



日本銀行ワーキングペーパーシリーズ

ビジネスサーベイにおける外れ値対応 全国企業短期経済観測調査（短観）のケース

石川篤史*

atsushi.ishikawa@boj.or.jp

遠藤峻介†

shunsuke.endou@boj.or.jp

白鳥哲哉‡

tetsuya.shiratori@boj.or.jp

No.10-J-12
2010年7月

日本銀行
〒103-8660 郵便事業（株）日本橋支店私書箱第30号

* 調査統計局

† 調査統計局

‡ 調査統計局

日本銀行ワーキングペーパーシリーズは、日本銀行員および外部研究者の研究成果をとりまとめたもので、内外の研究機関、研究者等の有識者から幅広くコメントを頂戴することを意図しています。ただし、論文の中で示された内容や意見は、日本銀行の公式見解を示すものではありません。

なお、ワーキングペーパーシリーズに対するご意見・ご質問や、掲載ファイルに関するお問い合わせは、執筆者までお寄せ下さい。

商用目的で転載・複製を行う場合は、予め日本銀行情報サービス局までご相談下さい。転載・複製を行う場合は、出所を明記して下さい。

ビジネスサーベイにおける外れ値対応 全国企業短期経済観測調査（短観）のケース

石川篤史*・遠藤峻介†・白鳥哲哉‡

2010年7月

【要 旨】

全国企業短期経済観測調査（短観）では、設備投資額等の計数項目の母集団推計において、外れ値対応を行っていないため、仮に、他のデータセットから著しく乖離したデータ（外れ値）が母集団を代表していない場合、母集団推計値やその前年比・修正率が母数の真値と乖離することになる。本稿では、前年比・修正率に対して影響が大きいデータを短観における外れ値と位置付けた上で、短観に適した外れ値への対応方法を検討した。検討に当たっては、短観の長所である景気変動を的確に捉える正確性・速報性・理解のし易さを維持しつつ、統計の作成負担にも留意した。具体的には、企業の回答値を母集団推計値の前年比・修正率への影響度合いを示す値に変換した上で、ノンパラメトリックな方法により、他のデータから大きく乖離したものを外れ値とみなした。短観のデータセットでは、上記の値が「0」になる比率が総じて高く、既存のノンパラメトリックな方法（quartile method 等）がうまく適用できないため、独自の対応方法として、1パーセンタイルと99パーセンタイルの距離を基準に外れ値を検出し、その影響を除外する方法を考案した。本稿の対応方法を短観の実データに適用したところ、影響の大きい外れ値の検出や除去に有用であることが示された。

キーワード： 短観、外れ値、層化抽出、ウェイト、stratum jumpers

* 日本銀行調査統計局（E-mail: atsushi.ishikawa@boj.or.jp）

† 日本銀行調査統計局（E-mail: shunsuke.endou@boj.or.jp）

‡ 日本銀行調査統計局（E-mail: tetsuya.shiratori@boj.or.jp）

本稿の作成に当たっては、西郷浩教授（早稲田大学）、椿広計教授（統計数理研究所）の両氏ならびに、一上響氏、亀田制作氏、関根敏隆氏、門間一夫氏をはじめとする日本銀行の各氏から有益なコメントを頂いた。また、本稿で考案した外れ値対応方法は、調査統計局企業統計担当スタッフとの議論に基づいている。ここに記して感謝したい。ただし、本稿に示されている意見は、筆者たち個人に属し、日本銀行の公式見解を示すものではない。また、ありうべき誤りはすべて筆者たち個人に属する。

目 次

1 . はじめに	1
2 . 外れ値対応の考え方と先行事例	1
(1) 外れ値の定義	1
(2) 外れ値の発生原因	2
(3) 外れ値対応の目的適合性	3
(4) 外れ値への対応方法	4
(イ) 外れ値の検出	4
(ロ) 外れ値の処理	6
(5) 外れ値対応の実例	9
3 . 短観における外れ値の発生と対応方法	9
(1) 短観における外れ値対応の基本的な考え方	9
(2) 短観の標本設計の概要と6区分ベースでの集計	11
(3) 短観における外れ値の捉え方	11
(4) 短観に適した外れ値検出方法	13
(イ) 既存の外れ値検出方法の適用	13
(ロ) 短観に適した外れ値検出の考え方	16
(ハ) 外れ度合いの判定方法	17
(5) 外れ値処理方法	17
4 . 実データを用いた実証分析	18
(1) データセット	18
(2) 外れ値の検出にかかる実証分析	19
(3) 外れ値の処理にかかる実証分析	20
5 . まとめ	22
補論/参考文献/図表	25

1．はじめに

全国企業短期経済観測調査（以下、短観）の調査内容は、業況判断DI等を算出する判断項目と、売上高や設備投資額等の計数項目に大きく分けられる。このうち、後者の計数項目については、層化抽出（stratified sampling）によって得られた標本企業の回答値から母集団推計を行っており、その母集団推計値の実額と前年比伸び率、前回調査との間の修正率を算出・公表している。この際、これらの計数項目において、標本企業の回答値が同じ層の他標本企業の回答値から著しく乖離していたとしても、その回答値が誤計数でなければ、特別に対応することなく、そのまま集計対象としている。

ところが、こうした推計方法を採用している場合、著しく乖離したデータ（いわゆる外れ値）が仮に母集団を代表していない時は、算出される母集団推計値が母集団の真の値と大きく乖離することになる。悉皆調査でないため、母集団の真の値は明らかでないが、ごく稀に発生する著しく乖離したデータを集計対象とすることについて、実務担当者からは検討の必要性が指摘されている。

また、景気判断において注目される前年比伸び率や前回調査との修正率が当該データの有無によって変動するため、今後、極めて乖離した外れ値が発生した場合、その影響度合いが分析者にとって無視できなくなる可能性もある。そこで、本稿では、このような外れ値へのより適切な対応方法を検討する。

本稿の構成は以下の通りである。第2章では、標本調査における外れ値対応の考え方および対応方法、その適用例をレビューする。第3章では、短観における外れ値の発生メカニズムとその特徴点を示した上で、第2章で紹介した対応方法を短観のデータに適用した場合の問題点、およびそれを克服するための短観における外れ値の具体的な対応方法を検討する。第4章では、短観の実データを用いて、第3章で考案した対応方法の実証分析を行う。第5章は、全体のまとめである。

2．外れ値対応の考え方と先行事例

（1）外れ値の定義

外れ値（outlier）とは、大部分のデータによって構成される集団から外れたデ

ータを指し、標本調査においてしばしば発生する¹。これは、そもそも標本単位の回答値が他のデータから大きく乖離しているものや、標本単位のウェイトにより重み付けされたデータが他のデータ集団から外れてしまうものがある（Fuller[2009]）。なお、ウェイトとは、標本抽出確率の逆数であり、標本単位が母集団の単位をどの程度（何単位分）代表しているかに他ならない（補論1参照）（Lohr[1999]）。

外れ値に関する統計処理については品質管理や医療分野等で研究が先行しているが、ビジネスサーベイの分野でも、海外を中心に研究が進んでいる。ビジネスサーベイでの外れ値については、1980年代以降研究が本格化しており、Chambers[1986]によって体系的に整理されている。すなわち、標本調査における外れ値は、代表的な外れ値（representative）と代表的でない外れ値（nonrepresentative）の2種類に大別され、前者は「代表性のある（母集団において唯一とみなせない）」値、後者は、「代表性がない（母集団において唯一とみなせる）」値と分類して、後続研究の足掛かりとなっている。なお、実際のデータには誤回答が紛れ込む可能性もあるため、そうした誤回答も代表的でない外れ値に含めている。したがって、実際の作業では、標本企業の回答値の正当性を確認し、誤計数であれば訂正した上で、正当な回答値であるが外れ値となる場合について、適切な対応を検討するステップが求められる。勿論、母集団の真の姿が分からないため、外れ値が代表的か代表的でないかを判断するのは難しい。つまり、標本企業のみ情報しか得られない場合、標本企業以外の企業が外れ値と近い値を示していない（外れ値が代表的でない）とは断定できない。

（2）外れ値の発生原因

前述の通り、標本調査における外れ値発生原因の1つには、記入ミスや設問の誤解釈による誤回答があるが、これに対しては、正しい計数に訂正することが求められる。一方で、回答計数自体が正しくても、外れ値になり得る。このような、正しい回答計数における外れ値の発生原因は大きく2つに分けることができる。

¹ United Nations Statistical Commission and Economic Commission for Europe(2000)では、「データセットの分布の裾に在るデータ」と定義している。

1 つには、標本設計方法に起因するものがある。例えば層化抽出においては、正確な層化が行われなかったため、回答値が他の標本単位の回答値から乖離したり、ウェイトにより重み付けされたデータが他のデータから乖離してしまう。層化は、母集団単位の属性等を基準に行うが、母集団情報の整備には限界があるため、限られた情報を基に行うしかなく、同一層において他の単位と比べて異質な単位が混在する可能性は排除できない。このような外れ値の発生を抑えるためには、標本設計方法の工夫等が求められる。

2 つ目の要因は、標本抽出時点では、正確な層化により、適切な水準の回答値、ウェイトであったものの、次の標本抽出までのタイムラグによって回答値が大きく変動し、回答値やウェイトにより重み付けされたデータが他のデータから乖離してしまう場合である。例えばビジネスサーベイでは、企業の急成長等によって企業の実態が標本抽出時点と乖離することが有り得る。層化抽出においては、企業の急激な変動のため、実質的には層を飛び越えて他の層に移ってしまったと考えられる。これは標本設計の経年劣化とも言え、層化抽出における stratum jumpers と呼ばれている。

ここで、stratum jumpers の具体例を示す。標本抽出の際に第 h 層（ウェイト：50）から抽出された t 年度の設備投資額が 1 億円の企業において、 $t+1$ 年度の設備投資額が 100 億円であったとする。設備投資額の水準からは、第 h' 層（ウェイト：5）にあるべき企業に変化したとみなされる。そうすると、本来、当社は第 h' 層のウェイト（5）により重み付けされる筋合いだが、実際には第 h 層のウェイト（50）で重み付けされてしまう。このため、重み付け後の設備投資額は $100 \text{ 億円} \times 50 = 5,000 \text{ 億円}$ （本来は、 $100 \text{ 億円} \times 5 = 500 \text{ 億円}$ ）と過大な金額になる。この結果、前年度との変化幅（重み付け後）は $5,000 \text{ 億円} - 50 \text{ 億円} = 4,950 \text{ 億円}$ となるため、当データを集計対象として算出される全体の設備投資額（母集団推計値）やその前年比は実態と大きく乖離する可能性が出てくる。

このように、標本設計時点だけでなく、時間の経過に伴って発生する外れ値への対策も外れ値対応のポイントである（Pfeffermann and Rao[2009]）。

（3）外れ値対応の目的適合性

外れ値は母集団推計値への影響力が強く、対応方法次第で母集団推計値が大きく異なる。このため、外れ値対応の目的を明確化し、目的に沿った対応方法

が求められる。

一般的に、標本調査においては、単純無作為抽出法 (simple random sampling (S R S)) 等を用いることにより、推定量のバイアスがゼロ (不偏推定量) となるよう標本設計を行う。こうした中で、外れ値対応を行う推定量は不偏性を失う。しかし、その一方、推定量の分散は抑えることができる。

このように、バイアスと分散がトレードオフの関係にある中で、外れ値対応の目的は、母集団推計値を母数の真値に近づけることと位置付けられる。なお、推定量のよさの基準としては、平均 2 乗誤差 (mean square error (M S E)) がよく用いられる (補論 2 参照) (永田 [1992])。

