棄却
	BとC	容認	棄却	棄却	棄却
G	AとB	棄却	棄却	棄却	棄却
	BとC	容認	棄却	棄却	棄却
L	AとB	棄却	棄却	容認	棄却
	BとC	容認	容認	棄却	棄却
N	AとB	棄却	棄却	容認	棄却
	BとC	容認	容認	棄却	棄却
R	AとB	棄却	棄却	容認	棄却
	BとC	容認	棄却	容認	棄却

## 5産業分の結果

比較対象	容認数	棄却数	棄却の割合
AとB	4	16	80%
BとC	8	12	60%

境界値の別	容認数	棄却数	棄却の割合
境界値付近	10	10	50%
全データ	2	18	90%

	容認数	棄却数	棄却の割合
全体	12	28	70%

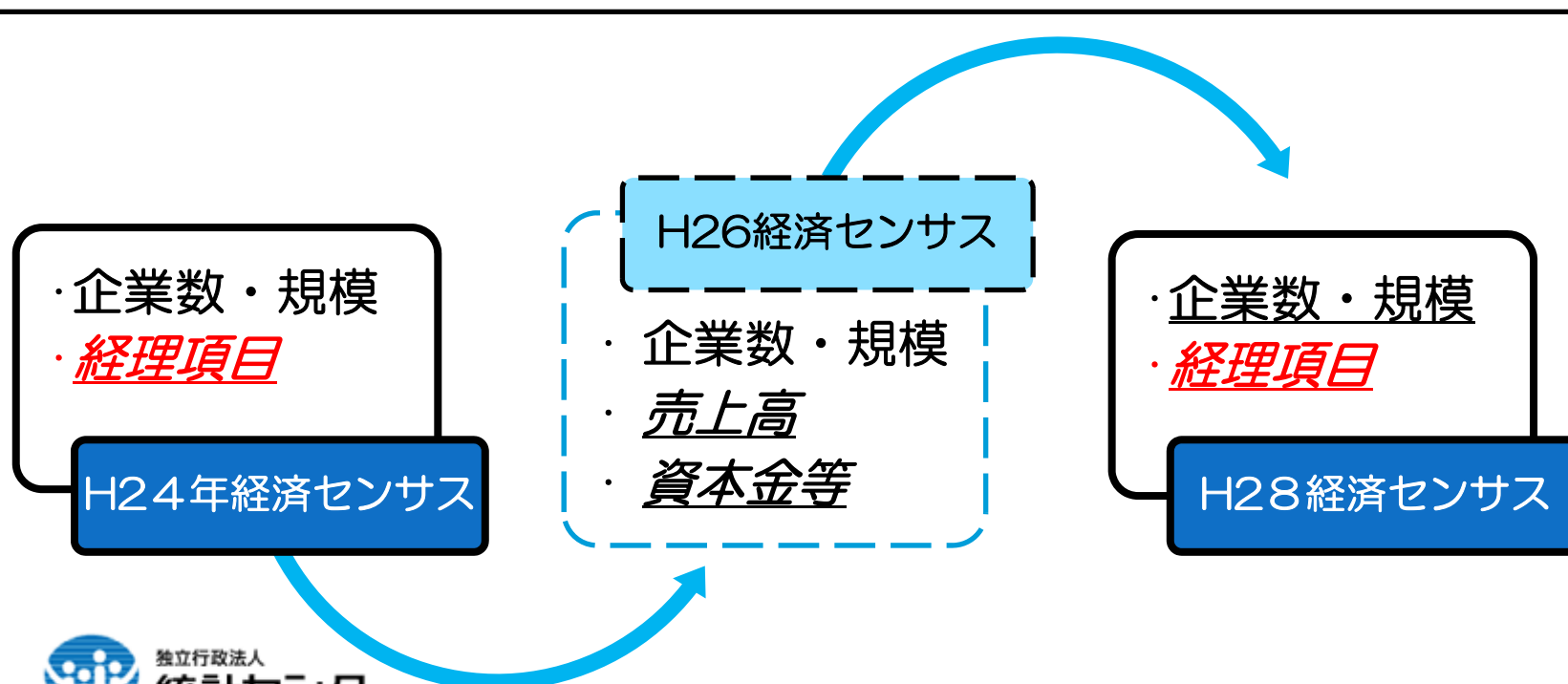
## 8. 経済センサスへの期待

### 【経済センサス - 基礎調査】

- 親子関係の調査項目も含まれている。

### 【経済センサス - 活動調査】

- 経理項目が含まれている。





# 検証結果と課題

---

- 中小企業と大企業ではChowの検定により売上高に違いが存在することが分かった。
- また、中小企業同士でも違いが発生することが分かった。
- 反面、中小企業と大企業以外の要因が売上高に与える影響を考慮する必要がある。
- そのため、今後は更に細分化した層に分けて検定等を行う必要がある。

# 今後の期待

---

1. 検証に用いるデータ数の特定
  - 単回帰分析から重回帰分析への展開
2. 他の企業の規模での検証
  - 親会社と子会社での比較
  - 連結企業の有無との比較 など
3. 評価方法の拡大
  - 判別分析など
4. 他の経理項目（給与、売上原価等）の分布の解析
  - 層化により売上高を説明する潜在的な経理項目を特定
5. 定性データで解析した境界値との融合

# 参考文献

---

- Ito, Takayuki, Yutaka Abe, and Tatsuo Noro. (2013). “ The Best Stratification to Impute Missing Values of Turnover in Economic Surveys,” *The 59th World Statistics Congress*, Hong Kong, China, 25-30 August 2013.
- 白川清美. (2013). 「定性データに基づく企業構造の境界値分析」, 2013年度統計関連学会連合大会講演報告集, 大阪大学豊中キャンパス.
- 中小企業庁ホームページ. (2013). 『中小企業・小規模企業者の定義』 . <http://www.chusho.meti.go.jp/faq/faq01.html>. (2013年6月27日アクセス).
- 松田芳郎, 清水雅彦, 舟岡史雄. (2003). 『講座ミクロ統計分析第4巻 企業行動の変容—ミクロデータによる接近』, 東京, 日本評論社.
- 蓑谷千凰彦, 縄田和満, 和合肇. (2007). 『計量経済学ハンドブック』, 東京, 朝倉書店.

ご清聴ありがとうございました。