ただし、実際の標本調査の多くの場合、真の母数が分からないため、M S E を算出することはできない。つまり、外れ値への対応方法には、唯一の正解が無いとも言え、それぞれのデータ特性に応じた実証分析やシミュレーション等を行うことにより、最適と思われる外れ値対応方法を検討する必要がある。

(4) 外れ値への対応方法

外れ値対応を導入する際には、どのような基準によって外れ値と判定するのかが問題となる。また、外れ値と判定された場合、そのデータをそのまま集計対象とするのか、他のデータに近づけるのか、集計から除外するのか等々、種々の選択肢からの判断が求められる。すなわち、外れ値の対応は、外れ値の検出 (detection) と 外れ値の処理 (treatment) の 2 つに分けて整理される。以下では、外れ値への対応方法について、検出と処理に分けて整理する。

(イ) 外れ値の検出

外れ値の検出は、定量的な評価となるため、各データの外れ度合いを示す指標が必要になる。この指標は、様々な算出方法が考案されている (図表 1)。

まず、もっとも基本的な手法が range edit と呼ばれる手法である。標本単位 i のデータを x_i 、位置の尺度 (平均等、度数分布の代表値) を m 、散らばりの尺度 (データがどのような散らばり方をしているかを示す尺度) を s とすると、外れ度合いを示す指標 d_i は、(1) 式のように示される。

$$d_i = \frac{x_i - m}{s} \quad (1)$$

ここで、 d_i が事前に設定した上限値または下限値を超えた場合、外れ値として検出する手法が range edit である。言い換えれば、定数 C_L 、 C_U を用いて許容区間($m - C_L s, m + C_U s$)を設定し、 x_i が許容区間の外にある時に外れ値とみなす。

ただし、range edit では、位置の尺度 m や散らばりの尺度 s について、外れ値に対して頑健でない平均や標準偏差を用いると、大きく乖離した外れ値の影響により、許容区間の幅が大きく振れてしまう (d_i が大きく変動する) ため、やや乖離した外れ値が検出されない状況も起こり得る。これは、masking effect と呼ばれている。そこで、外れ値に対して頑健であり、masking effect への対処が可能な方法として、位置の尺度 m 、散らばりの尺度 s を外れ値に対して頑健な尺度としたノンパラメトリックな手法が考案されている。

その代表例が quartile method と呼ばれる手法である。これは、 q_2 をメジアン、 q_1 を第1四分位数、 q_3 を第3四分位数とした時、位置の尺度を q_2 、散らばりの尺度を $(q_2 - q_1)$ および $(q_3 - q_2)$ と設定し、許容区間を $(q_2 - C_L(q_2 - q_1), q_2 + C_U(q_3 - q_2))$ とする手法である。この quartile method は、外れ値に対して頑健なノンパラメトリックな手法であり、range edit の一種と位置付けられる。実務的には、 C_L 、 C_U の設定方法が課題となるため、過去データを用いた実証分析等により設定する必要がある (Bernier and Nobrega[1998])。

	外れ値に対して頑健でない尺度	外れ値に対して頑健な尺度
位置の尺度 m	平均	メジアン、トリム平均 ²
散らばりの尺度 s	分散、標準偏差	四分位範囲

別の range edit として、前回調査との対比に着目する場合もある。すなわち、 t 期の調査において標本単位 i の前回回答値 ($t - 1$ 期) と今回回答値 (t 期) の比率を r_{it} とし、 r_{it} を用いて外れ値を検出する。例えば、標本単位 i において前回調査の回答値に比べて今回調査の回答値が非常に大きくなった場合は、 r_{it} が非常に大きくなるため、外れ値として検出される。

² データ全体から大きい方および小さい方から同数のデータを除き、残ったデータで算出した平均。

ただし、比率を用いる際は注意が必要である。一般的に、ビジネスサーベイでは、企業規模が小さい程、 r_{it} が振れ易いと言われている。このため、大企業に比べて中小企業ばかり外れ値として検出されてしまう傾向にある。こうした傾向は、size masking effect と呼ばれている。また、 r_{it} の分布は、左端が0となる一方、右端は理論上 1 になるので（回答値が非負値の場合）、分布の左端からは外れ値が検出され難いとの指摘もある。

こうした中、Hidiroglou and Berthelot[1986]では、定期的を実施するビジネスサーベイの外れ値検出において、上述の問題に対処するための手法を提案しており、ビジネスサーベイにおける外れ値対応の先駆けとなった（Hidiroglou-Berthelot method（補論3参照））。Hidiroglou-Berthelot method は、上記の quartile method を修正した手法と位置付けられる。具体的には、 r_{it} を変数変換したり、独自に散らばりの尺度を設定することで、データ分布の両端から外れ値を検出する、size masking effect に対処する、散らばりの尺度である $(q_2 - q_1)$ 、 $(q_3 - q_2)$ が極めて小さい時にも適切に外れ値を検出すること等を企図している。

上述の3つの手法は、ノンパラメトリックな手法であるが、外れ値の検出には、パラメトリックな手法であるスミルノフ・グラブス検定等が用いられることもある。しかし、データに対して正規性等、特定の分布が仮定できない場合は、こうした手法の適用は難しい。

標本調査における外れ値の検出手法については、これら以外にも様々な手法が提案・検討されているが、いずれの手法にも長所・短所があり、対象になるデータ特性を踏まえた上で、最適な手法の選択・設定を行う必要がある。

（ロ）外れ値の処理

次に、実際に検出された外れ値を母集団推計に際してどう処理するかが課題となる。外れ値の処理方法には様々な手法があるが、その特徴によって大まかに分類すると、標本単位のウェイトを修正する方法（weight modification）、標本単位の回答値を修正する方法（value modification）、標本単位のウェイトと回答値の双方を修正する方法の3つと³、外れ値自体を修正するのではなく、

³ 外れ値の影響を制限するために、回答値やウェイトにより重み付けされたデータを事前に設定した値等に置き換える方法は、winsorization と呼ばれている。

外れ値の影響を緩和する推定方法(robust prediction)に分けることができる(図表2)。以下では、その概要を示す。

標本単位のウェイトを修正する方法

この方法では、回答値のウェイトを調整し、外れ値の影響度を低下させる。この際、回答値の修正は行わない。例えば、ウェイトによって重み付けされた標本単位のデータが非常に大きいため外れ値として検出された際は、当該回答値のウェイトを低下させることによりデータを下方修正する(weight reduction)。こうしてウェイトを低下させた分、外れ値でない単位のウェイトは増加させる。

ここで、単純な例を示す。第 h 層の母集団単位数 N_h 、標本単位数 n_h の層化抽出において、第 h 層のウェイト $w_h (= N_h/n_h)$ は全て同値となる。標本単位 i の回答値を x_{hi} とすると、第 h 層の合計値 X_h の母集団推計値は(2)式のように示される。

$$\widehat{X}_h = w_h \sum_{i=1}^{n_h} x_{hi} \quad (2)$$

この標本のうち、外れ値として検出された単位のウェイトを w'_h に修正することを考える。外れ値の集合を s_1 、外れ値でない集合を s_2 、 s_1 に含まれる標本単位数を n_{h1} 、 s_2 に含まれる標本単位数を $n_{h2} (= n_h - n_{h1})$ とし、ウェイトの総和が N_h となることを踏まえると、外れ値処理後における X_h の母集団推計値は(3)式で示される。

$$\widehat{X}_h = w'_h \sum_{s_1} x_{hi} + \frac{N_h - n_{h1}w'_h}{n_{h2}} \sum_{s_2} x_{hi} \quad (3)$$

これは、外れ値か否かによって、事後層化(poststratification)を行っていると位置付けられる。特に、 w'_h を「0」に設定すると、外れ値を集計から外す扱いとなる。ただし、ウェイトを修正する方法では、適切なウェイトの設定が難しいほか、同一層の標本単体に異なるウェイトが設定されてしまう。

標本単位の回答値を修正する方法

この方法は、標本単位のウェイトは修正せずに、回答値を修正することで、

外れ値を処理する。主な修正方法としては、(a) 他の標本単位の値を代入する、(b) 事前に設定した値 (cutoff point) を超えた場合 (上限値を上回った場合および下限値を下回った場合) cutoff point を代入する、(c) 外れ値は推定に用いるデータとして不適当であるとみなし、欠測値として捉えて欠測値補完 (imputation) を用いる方法等がある (Chambers and Ren[2004])。このうち、欠測値補完については、他の標本単位の回答値を代入、同一単位の過去の回答値を代入、層の平均値を代入等、様々な手法が考案されている (図表 3) (宇都宮・園田[2001])。

これらの回答値を修正する方法は、他の標本単位の影響を与えないため、実務上、比較的簡単に適用可能であるが、どの値を代入するか、どのようにして cutoff point を設定するかが課題となる。

標本単位のウェイトと回答値の双方を修正する方法

上記の他に、標本単位のウェイトを修正する方法と標本単位の回答値を修正する方法を組み合わせた方法もある。つまり、ウェイトによって重み付けされた標本単位のデータが cutoff point (上限値を K_U 、下限値を K_L とする) を超えた場合、cutoff point を代入するのではなく、cutoff point を超えた部分については、ウェイトを「1」として cutoff point に加え、その値を代入する ((4) 式)。

$$w_i x_{hi} = \begin{cases} w_i x_{hi} & \text{if } K_L < w_i x_{hi} < K_U \\ K_U + \left(x_{hi} - \frac{K_U}{w_i}\right) & \text{if } K_U \leq w_i x_{hi} \\ K_L - \left(\frac{K_L}{w_i} - x_{hi}\right) & \text{if } w_i x_{hi} \leq K_L \end{cases} \quad (4)$$

外れ値の影響を緩和する推定方法

この方法は、標本単位の回答値やそのウェイトを修正するのではなく、推定値を算出する際、そもそも外れ値の影響を受け難い統計モデルを用いることで外れ値の影響を緩和するものである。簡単な一例として、(5) 式を用いた推定方法がある。ここで、標本単位 i の回答値を x_i 、標本単位数を n 、母集団単位数を N 、外れ値に対して頑健な位置の尺度 (メジアンやトリム平均等) を \tilde{x} とし、推定対象である総合計 X を以下により求める。

$$\hat{X} = \sum_1^n x_i + (N - n)\tilde{x} \quad (5)$$

これは、 n 個の標本単位の回答値を使いつつ、 $(N - n)$ 個の非標本単位の推定値として、位置の尺度 \tilde{x} を代入することにより、外れ値の影響を緩和している。この他にも、様々な統計モデルが提案されているが、それらの統計モデルは、しばしば構造が複雑になってしまうとの問題が指摘されている⁴。このため、ビジネスサーベイにおける適用例は殆どみられない。

(5) 外れ値対応の実例

外れ値対応について、国内の経済統計調査では、明確な基準を設けて外れ値の検出・処理を行っている事例はあまりみられない。一方で、海外においては、外れ値対応を導入している事例がみられる(図表4)。

例えば、カナダの、『Survey of Employment, Payrolls and Hours (SEPH)』(Statistics Canada)、『National Construction Industry Wage Rate Survey』(Statistics Canada)、アメリカの、『State and Metro Area Employment, Hours, & Earnings』(Bureau of Labor Statistics)、オーストラリアの、『Consumer Price Index』(Australian Bureau of Statistics)等がある。

具体的な対応方法は、 \hat{X} は、検出方法に Hidiroglou-Berthelot method や quartile method 等を、修正方法に weight modification 等を用いている。 \tilde{x} は、位置の尺度に「平均」、散らばりの尺度に「標準偏差」を用いた range edit により外れ値を検出し、ケースバイケースで処理を行っている。 \hat{X} では、weight reduction により、外れ値の影響力を低下させている。 \tilde{x} では、winsorization によって、外れ値の次に大きい値と同値に修正している。このように、外れ値の対応方法は統計によって区々である。

3. 短観における外れ値の発生と対応方法

(1) 短観における外れ値対応の基本的な考え方

短観は層化抽出を用いた標本調査であり、短観の母集団は総務省の「事業所・

⁴ 詳細は、Barnett [1994]、Barnett and Lewis[1994]参照。

企業統計調査」をベースとした全国の資本金 2 千万円以上の民間企業（約 21 万社）、標本の大きさは約 1.1 万社である。これらの標本企業を、業種、資本金および雇用者数の 3 変数によって層化している。

短観では、何重もの Micro-Editing、Macro-Editing によって回答計数の正当性を確認しているため、企業の回答値が誤りである可能性は基本的に排除している⁵。一方で、標本設計については、層化時点では適切に層化しているものの、調査回毎に標本抽出を行っている訳ではなく、一定期間、母集団・標本企業を固定しているため、層化時点と調査時点でタイムラグが生じている⁶。このために、企業の急成長や形態の変化等に起因する stratum jumpers の発生は不可避であり、外れ値の発生要因となり得る。

そこで、短観においては、以下のような統計の優位性を損なうことのないように留意しながら、外れ値対応を行う必要がある。

正確性の維持

積極的に母集団推計値やその前年比・修正率を均してしまうと、景気の転換点やその兆候を捉えきれなくなり、必ずしもユーザーニーズを満たした統計とはならない。基本的には、標本企業から得られた調査項目毎の回答は、それぞれの母集団の姿を適切に反映していると考えられるため、そのまま集計に用い、影響が極めて大きな値となる調査項目のみが検出される工夫が必要となる。

理解し易さの維持

短観はDIの算出方法や調査項目の定義等、内容の分かり易さも特徴である。また、計数項目における母集団推計方法もユーザーが理解し易いシンプルかつクリアな方法となっている。外れ値対応の方法も、過度に複雑でないものが望ましい。また、透明性を担保するため、予め恣意的な判断余地を排除する工夫が必要となる。

⁵ Micro-Editing とは、個々のデータをチェックしてエラーを特定すること。Macro-Editing とは、集計値やデータの分布からエラーを特定すること。

⁶ 短観では、継続標本をベースにすることにより、高い回答率、誤回答の削減を実現し、非標本誤差の最小化を図っているほか、母集団情報の更新および標本の再抽出が前年比・修正率に与える影響を抑えている。

作成負担の抑制

短観は、毎年3、6、9、12月と四半期毎の調査であるが、できる限り直近の景気動向を把握するため、調査期間最終日の翌日に公表する俊敏な調査が求められている。過度に手間がかかる対応方法では、統計作成事務にかかる負担が増加し、タイムリーな短観の調査・公表に支障を来す可能性がある。そこで、統計作成の実務コストが大きく嵩まない方法が望まれる。

(2) 短観の標本設計の概要と6区分ベースでの集計

冒頭にも述べたように、短観では、母集団推計を行っている売上高や設備投資額等の計数項目は、標本から層の平均を推定し、その平均を層の母集団企業数で乗じることで層の合計を推定する。こうして推定された各層の合計を足し合わせることで、全体の合計を推定している。これは、各企業の回答値をウェイトで重み付けした値の合計と同値である。

また、計数項目の集計は、母集団全体の合計のみでなく、企業規模（資本金を基準に、大企業＜資本金10億円以上＞、中堅企業＜同1億円以上10億円未満＞、中小企業＜同2千万円以上1億円未満＞）や業種別にも行っている。この中で、特に、統計ユーザーの注目度が高いのは、大・中堅・中小企業別、製造・非製造業別の6区分（以下、6区分ベースという）における母集団推計値の前年比（(6)式）および修正率（(7)式）である。

$$\text{前年比(\%)} = \frac{\text{今年度の母集団推計値} - \text{前年度の母集団推計値}}{\text{前年度の母集団推計値}} \times 100 \quad (6)$$

$$\text{修正率(\%)} = \frac{\text{今回調査の母集団推計値} - \text{前回調査の母集団推計値}}{\text{前回調査の母集団推計値}} \times 100 \quad (7)$$

(3) 短観における外れ値の捉え方

先述の通り、短観では、母集団推計値の前年比や修正率の注目度が高く、その中でも、6区分ベースに対する統計ユーザーの注目が高い。そこで、本稿では、外れ値対応を行う際には、各企業の回答値について、6区分ベースでの母

集団推計値の前年比や修正率に対する影響度合いに着目した⁷。

6区分ベース毎に外れ値対応を行うことは、size masking effectのように各企業の規模によって影響度合いに違いが生じる可能性に対処する観点や、各区分1千社以上の標本企業が確保され、サンプルサイズ上の問題がないことからも適切と判断できる。

具体的には、 t および $t-1$ を時期（前年比の場合は、今年度および前年度、修正率の場合は、今回調査および前回調査）第 i 番目の企業における回答値を x_{it} 、そのウェイトを w_{it} 、6区分ベースの母集団推計値を X_{jt} とすると、第 i 番目の企業が j 区分($j = 1, \dots, 6$)の前年比、または、修正率に及ぼす影響度は、

$$y_{it} = \frac{\text{重み付けされたデータの変化幅}}{|\text{前年度(前回)の6区分ベースの母集団推計値}|} \times 100$$
$$= \frac{w_{it}x_{it} - w_{it-1}x_{it-1}}{|X_{jt-1}|} \times 100 \quad (8)$$

と定義される。

ここで、2時点 t 期および $t-1$ 期におけるウェイト w_{it} 、 w_{it-1} は、標本企業数が同じであれば同値となるが、通常、調査回によって標本企業数は変わることが多いため、2つのウェイトは異なる⁸。しかし、そうしたウェイトの変化によってもたらされる「重み付けされたデータの変化幅」は、当該企業の実質的な変動ではない。そこで、予め、2時点のいずれかで回答が存在しない場合には、cell mean imputationによる欠測値補完を行うことにより標本企業数を揃える。こ

⁷ 経常利益や当期純利益の母集団推計値は負の値を取り得るため、実際には推計値の絶対値に対する影響度をみる。また、毎調査回、前年比と修正率の両方について外れ値対応を行うことも考えられるが、修正率の変動を抑えれば、前年比の変動幅も抑制されるため、修正率のみに着目する。ただし、初回調査（3月調査）においては、修正率が存在しないため、前年比に着目する。

⁸ 母集団企業数は、「事業所・企業統計調査」の実施（2、3年毎）に合わせて見直しを行っているため、2時点 t 期および $t-1$ 期におけるウェイトの変化は標本企業数に依存する。なお、母集団企業の見直しの際は、見直し前後両ベースのデータを用いて2系列の前年比を算出しているため、 t 期および $t-1$ 期の母集団企業数は常に一致している。

うして算出されたウェイトを w_i とする。つまり、当ウェイトは、各層の母集団企業数を t 期、 $t-1$ 期の少なくとも一方で回答が存在する企業数で除したものとなる。こうした調整の結果、(8)式は、

$$y_{it} = \frac{w_i \cdot (x_{it} - x_{it-1})}{|x_{jt-1}|} \times 100 \quad (9)$$

と表すことができる。この、6区分ベースに対する影響度合い y_{it} のばらつき度合いをみて、他のデータから非常に大きく外れているものが外れ値となる。

(4) 短観に適した外れ値検出方法

本節では、短観に適した外れ値の検出方法を検討する。短観の外れ値検出については、前節で示した影響度合い y_{it} が他のデータから非常に大きく外れているもののみを効果的に検出できる方法が望ましい。そこで、まずは、2.(4)(イ)項で示した既存の外れ値検出方法が、短観の外れ値検出に適するかを検討するため、短観の実データを用いた検証を行った。

(イ) 既存の外れ値検出方法の適用

quartile method を用いる場合

まず、quartile method を用いる場合を検討する。短観の主要項目(売上高・経常利益・当期純利益・設備投資額・ソフトウェア投資額)について、10調査回分(07年3月調査から09年6月調査)の影響度合い y_{it} の分布をみると、 $y_{it} = 0$ (企業の回答値が前年度や前回調査と同値)となる企業の割合が非常に高く、また、必ずしも、影響度合い y_{it} の分布は左右対称にならない(理論上、売上高や設備投資額等、回答値が負の値を取らない場合、マイナス側は最大100、プラス側は最大)ため、正規分布とみなすことは困難である⁹(図表5-1~5-2、6-1~6-10)。特に中小企業では、そもそも設備投資やソフトウェア投資を行っていない企業が多いため、その変化幅も「0」となるケースが散見される。例えば、中小企業・製造業のソフトウェア投資額をみると、平均して9割近くの影響度合い y_{it} が「0」となっていることが分かる(図表5-2(4))。

⁹ 因みに、影響度合い y_{it} は、負の値を取り得るため、そのまま対数をとることができない。

このように、影響度合い y_{it} は、「0」の割合が非常に高く、四分位範囲が「0」となってしまう場合が多いため、位置の尺度をメジアン、散らばりの尺度を第1四分位数とメジアンの距離、または、第3四分位数とメジアンの距離とした quartile method をそのまま適用することは難しい。

比率を用いる場合

次に、比率を用いた検出方法について検証する。具体的には、2009年6月調査における売上高の2009年度修正率について、各企業の前回調査（2009年3月調査）と今回調査（2009年6月調査）の回答値から比率 r_{it} （今回回答値/前回回答値）を算出し、比率 r_{it} が事前に設定した境界値（上限値： T_U 、下限値： T_L とする）を超えたものを外れ値とみなす¹⁰。観察されたポイントは以下の通り（図表7-1～7-2、図表8）。

- ・ 比率 r_{it} が他のデータから大きく乖離しても、必ずしも影響度合い y_{it} が大きく外れている状況には無い（図表7-1～7-2）。例えば、比率 r_{it} が大きく乖離していても影響度合い y_{it} が「0」近傍であるデータがいくつかみられる。つまり、外れ値として検出すべきでないデータを外れ値として検出してしまうリスクがある。
- ・ 一方で、比率 r_{it} が相対的に大きくないにも拘わらず、影響度合い y_{it} が相対的に大きいデータもみられる（図表7-1～7-2）。つまり、外れ値として検出すべきデータを検出し損ねてしまうリスクがある。
- ・ 比率 r_{it} が上限値を超える割合は、企業規模が小さくなるにつれて、高くなる傾向にあり、企業規模毎に区分しない場合に、size masking effect が起きている可能性が示されている（図表8（1））。

このように、比率を用いる方法では、影響度合い y_{it} が大きく外れているものを効果的に検出することができない。また、短観の調査項目には、設備投資額やソフトウェア投資額のように回答値が「0」となる割合が高い項目や、経常利益や当期純利益のように回答値が負の値を取り得る項目があり、比率 r_{it} の算出が難しくなるため、比率を用いる方法は適さないと考えられる。

¹⁰ 前回回答値が「0」の場合、比率 r_{it} が算出できないため、当実証のデータセットからは除いている。

Hidiroglou-Berthelot method を用いる場合

上と同様に、(補論3)で示した Hidiroglou-Berthelot method についても検証する。上記の比率を用いる方法と同様に、2009年6月調査における売上高の2009年度修正率に対して外れ値の検出を行った¹¹。観察されたポイントは以下の通り(図表9-1~9-2、図表10)。

- ・比率 r_{it} と同様、比率 r_{it} を変数変換した E_{it} においても、他のデータから大きく乖離したデータであっても必ずしも影響度合い y_{it} が大きく外れている状況には無い(図表9-1~9-2)。例えば、比率 r_{it} を変数変換した E_{it} が大きく乖離していても影響度合い y_{it} が「0」近傍であるデータがいくつかみられる。つまり、外れ値として検出すべきでないデータを外れ値として検出してしまうリスクがある。
- ・一方で、比率 r_{it} を変数変換した E_{it} が相対的に大きくないにも拘わらず、影響度合い y_{it} が相対的に大きいデータもみられる(図表9-1~9-2)。つまり、外れ値として検出すべきデータを検出し損ねてしまうリスクがある。
- ・比率 r_{it} を変数変換した E_{it} の分布について、メジアンおよび第3四分位数がともに「0」となっている区分がみられる(中堅企業・製造業、中堅企業・非製造業、中小企業・製造業)(図表10(1))。このため、散らばりの尺度 D_U が「0」となってしまい、適切な許容区間の設定ができない。この結果、上限値が「0」となるため、僅かな変動のデータが外れ値として検出されてしまい、2割前後のデータが外れ値となっている(図表10(3))。

このように、Hidiroglou-Berthelot method についても、影響度合い y_{it} が大きく外れているものを効果的に検出することは難しい。また、最初の段階で前回回答値と今回回答値の比率を算出する必要があるため、回答値が「0」となる割合が高い設備投資額、ソフトウェア投資額や、負の値を取り得る経常利益、当期純利益への適用は難しいと考えられる。

¹¹ 未知パラメータ U 、 A については、Belcher[2003]、Banim[2000]、Hunt[1999]を参考に、それぞれ、0.2、0.05 とした。また、 c については、20、30、40 の3通りで試算を行った。なお、 $r_{it} < r_{it}^M$ の場合、 $r_{it} = 0$ の時に s_{it} が算出できないため、当実証のデータセットからは除いている。

仮説検定を用いる場合

最後に、仮説検定を用いる方法については、短観の回答値やそこから算出される影響度合い y_{it} に対して正規分布等、特定の分布を仮定することが難しいため、パラメトリックな仮説検定は適切でないと考えられる。

こうした結果を踏まえると、短観のデータで効率的に外れ値を検出するためには、range edit を影響度合い y_{it} に対して直接適用することが望ましいと考えられる。

(口) 短観に適した外れ値検出の考え方

以上の結果を踏まえて、影響度合い y_{it} の外れ度合いを定量的に把握するため、外れ度合いの指標 z_{it} を定義する。ここでは、masking effect に対処するため、位置の尺度および散らばりの尺度に、外れ値に対して頑健な尺度を用いる。具体的には、位置の尺度は、母集団推計値に対する影響度合い y_{it} の分布が左右非対称であることを踏まえ、メジアンやトリム平均ではなく、1パーセンタイル(d_j^1 とする)および99パーセンタイル(d_j^{99} とする)を用いる。また、散らばりの尺度については、四分位範囲の代わりに、1パーセンタイルと99パーセンタイルの距離を単位距離 D_j と定義し、散らばりの尺度とする。

$$D_j = d_j^{99} - d_j^1 \quad (10)$$

この単位距離 D_j であれば、安定的に $D_j > 0$ となるため、散らばりの尺度として用いることができる(図表 11 - 1 ~ 11 - 2)。なお、仮に $D_j = 0$ となった時は、直近4調査回の平均値を用いることにする。以上より、 z_{it} は下式の通り定義される。

$$z_{it} = \begin{cases} \frac{y_{it} - d_j^{99}}{D_j} & \text{if } d_j^{99} \leq y_{it} \\ \frac{d_j^1 - y_{it}}{D_j} & \text{if } y_{it} \leq d_j^1 \\ 0 & \text{if } d_j^1 < y_{it} < d_j^{99} \end{cases} \quad (11)$$

つまり、実質的に外れ値の検出対象となるのは、全データのうち、影響度合い y_{it} の上位・下位、それぞれ1パーセントのデータのみである。

(八) 外れ度合いの判定方法

最後に、外れ度合いの指標 z_{it} を用いて外れ値の判定方法を検討する。具体的には、過去データを用いた実証分析により事前に基準を設定し(基準Cとする)、 z_{it} が基準Cを超えた場合、外れ値として検出することとする(図表12)。

この際、検出方法の分かり易さを考慮し、全区分・項目においてCは同値とする。なお、外れ度合いの指標 z_{it} 、すなわち、 D_j 、 d_j^1 、 d_j^{99} は6区分ベース・調査項目毎に算出されるため、既に、各区分・調査項目のデータ特性を反映していると言える。

(5) 外れ値処理方法

本稿では、短観の外れ値検出の際、短観の目的に照らして、外れ値の検出数が嵩まないことを優先した。つまり、調査項目毎に非常に影響力の大きいデータのみを外れ値として検出するため、母集団において他に同様なデータが存在する蓋然性は低いと考えられる。加えて、当該外れ値は、stratum jumpers であるため、当該調査項目については、適切でない層に層化されていると判断される。換言すれば、外れ値として検出された当該調査項目のデータは代表的でないとの判断になる。一方で、当該企業の他の調査項目は外れ値として検出されなければ、代表的であると判断する。こうした点を踏まえると、短観の場合には、仮に企業の回答が正しい値であったとしても、外れ値として検出された以上、回答値を母集団推計には用いない、すなわち、外れ値として検出された調査項目は欠測値とみなすことが適当だと考えられる¹²。

現行短観では欠測値補完方法について、今年度または今回調査の回答値に対して、前年度または前回調査と同値を代入する方法($x_{it} = x_{it-1}$ 、とする)を2004年から採用している。このため、検出された外れ値の処理方法は、現行ルール

¹² 外れ値を他層に移した上で推定に用いる点については、短観では層化変数に「業種」、「資本金」、「雇用者数」を用いている中、ある調査項目が外れ値と検出された場合でも、層の設定は層化変数に依存するため、当該外れ値が属する真の層を一意的に決めることが難しいという問題がある。

と同じ欠測値補完方法を適用することが整合的である。この方法では、外れ値の母集団推計値の前年比や修正率に対する影響度合い y_{it} は「0」となり、外れ値の影響が完全に排除されることになる¹³（この処理方法を以下では、「(a)横置き補完」と呼ぶ）。

4．実データを用いた実証分析

本章では、上記の外れ値の検出・処理方法に基づき、短観の実データに対して、外れ値の検出・処理を行った。検出については、事前に設定する基準Cの値と、外れ値として検出されるデータの個数について考察した。処理については、検出されたデータに対して3.(5)節で示した外れ値の処理を行い、母集団推計値への影響を試算した。

(1) データセット

実証分析に用いたデータセットについて説明する。まず、調査項目は、短観の主要項目である売上高・経常利益・当期純利益・設備投資額・ソフトウェア投資額の5項目とする。次に、データ期間は、07年3月調査から09年6月調査の10調査回分とする。よって、売上高・経常利益・当期純利益は90系列（15系列×6区分）、設備投資額・ソフトウェア投資額は84系列（14系列×6区分）となる¹⁴。

¹³ 理論的には、 $y_{it} = 0$ の値が再び外れ値となる場合も有り得る。例えば、ほぼ全ての企業が前回と比べて極めて大幅に上方修正を行った結果、 y_{it} の平均値が「0」から極めて大きく乖離すると、 $y_{it} = 0$ における z_{it} の値が極めて大きくなる。しかし、過去のデータ特性をみる限り、実際にこのような状況は想定され難い。

¹⁴ 設備投資額・ソフトウェア投資額は、リース会計基準の変更に伴い調査項目の定義を変更したため、09年3月調査の08年度修正率は算出していない。このため、他調査項目と比べて系列数が少なくなっている。

データセット	
項目 (5 項目)	売上高・経常利益・当期純利益・設備投資額・ソフトウェア投資額
データ期間 (10 調査回)	2007 年 3 月調査から 2009 年 6 月調査まで 売上高・経常利益・当期純利益：90 系列 (15 系列 × 6 区分) 設備投資額・ソフトウェア投資額：84 系列 (14 系列 × 6 区分)

(2) 外れ値の検出にかかる実証分析

(検出データ数)

上記データセットに対して外れ値の検出を行った。基準 C の設定と外れ値の検出データ数からみるポイントは以下の通り (図表 13 - 1 ~ 13 - 10、14)。

- ・ C = 25 とすると外れ値の検出数は全項目の合計で 14 個となった。つまり、平均すると調査回毎に 1.4 個の外れ値が発生することになる (図表 14)。
- ・ C = 50 とすると検出数は大きく減少し、設備投資額およびソフトウェア投資額が 1 個ずつの計 2 個となった。C = 25 と比べると、前年比や修正率に対して、非常に大きな影響力をもつデータのみ外れ値として検出された (図表 14)。
- ・ C = 100 とすると検出は設備投資額の 1 個のみとなり、C = 125 とすると一切検出されなかった (図表 14)。

(調査項目別の特徴)

上記検出データ数を調査項目別にやや詳しくみると、売上高については、C = 25 においても、外れ値は検出されなかった。母集団推計値に対する影響度合い y_{it} が安定的に一定の範囲内で収まっている中、単位距離 D_j も各調査回において大きく変動しないため、外れ度合いの指標 z_{it} が相対的に大きな値にならないためと考えられる。

利益項目 (経常利益・当期純利益) については、C = 25 において、2 個外れ値が検出された。利益項目は売上高や設備投資額等と異なり、各企業の回答値が負の値を取り得るため、母集団推計値の絶対値が比較的小さい値となることで影響度合い y_{it} が振れ易くなる。例えば、図表 6 - 5 (1) より、2009 年 3 月調

査の当期純利益（大企業・製造業・2009年度前年比）をみると、6区分ベースの当期純利益（2008年度）の実額が平年比極端に小さいことも影響して、影響度合い y_{it} の最大値が「2,291」（大企業・製造業の当期純利益の前年比に対する当該データの寄与度が+2,291%ポイントという意味）と、非常に大きな値を示している。しかし、他の企業についても当期純利益の変動が相対的に大きい企業が多いため、単位距離 D_j も他調査項目に比べて大きい値を示している（図表11-1）。その結果、当該データの外れ度合いの指標 z_{it} をみると、一定の範囲内に収まっており、外れ値として検出する必要は無いとの結果となった。このように、仮に影響度合い y_{it} が大きく振れた場合でも、外れ度合いの指標 z_{it} の算出の際は、単位距離 D_j により均されるため、実質的にデータ分布から外れていなければ、外れ値として検出されない仕組みとなっていることが確認された。

投資項目（設備投資額、ソフトウェア投資額）については、 $C=25$ において、12個、 $C=50$ において、2個の外れ値が検出された。設備投資額の中小企業・製造業やソフトウェア投資額の中堅企業・非製造業では、外れ度合いの指標 z_{it} が非常に大きいデータが存在している（図表13-7（3）、13-10（2））。投資項目は、利益項目と比べて全体的に影響度合い y_{it} は安定している。ただし、稀にはあるが影響度合い y_{it} の絶対値が「10」前後のものも存在しているほか、「5~10」程度のものは散見されている。単位距離 D_j は比較的安定しているため、外れ度合いの指標 z_{it} は、影響度合い y_{it} に連動する傾向がみられた。

以上の実証結果は、売上高や利益項目は外れ値が比較的発生し難い一方、設備投資額やソフトウェア投資額は、特に中堅・中小企業において、外れ値が相対的に発生し易い可能性を示唆している。このように基準 C の設定値により、外れ値の検出数が変わるため、外れ値対応のルール化においては、境界値の決め方が重要になる。

（3）外れ値の処理にかかる実証分析

次に、外れ値として検出されたデータについて、3.（5）節で述べた外れ値処理方法（「(a)横置き補完」）を適用し、母集団推計値への影響を試算した。その際、欠測値補完方法の違いによる母集団推計値への影響を比較するため、「(a)横置き補完」以外の補完方法についても試算した。1つ目は、欠測値が属する層の平均値を代入する cell mean imputation（すなわち、当該回答値のウェイトを

「0」として集計から外すのと同義)である(この処理方法を以下では、「(b) 平均値補完」と呼ぶ)。2つ目は、短観のユーザーが母集団推計値の前年比や修正率に着目していることを踏まえ、外れ値が属する層全体の前年比や修正率を利用した処理方法である(この処理方法を以下では、「(c) 伸び率補完」と呼ぶ)。具体的には、外れ値が発生した層において、外れ値以外の回答値から算出される層全体の前年比または修正率を外れ値の前年度計数および前回計数に乘じ、この計数を今年度または今回調査の回答値に対して代入する。結果として、外れ値は当該層の前年比や修正率に影響を与えないため、外れ値処理後における当該層全体の前年比または修正率は、外れ値以外の回答値から算出した場合と同値となる。

以下では、本実証分析において外れ度合いの指標 z_{it} の値が大きい上位3つのケース(以下の ~)を外れ値とみなし、それぞれの処理方法が6区分ベースの修正率にどの程度影響を及ぼすか試算した。

2007年6月調査、中小企業・製造業、2006年度設備投資額(外れ度合いの指標 $z_{it} = +104.81$)

2007年6月調査、中堅企業・非製造業、2007年度ソフトウェア投資額(外れ度合いの指標 $z_{it} = +82.18$)

2008年9月調査、中小企業・製造業、2008年度設備投資額(外れ度合いの指標 $z_{it} = +46.11$)

試算結果の概要は以下の通りである(図表15)。

(a) 横置き補完の場合、外れ値の6区分修正率への影響を完全に取り除いた(影響度合い y_{it} を「0」とした)ため、6区分修正率がそれぞれ10%ポイント前後変化している。

(b) 平均値補完の場合、図表15(2)の や で、外れ値処理前と比べて、影響度合い y_{it} の正負が逆転している。これは、当該回答値が実際には上方修正されたにも拘わらず、平均値が補完されたことにより下方修正されたことを意味する。一方で、 では、当該外れ値が、外れ値処理前よりも更に他のデータから乖離する結果となっている。

(c)伸び率補完の場合、図表 15(2)の や のケースでは、影響度合い y_{it} が相対的に小さい値となり(0.07、 +0.04) (a)横置き補完と比べて、相対的に大きな差はみられない。しかし、 のケースでは、外れ値処理後の影響度合い y_{it} が再び相対的に大きな値を示している ($y_{it} = +7.38$)。これは、一度外れ値処理を行ったデータが、再び外れ値となり得る可能性を示している。

3つの処理方法の特徴を比較すると、(a)横置き補完は、外れ値の影響を取り除くという所与の目的は達成している。また、実務上、シンプルである。しかし、この方法は、正・負何れの方向にせよ、外れ値以外のデータで算出した前年比や修正率の大きさ(絶対値)を小さくしてしまう。(b)平均値補完では、外れ値として検出されたデータの前年度または前回調査回答値が同一層の他の回答値と比べて大きく異なっていた場合、今年度または今回調査回答値に層の平均値を代入することにより、前年度または前回調査から大きな変動が起きてしまう。(c)伸び率補完は、外れ値を除いて算出した「外れ値が属する層全体の前年比や修正率」に影響を与えない。ただし、外れ値が同一層の他の回答値に比べて大きな値を示している場合、層全体の前年比や修正率を乗じた値が、母集団推計値に大きな変動を与え得るという特徴があるほか、経常利益のように負の値を取り得る項目について伸び率を定義できない場合の扱い等も検討する必要がある。

ただし、外れ値が発生した時点での母集団の真の姿が分からないため、実証分析で最適な処理方法を1つに選択することは困難であるほか、シミュレーションを行う場合でも、シミュレーションの結果から最適と判断される処理方法は仮想母集団でのデータ設定等に依存するため、統計学的に最適な処理方法を1つに決めることは難しい。しかし、一定のルールに基づき極めて影響の大きいデータのみが外れ値として検出されるため、一切の外れ値処理を行わない現行の方法と比べて、外れ値の処理を行う方が相対的に優位であることを示唆していると考えられる。

5 . まとめ

本稿では、短観における外れ値を母集団推計値の前年比・修正率に対する影響が極めて大きいものと位置付け、短観の実データを用いた実証分析を通じて、短観の特性に見合った検出・処理方法を検討した。この際、既存の手法を勘案

しつつ、短観の正確性、理解し易さの維持、作業負担の抑制を念頭に、短観に馴染み易い方法を検討した。具体的には、検出方法については、既存の手法(range edit) をベースに、「0」の割合が非常に高いノンパラメトリックなデータに適用できるよう、位置の尺度や散らばりの尺度を1パーセンタイルおよび99パーセンタイルを用いて定義することにより、安定的な検出が行える方法とした。検出された外れ値は欠測値とみなし、現行短観が採用している欠測値補完方法により処理する手法とした。

実証分析の結果をみると、本稿で示した外れ値「検出」方法については、非常に影響力の大きいデータの検出に対して有効であり、外れ値「処理」方法についても、外れ値の影響除去に対して有効となることが示された。また、本稿で示した外れ値対応方法は、実務負担も限定的であり、短観へ導入し易い結果となっている。

今後、実際に外れ値対応を導入する場合には、以下のような実務的な課題についても検討する必要がある。

1つ目は、外れ値の検出の際に設定する基準 C の値を1つに決める必要がある。しかし、基準 C の値を決める際に、客観的な判断基準を設定することが困難であるため、短観の特徴や外れ値対応の目的に照らした上で、「検出すべき全ての外れ値は検出できないことのリスクよりも、本来外れ値ではない多くのデータまでを外れ値として検出してしまうことのリスクを回避する」との立場から、外れ値の検出数が限定的になるように基準 C を設定するのが適当と考えられる。

2つ目は、項目間の整合性の問題である。調査項目毎に外れ値の検出を行うため、検出された企業の全調査項目の回答値を外れ値処理するのではなく、検出された項目のみ外れ値処理を行えばよいが、内訳項目を有する調査項目には留意が必要である。例えば、設備投資額と、その内訳項目である土地投資額について、それぞれ外れ値対応を適用した場合、設備投資額と土地投資額のうち片方のみが外れ値として検出・処理された結果、計数の大小関係が逆転する等、項目間で不整合が起きてしまう可能性がある。よって、土地投資額については単独での検出は行わず、設備投資額が検出された場合に限り、土地投資額も同様の処理(前年度あるいは前回と同値を代入)を行う等のルールが必要となる。

3つ目は、外れ値の次調査回以降の扱いについても検討する必要がある。外れ値として検出されたデータは適切な層に層化されていないと整理されるため、次回の層化を行うまでは、当該層の標本単位とみなすのは望ましくない。つまり、外れ値処理を継続することにより、母集団推計値に対して意図しない影響が発生する可能性を排除することが重要である。

また、一旦導入した外れ値対応方法は、一定期間継続して適用すること、更にその内容を対外的に開示すること等によって、統計全体の透明性を高めていくことも重要である。

(補論1) 層化抽出について

層化抽出とは、母集団を層 (strata) と呼ばれる重ならない部分母集団に分割 (層化) し、各層からそれぞれ独立に SRS を行うものである。層化の基準には、母集団単位の属性等を用いる。層化抽出では、偏りの無い標本抽出を行う、標本の管理コストを抑制する、層化により分散が低下するため推定値の精度が向上するといったメリットが期待される。

大きさ N の母集団を H 個の層に分け、各層の大きさを N_h ($h = 1, \dots, H$) とすると、

$$N = \sum_{h=1}^H N_h$$

となる。第 h 層における i 番目の母集団単位の値を x_{hi} とすると、第 h 層の合計 X_h は、

$$X_h = \sum_{i=1}^{N_h} x_{hi}$$

となる。また、母集団の合計 X は、

$$X = \sum_{h=1}^H X_h$$

で示される。ここで、第 h 層における母集団単位の平均を \bar{x}_h とすると、第 h 層の母分散 S_h^2 は、

$$S_h^2 = \sum_{i=1}^{N_h} \frac{(x_{hi} - \bar{x}_h)^2}{N_h - 1}$$

と表される。各層から SRS により標本を抽出し、母集団の母数を推定する。

ここで、SRSにより第 h 層から抽出された標本を A_h 、標本単位数を n_h とすると、 \bar{x}_h 、 X_h 、 S_h^2 、 X の推定量はそれぞれ下式のように表せる。

$$\widehat{\bar{x}}_h = \frac{\sum_{i \in A_h} x_{hi}}{n_h}$$

$$\widehat{X}_h = \frac{N_h}{n_h} \sum_{i \in A_h} x_{hi} = N_h \widehat{\bar{x}}_h$$

$$\widehat{S}_h^2 = s_h^2 = \sum_{i \in A_h} \frac{(x_{hi} - \widehat{\bar{x}}_h)^2}{n_h - 1}$$

$$\widehat{X} = \sum_{h=1}^H \widehat{X}_h = \sum_{h=1}^H N_h \widehat{\bar{x}}_h = \sum_{h=1}^H \sum_{i \in A_h} \frac{N_h}{n_h} x_{hi}$$

ここで、 N_h/n_h を w_{hi} とおくと、上式は

$$\widehat{X} = \sum_{h=1}^H \sum_{i \in A_h} w_{hi} x_{hi}$$

となる。 w_{hi} は、ウェイトと呼ばれ、第 h 層における標本の抽出確率の逆数であり、ウェイトがかかる標本単位が、母集団単位をどの程度代表しているか示している。例えば、 $N_h = 1000$ 、 $n_h = 10$ の時、 $w_{hi} = 100$ となるため、第 h 層の各標本単位は、当単位の他に標本に含まれない 99 の単位を代表していることになる。つまり、ウェイトの和は、母集団の大きさと同値となる。

$$\sum_{h=1}^H \sum_{i \in A_h} w_{hi} = N$$

また、母集団平均の推定量は、ウェイトを用いて、

$$\hat{\bar{X}} = \frac{\sum_{h=1}^H \sum_{i \in A_h} W_{hi} X_{hi}}{\sum_{h=1}^H \sum_{i \in A_h} W_{hi}}$$

と表すことができる。

(補論 2) 平均 2 乗誤差について

パラメータ θ の推定量 $\hat{\theta}$ の悪さかげんを測る尺度として損失関数(loss function) $L(\hat{\theta}, \theta)$ を考える。これは、 $L(\hat{\theta}, \theta) \geq 0$ で $|\hat{\theta} - \theta|$ について単調増加し、 $\hat{\theta} = \theta$ の時ゼロとなるものと定義する。また、損失関数の期待値を危険関数(risk function) $R(\hat{\theta}, \theta)$ と定義する。

$$R(\hat{\theta}, \theta) = E\{L(\hat{\theta}, \theta)\}$$

ここで、下式を 2 乗損失関数と定義する。

$$L(\hat{\theta}, \theta) = (\hat{\theta} - \theta)^2$$

また、2 乗損失関数の期待値(危険関数) を平均 2 乗誤差と呼ぶ。

$$R(\hat{\theta}, \theta) = E\{(\hat{\theta} - \theta)^2\}$$

平均 2 乗誤差は、推定量の分散とバイアスの和として表現でき、分散とバイアスを同時に評価するものとなっている。

$$\begin{aligned} E\{(\hat{\theta} - \theta)^2\} &= E\left\{\left[(\hat{\theta} - E(\hat{\theta})) + (E(\hat{\theta}) - \theta)\right]^2\right\} \\ &= E\left\{(\hat{\theta} - E(\hat{\theta}))^2\right\} + (E(\hat{\theta}) - \theta)^2 + 2E\left\{(\hat{\theta} - E(\hat{\theta}))(E(\hat{\theta}) - \theta)\right\} \\ &= E\left\{(\hat{\theta} - E(\hat{\theta}))^2\right\} + (E(\hat{\theta}) - \theta)^2 + 2(E(\hat{\theta}) - \theta)E(\hat{\theta} - E(\hat{\theta})) \\ &= E\left\{(\hat{\theta} - E(\hat{\theta}))^2\right\} + (E(\hat{\theta}) - \theta)^2 \\ &= V(\hat{\theta}) + \{b(\hat{\theta})\}^2 \end{aligned}$$

つまり、推定量のよさは、様々な尺度で測られるが、平均 2 乗誤差は、不偏でなくても、「真の母数の値からの差の期待値」ができるだけ小さい方がよいと考えるものである。

(補論 3) Hidiroglou-Berthelot method

標本単位 i の時期 t における回答値を x_{it} とおくと、前回回答値との比率 r_{it} は、

$$r_{it} = \frac{x_{it}}{x_{it-1}}$$

と表される。ここで、 r_{it} の分布は左右非対称であるため、分布の左端からは外れ値が検出され難いとの指摘を踏まえ、データ分布の両端から外れ値が検出されるように下式の通り変数変換を行う。ただし、 r_{it} の分布のメジアンを r_{it}^M とする。

$$s_{it} = \begin{cases} 1 - \frac{r_{it}^M}{r_{it}} & \text{if } 0 < r_{it} < r_{it}^M \\ \frac{r_{it}}{r_{it}^M} - 1 & \text{otherwise } r_{it} \geq r_{it}^M \end{cases}$$

ここで、回答値 (x_{it} 、 x_{it-1}) が小さい程、 r_{it} は変動し易い傾向にあるため、回答値の比率のみでなく、回答値の大きさも考慮するのが望ましい。このため、

$$E_{it} = s_{it} \{ \max(x_{it-1}, x_{it}) \}^U$$

と変数変換を行う。ただし、 U は回答値の大きさをどの程度反映させるか調整するための定数であり、 $0 \leq U \leq 1$ とする。

こうして得られた E_{it} の分布を基に、外れ値の検出を行う。まず、散らばりの尺度 D_L および D_U を

$$D_L = \max(E_{it}^M - E_{it}^{25}, |AE_{it}^M|)$$

$$D_U = \max(E_{it}^{75} - E_{it}^M, |AE_{it}^M|)$$

と定義する。ただし、 E_{it}^{25} 、 E_{it}^M 、 E_{it}^{75} をそれぞれ E_{it} の第1四分位数、第2四分位数(メジアン)、第3四分位数とする。また、 E_{it} が特定の値近傍に集中した場合、 $(E_{it}^M - E_{it}^{25})$ や $(E_{it}^{75} - E_{it}^M)$ が非常に小さくなってしまう。そこで、 AE_{it}^M を設定することで、 $(E_{it}^M - E_{it}^{25})$ や $(E_{it}^{75} - E_{it}^M)$ が非常に小さい時も適切に散らばりの尺度を算出できるようにしている。なお、実際には、 $A = 0.05$ と設定されることが多い。

次に、許容区間を

$$(E_{it}^M - cD_L, E_{it}^M + cD_U)$$

と設定する。ただし、 c は許容区間の幅を調整するための定数である。ここで、 E_{it} が許容区間の外にある時、外れ値として検出する。

【参考文献】

宇都宮浄人・園田桂子(2001). 「「全国企業短期経済観測調査」における欠測値補完の検討」, 日本銀行ワーキングペーパー.

永田靖(1992). 『入門統計解析法』, 日科技連.

Belcher, R., (2003). "Application of the Hidioglou-Berthelot Method of Outlier Detection for Periodic Business Surveys". *Proceedings of the Survey Methods Section*, 25-30.

Banim, J., (2000). "An Assessment of Macro Editing Methods". *UN/ECE Work Session on Statistical Data Editing*, Working Paper No. 7 .

Barnett, V., (1994). "Outlier in Sample Survey". *Journal of Applied Statistics*, 21, 373-381.

Barnett, V. and Lewis, T., (1994). *Outliers in Statistical Data*, Wiley.

Bernier, J. and Nobrega, K., (1998). "Outlier Detection in Asymmetric Samples:A Comparison of an Inter-quartile Range Method and a Variation of a Sigma Gap Method". *Proceedings of the Survey Methods Section*, 137-141.

Chambers, R.L., (1986). "Outlier Robust Finite Population Estimation". *Journal of the American Statistical Association*, 81, 1063-1069.

Chambers, R.L. and Ren, R., (2004). "Outlier Robust Imputation of Survey Data". *Proceedings of the Section on Survey Research Methods of the American Statistical Association*, 3336-3344.

Fuller, W. A., (2009). *Sampling Statistics*, Wiley.

Hidioglou, M.A., and Berthelot, J.-M., (1986). "Statistical Editing and Imputation for Periodic Business Surveys". *Survey Methodology*, 12, 73-83.

Hunt, J.W., (1999). "Detecting Outliers in the Monthly Retail Trade Survey Using the Hidioglou-Berthelot Method". *Proceedings of the Section on Survey Research Methods of the American Statistical Association*, 539-543.

Lohr, S.L., (1999). *Sampling: Design and Analysis*, Duxbury.

Pfeffermann, D. and Rao, C.R., (2009). *Handbook of Statistics 29*, North Holland.

(図表 1)

外れ値の主な検出方法

外れ値の検出方法	
許容区間を設定する方法 (range edit)	事前に設定した上限値または下限値を超えた場合、外れ値として検出。すなわち、平均やメジアン等の位置の尺度と分散や四分位範囲等の散らばりの尺度を用いて許容区間を設定し、許容区間外のデータを外れ値として検出する。
quartile method	メジアンを位置の尺度、第 1 四分位数とメジアンの距離および第 3 四分位数とメジアンの距離をそれぞれ散らばりの尺度として許容区間を設定し、事前に設定した値を超えた場合、外れ値として検出する。
比率を用いる方法	前回調査の回答値と今回調査の回答値といった 2 つのデータの比率を算出。比率が事前に設定した値を超えた場合、外れ値として検出する。
Hidiroglou- Berthelot method	「比率を用いる方法」のように比率を算出した後、変数変換を行い、データセットを設定する。散らばりの尺度等を独自に設定することで、適切な外れ値検出を企図している (詳細は補論 3)
仮説検定を用いる方法	仮説検定を適用する。パラメトリックな手法であるスミルノフ・グラブス検定等がある。

(図表 2)

外れ値の処理方法

外れ値の処理方法	
外れ値自体を修正する方法	
標本単位のウェイトを修正する方法	回答値の修正は行わず、ウェイトを調整することで、外れ値の影響度を低下させる。外れ値のウェイトを修正した分、他の標本単位のウェイトも修正が必要となる。
標本単位の回答値を修正する方法	標本単位のウェイトは修正せずに、回答値を修正する。具体的には、(a)他の標本単位の値を代入、(b)事前に設定した値 (cutoff point) を超えた場合、cutoff point を代入、(c) 欠測値と捉えて欠測値補完 (imputation) を用いる等がある (詳細は図表 3)。
標本単位のウェイトと回答値の双方を修正する方法	ウェイトによって重み付けされた標本単位のデータが cutoff point を超えた場合、cutoff point を代入するのではなく、cutoff point を超えた部分については、ウェイトを「 1 」として cutoff point に加え、その値を代入する。
外れ値の影響を緩和する推定方法	標本単位の回答値やウェイトは修正せず、推定の際に外れ値の影響を受け難い統計モデルを用いることで外れ値の影響を緩和する。

(図表 3)

主な欠測値補完方法

	欠測値補完方法
deductive imputation	同一単位において、調査項目間の論理的な根拠により補完する。例えば、負の値を取り得ない項目の回答値が「0」であるなら、その内訳項目が欠測値である時、「0」を代入するというもの。
cell mean imputation	欠測値の発生は、同一層において無作為に発生するとの仮定の下、欠測値が発生した層の平均値を代入する。このため、実際には取り得ない値が代入されることがある（整数項目に小数が代入される等）。また、同一層では全ての欠測値に同じ値を代入するため、データが一か所に集中してしまい（spike）分散の推定値が過小評価となる。
hot-deck imputation	欠測値が発生した調査回における同一層の単位の回答値を代入する。具体的には、予め、標本単位に一定の条件の下で順序付けを行い、欠測値が発生した場合、1つ前の標本単位の値を代入する方法（sequential hot-deck imputation）、欠測値と同一層の標本から無作為に単位を抽出し、その単位の回答値を代入する方法（random hot-deck imputation）、標本単位間の距離を定義し、欠測値と最も近い単位の回答値を代入する方法（nearest-neighbor hot-deck imputation）等がある。
regression imputation	hot-deck が、他の標本単位の値をそのまま用いるのに対し、regression imputation では、回帰分析により推定した値を代入する。勿論、代入値の精度は、回帰モデルの性能次第となる。
cold-deck imputation	前回調査の回答値や時系列データ等、欠測値が発生した調査回以外の情報を用いて補完する。
substitution	欠測値が発生した場合、その代替単位として当初標本として抽出されなかった集団から再抽出を行い補完する。ただし、substitution を用いると代替単位の抽出にバイアスがかかる傾向があるため、抽出の際は注意が必要となる。
multiple imputation	欠測値に対して複数の値を代入することにより複数のデータセットを作成する。各データセットを欠測値が無いデータセットとして分析を行い、そこから得られた結果を統合する方法である。

(図表 4)

海外における外れ値対応事例

統計名	作成機関	国	外れ値の対応方法
Survey of Employment, Payrolls and Hours (SEPH)	Statistics Canada	カナダ	Hidioglou-Berthelot method や quartile method 等により外れ値を検出。weight modification 等により外れ値を修正。
National Construction Industry Wage Rate Survey	Statistics Canada	カナダ	位置の尺度に「平均」、散らばりの尺度に「標準偏差」を用いた range edit により外れ値を検出し、ケースバイケースで外れ値を処理。
State and Metro Area Employment, Hours, & Earnings	Bureau of Labor Statistics	アメリカ	推定値に対する外れ値の影響を低下させるため、weight reduction により、外れ値を処理。
Consumer Price Index	Australian Bureau of Statistics	オーストラリア	winsorization により、外れ値を処理。代入値は、外れ値の次に大きい値と同値。

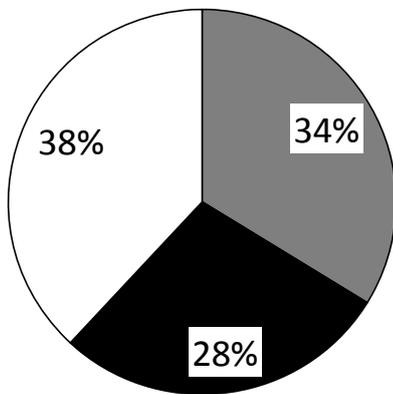
母集団推計値への影響度合い y_{it} の分布(10 調査回平均値) (1)

売上高や利益項目に比べ、設備投資額やソフトウェア投資額は、0の割合が非常に高い。特に、中小企業では、設備投資やソフトウェア投資を行っていない企業が多いため、その変化幅も0となるケースが散見される。

< 大企業・製造業 >

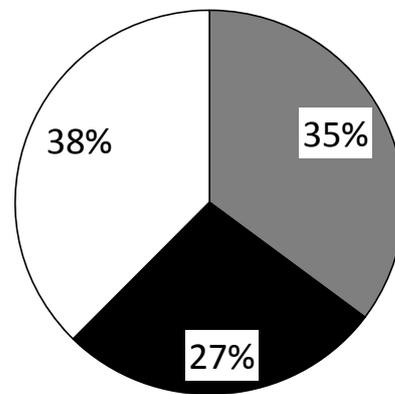
(1) 売上高

■ 正数の割合 ■ 0の割合 □ 負数の割合



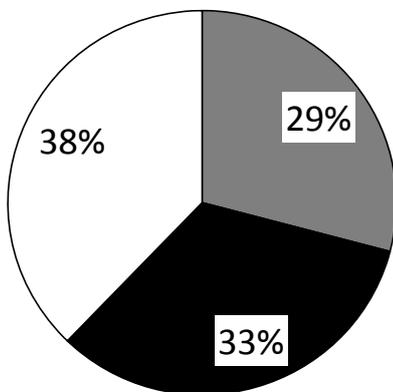
(2) 経常利益

■ 正数の割合 ■ 0の割合 □ 負数の割合



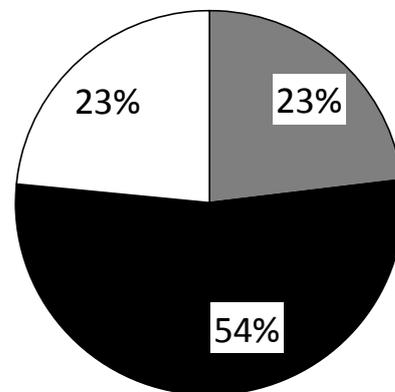
(3) 設備投資額

■ 正数の割合 ■ 0の割合 □ 負数の割合



(4) ソフトウェア投資額

■ 正数の割合 ■ 0の割合 □ 負数の割合



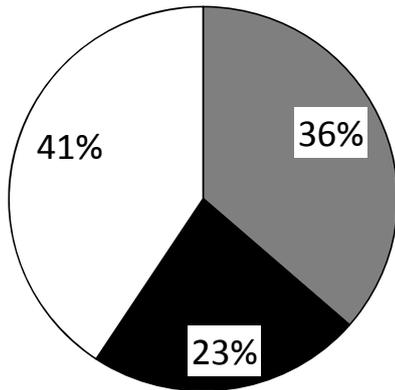
(図表 5 - 2)

母集団推計値への影響度合い y_{it} の分布(1 0 調査回平均値) (2)

< 中小企業・製造業 >

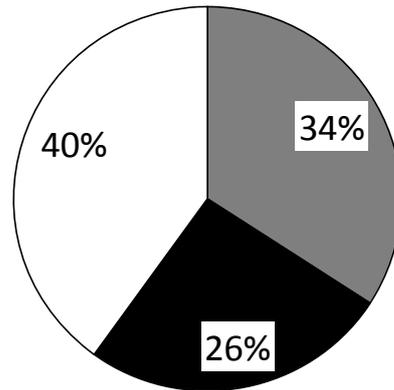
(1) 売上高

■ 正数の割合 ■ 0の割合 □ 負数の割合



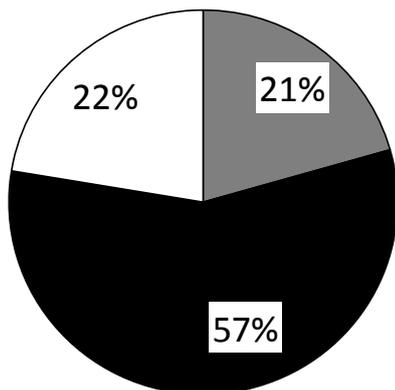
(2) 経常利益

■ 正数の割合 ■ 0の割合 □ 負数の割合



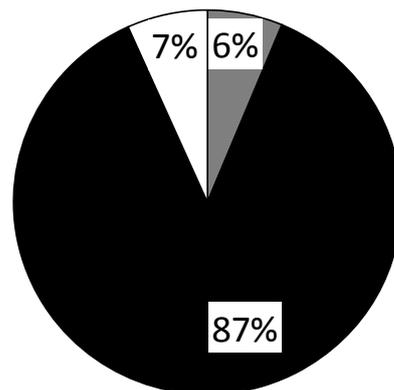
(3) 設備投資額

■ 正数の割合 ■ 0の割合 □ 負数の割合



(4) ソフトウェア投資額

■ 正数の割合 ■ 0の割合 □ 負数の割合



(図表 6 - 1)

母集団推計値への影響度合い y_{it} の分布 (1)

(1) 売上高：大企業・製造業

		データ数	最大値	最小値	平均値	正数の数、割合(%)	0の数、割合(%)	負数の数、割合(%)
2007年								
3月調査	2007年度前年比	1245	0.138	-0.138	0.001	662 (53.2)	365 (29.3)	218 (17.5)
	2006年度修正率	1245	0.238	-0.034	0.001	668 (53.7)	177 (14.2)	400 (32.1)
6月調査	2007年度修正率	1245	0.287	-0.049	0.002	609 (48.9)	318 (25.5)	318 (25.5)
	2006年度修正率	1239	0.155	-0.027	0.001	312 (25.2)	687 (55.4)	240 (19.4)
12月調査	2007年度修正率	1235	0.107	-0.048	0.001	511 (41.4)	270 (21.9)	454 (36.8)
2008年								
3月調査	2007年度修正率	1233	0.129	-0.246	0.000	318 (25.8)	597 (48.4)	318 (25.8)
	2008年度前年比	1226	0.205	-0.110	0.002	609 (49.7)	394 (32.1)	223 (18.2)
6月調査	2007年度修正率	1226	0.096	-0.054	0.000	587 (47.9)	164 (13.4)	475 (38.7)
	2008年度修正率	1226	0.220	-0.108	0.002	553 (45.1)	306 (25.0)	367 (29.9)
9月調査	2008年度修正率	1219	0.320	-0.055	0.001	223 (18.3)	679 (55.7)	317 (26.0)
12月調査	2008年度修正率	1215	0.092	-0.438	-0.003	261 (21.5)	198 (16.3)	756 (62.2)
2009年								
3月調査	2008年度修正率	1214	0.066	-0.419	-0.007	79 (6.5)	256 (21.1)	879 (72.4)
	2009年度前年比	1207	0.088	-0.405	-0.005	202 (16.7)	353 (29.2)	652 (54.0)
6月調査	2008年度修正率	1208	0.099	-0.114	-0.001	368 (30.5)	151 (12.5)	689 (57.0)
	2009年度修正率	1208	0.120	-0.653	-0.008	274 (22.7)	259 (21.4)	675 (55.9)

(2) 売上高：中堅企業・製造業

		データ数	最大値	最小値	平均値	正数の数、割合(%)	0の数、割合(%)	負数の数、割合(%)
2007年								
3月調査	2007年度前年比	1236	0.245	-0.214	0.002	745 (60.3)	98 (7.9)	393 (31.8)
	2006年度修正率	1236	0.205	-0.071	0.001	635 (51.4)	246 (19.9)	355 (28.7)
6月調査	2007年度修正率	1236	0.220	-0.080	0.001	395 (32.0)	549 (44.4)	292 (23.6)
	2006年度修正率	1228	0.073	-0.233	0.000	463 (37.7)	345 (28.1)	420 (34.2)
12月調査	2007年度修正率	1219	0.067	-0.030	0.000	573 (47.0)	138 (11.3)	508 (41.7)
2008年								
3月調査	2007年度修正率	1220	0.141	-0.514	0.000	467 (38.3)	283 (23.2)	470 (38.5)
	2008年度前年比	1201	0.157	-0.097	0.002	741 (61.7)	92 (7.7)	368 (30.6)
6月調査	2007年度修正率	1203	0.200	-0.066	0.001	601 (50.0)	239 (19.9)	363 (30.2)
	2008年度修正率	1203	0.199	-0.074	0.001	403 (33.5)	506 (42.1)	294 (24.4)
9月調査	2008年度修正率	1198	0.133	-0.088	0.000	401 (33.5)	318 (26.5)	479 (40.0)
12月調査	2008年度修正率	1193	0.029	-0.145	-0.002	381 (31.9)	116 (9.7)	696 (58.3)
2009年								
3月調査	2008年度修正率	1189	0.048	-0.181	-0.005	183 (15.4)	161 (13.5)	845 (71.1)
	2009年度前年比	1178	0.205	-0.291	-0.007	272 (23.1)	80 (6.8)	826 (70.1)
6月調査	2008年度修正率	1178	0.133	-0.111	0.000	455 (38.6)	214 (18.2)	509 (43.2)
	2009年度修正率	1178	0.078	-0.192	-0.002	268 (22.8)	418 (35.5)	492 (41.8)

(3) 売上高：中小企業・製造業

		データ数	最大値	最小値	平均値	正数の数、割合(%)	0の数、割合(%)	負数の数、割合(%)
2007年								
3月調査	2007年度前年比	2035	0.138	-0.116	0.001	1085 (53.3)	182 (8.9)	768 (37.7)
	2006年度修正率	2036	0.074	-0.065	0.001	953 (46.8)	592 (29.1)	491 (24.1)
6月調査	2007年度修正率	2036	0.117	-0.054	0.001	672 (33.0)	821 (40.3)	543 (26.7)
	2006年度修正率	2017	0.211	-0.064	0.000	752 (37.3)	549 (27.2)	716 (35.5)
12月調査	2007年度修正率	2005	0.156	-0.185	0.000	857 (42.7)	309 (15.4)	839 (41.8)
2008年								
3月調査	2007年度修正率	2007	0.075	-0.344	0.000	767 (38.2)	490 (24.4)	750 (37.4)
	2008年度前年比	1981	0.340	-0.099	0.001	997 (50.3)	215 (10.9)	769 (38.8)
6月調査	2007年度修正率	1981	0.070	-0.101	0.000	881 (44.5)	555 (28.0)	545 (27.5)
	2008年度修正率	1981	0.157	-0.104	0.000	601 (30.3)	764 (38.6)	616 (31.1)
9月調査	2008年度修正率	1959	0.126	-0.135	0.000	627 (32.0)	465 (23.7)	867 (44.3)
12月調査	2008年度修正率	1947	0.063	-0.229	-0.001	635 (32.6)	244 (12.5)	1068 (54.9)
2009年								
3月調査	2008年度修正率	1933	0.084	-0.101	-0.002	416 (21.5)	347 (18.0)	1170 (60.5)
	2009年度前年比	1911	0.115	-0.412	-0.005	421 (22.0)	170 (8.9)	1320 (69.1)
6月調査	2008年度修正率	1916	0.125	-0.148	0.000	731 (38.2)	478 (24.9)	707 (36.9)
	2009年度修正率	1916	0.133	-0.269	-0.002	445 (23.2)	624 (32.6)	847 (44.2)

(図表 6 - 2)

母集団推計値への影響度合い y_{it} の分布 (2)

(1) 売上高：大企業・非製造業

		データ数	最大値	最小値	平均値	正数の数、割合(%)	0の数、割合(%)	負数の数、割合(%)
2007年								
3月調査	2007年度前年比	1222	0.281	-0.205	0.001	582 (47.6)	382 (31.3)	258 (21.1)
	2006年度修正率	1223	0.083	-0.123	0.000	619 (50.6)	225 (18.4)	379 (31.0)
6月調査	2007年度修正率	1223	0.091	-0.252	0.001	537 (43.9)	337 (27.6)	349 (28.5)
	2007年度修正率	1217	0.055	-0.101	0.000	256 (21.0)	720 (59.2)	241 (19.8)
12月調査	2007年度修正率	1212	0.140	-0.070	0.001	443 (36.6)	313 (25.8)	456 (37.6)
2008年								
3月調査	2007年度修正率	1211	0.071	-0.110	0.000	290 (23.9)	530 (43.8)	391 (32.3)
	2008年度前年比	1199	0.386	-0.083	0.001	543 (45.3)	389 (32.4)	267 (22.3)
6月調査	2007年度修正率	1199	0.145	-0.024	0.001	528 (44.0)	210 (17.5)	461 (38.4)
	2008年度修正率	1199	0.266	-0.070	0.002	459 (38.3)	357 (29.8)	383 (31.9)
9月調査	2008年度修正率	1189	0.192	-0.102	0.001	183 (15.4)	680 (57.2)	326 (27.4)
12月調査	2008年度修正率	1186	0.241	-0.177	0.000	329 (27.7)	270 (22.8)	587 (49.5)
2009年								
3月調査	2008年度修正率	1182	0.218	-0.690	-0.003	169 (14.3)	387 (32.7)	626 (53.0)
	2009年度前年比	1171	0.103	-0.612	-0.004	308 (26.3)	395 (33.7)	468 (40.0)
6月調査	2008年度修正率	1172	0.137	-0.414	-0.002	356 (30.4)	197 (16.8)	619 (52.8)
	2009年度修正率	1172	0.287	-0.913	-0.006	305 (26.0)	320 (27.3)	547 (46.7)

(2) 売上高：中堅企業・非製造業

		データ数	最大値	最小値	平均値	正数の数、割合(%)	0の数、割合(%)	負数の数、割合(%)
2007年								
3月調査	2007年度前年比	1690	0.378	-0.052	0.002	916 (54.2)	200 (11.8)	574 (34.0)
	2006年度修正率	1692	0.484	-0.254	0.000	772 (45.6)	399 (23.6)	521 (30.8)
6月調査	2007年度修正率	1692	0.397	-0.242	0.001	539 (31.9)	718 (42.4)	435 (25.7)
	2007年度修正率	1675	0.117	-0.407	0.000	493 (29.4)	551 (32.9)	631 (37.7)
12月調査	2007年度修正率	1660	0.089	-0.085	0.000	656 (39.5)	278 (16.7)	726 (43.7)
2008年								
3月調査	2007年度修正率	1652	0.078	-0.170	0.000	551 (33.4)	437 (26.5)	664 (40.2)
	2008年度前年比	1636	0.174	-0.359	0.001	874 (53.4)	193 (11.8)	569 (34.8)
6月調査	2007年度修正率	1636	0.302	-0.069	0.000	717 (43.8)	385 (23.5)	534 (32.6)
	2008年度修正率	1636	0.312	-0.171	0.001	464 (28.4)	686 (41.9)	486 (29.7)
9月調査	2008年度修正率	1622	0.693	-0.248	0.000	426 (26.3)	538 (33.2)	658 (40.6)
12月調査	2008年度修正率	1618	0.161	-0.172	-0.001	502 (31.0)	269 (16.6)	847 (52.3)
2009年								
3月調査	2008年度修正率	1610	0.077	-0.228	-0.002	346 (21.5)	346 (21.5)	918 (57.0)
	2009年度前年比	1585	0.200	-0.682	-0.003	528 (33.3)	218 (13.8)	839 (52.9)
6月調査	2008年度修正率	1586	0.159	-0.347	0.000	568 (35.8)	371 (23.4)	647 (40.8)
	2009年度修正率	1586	0.157	-0.368	-0.001	383 (24.1)	608 (38.3)	595 (37.5)

(3) 売上高：中小企業・非製造業

		データ数	最大値	最小値	平均値	正数の数、割合(%)	0の数、割合(%)	負数の数、割合(%)
2007年								
3月調査	2007年度前年比	3457	0.241	-0.089	0.000	1568 (45.4)	543 (15.7)	1346 (38.9)
	2006年度修正率	3458	0.066	-0.234	0.000	1401 (40.5)	1128 (32.6)	929 (26.9)
6月調査	2007年度修正率	3458	0.094	-0.234	0.000	984 (28.5)	1555 (45.0)	919 (26.6)
	2007年度修正率	3418	0.099	-0.064	0.000	1022 (29.9)	1183 (34.6)	1213 (35.5)
12月調査	2007年度修正率	3368	0.104	-0.156	0.000	1297 (38.5)	688 (20.4)	1383 (41.1)
2008年								
3月調査	2007年度修正率	3406	0.142	-0.063	0.000	1194 (35.1)	974 (28.6)	1238 (36.3)
	2008年度前年比	3360	0.082	-0.118	0.000	1499 (44.6)	518 (15.4)	1343 (40.0)
6月調査	2007年度修正率	3363	0.062	-0.113	0.000	1368 (40.7)	1081 (32.1)	914 (27.2)
	2008年度修正率	3363	0.104	-0.114	0.000	906 (26.9)	1445 (43.0)	1012 (30.1)
9月調査	2008年度修正率	3318	0.205	-0.072	0.000	932 (28.1)	1057 (31.9)	1329 (40.1)
12月調査	2008年度修正率	3272	0.123	-0.224	0.000	1139 (34.8)	655 (20.0)	1478 (45.2)
2009年								
3月調査	2008年度修正率	3331	0.125	-0.247	-0.001	866 (26.0)	808 (24.3)	1657 (49.7)
	2009年度前年比	3284	0.129	-0.327	-0.002	947 (28.8)	461 (14.0)	1876 (57.1)
6月調査	2008年度修正率	3285	0.084	-0.131	0.000	1213 (36.9)	1014 (30.9)	1058 (32.2)
	2009年度修正率	3285	0.163	-0.121	0.000	832 (25.3)	1243 (37.8)	1210 (36.8)

(図表 6 - 3)

母集団推計値への影響度合い y_{it} の分布 (3)

(1) 経常利益：大企業・製造業

		データ数	最大値	最小値	平均値	正数の数、割合(%)	0の数、割合(%)	負数の数、割合(%)
2007年								
3月調査	2007年度前年比	1245	0.761	-0.761	0.000	567 (45.5)	376 (30.2)	302 (24.3)
	2006年度修正率	1245	0.563	-0.843	0.002	700 (56.2)	171 (13.7)	374 (30.0)
6月調査	2007年度修正率	1245	0.621	-0.842	0.002	531 (42.7)	252 (20.2)	462 (37.1)
	2007年度修正率	1239	0.631	-0.263	0.002	316 (25.5)	690 (55.7)	233 (18.8)
12月調査	2007年度修正率	1235	0.320	-0.381	0.001	475 (38.5)	263 (21.3)	497 (40.2)
2008年								
3月調査	2007年度修正率	1233	0.240	-0.370	-0.001	277 (22.5)	585 (47.4)	371 (30.1)
	2008年度前年比	1226	0.592	-0.381	0.000	505 (41.2)	399 (32.5)	322 (26.3)
6月調査	2007年度修正率	1226	0.744	-0.736	-0.002	507 (41.4)	171 (13.9)	548 (44.7)
	2008年度修正率	1226	0.744	-3.284	-0.009	412 (33.6)	257 (21.0)	557 (45.4)
9月調査	2008年度修正率	1219	1.030	-0.310	0.000	223 (18.3)	665 (54.6)	331 (27.2)
12月調査	2008年度修正率	1215	0.706	-2.326	-0.013	323 (26.6)	203 (16.7)	689 (56.7)
2009年								
3月調査	2008年度修正率	1214	0.552	-3.183	-0.042	137 (11.3)	247 (20.3)	830 (68.4)
	2009年度前年比	1207	2.352	-1.710	-0.016	412 (34.1)	371 (30.7)	424 (35.1)
6月調査	2008年度修正率	1208	3.129	-1.929	0.002	600 (49.7)	150 (12.4)	458 (37.9)
	2009年度修正率	1208	3.555	-9.704	-0.019	495 (41.0)	211 (17.5)	502 (41.6)

(2) 経常利益：中堅企業・製造業

		データ数	最大値	最小値	平均値	正数の数、割合(%)	0の数、割合(%)	負数の数、割合(%)
2007年								
3月調査	2007年度前年比	1234	0.752	-0.744	0.001	653 (52.9)	138 (11.2)	443 (35.9)
	2006年度修正率	1235	0.712	-0.485	0.003	595 (48.2)	258 (20.9)	382 (30.9)
6月調査	2007年度修正率	1235	0.265	-0.350	0.000	402 (32.6)	437 (35.4)	396 (32.1)
	2007年度修正率	1227	0.705	-0.411	0.001	411 (33.5)	349 (28.4)	467 (38.1)
12月調査	2007年度修正率	1218	0.397	-0.156	-0.001	515 (42.3)	146 (12.0)	557 (45.7)
2008年								
3月調査	2007年度修正率	1219	0.277	-0.482	-0.002	416 (34.1)	271 (22.2)	532 (43.6)
	2008年度前年比	1199	0.476	-0.823	0.003	642 (53.5)	130 (10.8)	427 (35.6)
6月調査	2007年度修正率	1201	0.359	-1.452	-0.002	525 (43.7)	239 (19.9)	437 (36.4)
	2008年度修正率	1201	0.397	-0.747	-0.002	387 (32.2)	393 (32.7)	421 (35.1)
9月調査	2008年度修正率	1196	0.434	-0.570	-0.003	362 (30.3)	323 (27.0)	511 (42.7)
12月調査	2008年度修正率	1191	0.479	-0.810	-0.011	380 (31.9)	119 (10.0)	692 (58.1)
2009年								
3月調査	2008年度修正率	1187	0.498	-1.731	-0.030	227 (19.1)	183 (15.4)	777 (65.5)
	2009年度前年比	1176	0.982	-2.107	-0.017	514 (43.7)	116 (9.9)	546 (46.4)
6月調査	2008年度修正率	1176	0.495	-1.449	-0.001	524 (44.6)	214 (18.2)	438 (37.2)
	2009年度修正率	1176	1.066	-1.181	-0.002	422 (35.9)	342 (29.1)	412 (35.0)

(3) 経常利益：中小企業・製造業

		データ数	最大値	最小値	平均値	正数の数、割合(%)	0の数、割合(%)	負数の数、割合(%)
2007年								
3月調査	2007年度前年比	2030	0.708	-0.516	0.004	976 (48.1)	393 (19.4)	661 (32.6)
	2006年度修正率	2031	0.230	-0.279	0.002	817 (40.2)	610 (30.0)	604 (29.7)
6月調査	2007年度修正率	2031	0.247	-0.421	-0.001	620 (30.5)	699 (34.4)	712 (35.1)
	2007年度修正率	2014	0.383	-0.341	-0.002	595 (29.5)	621 (30.8)	798 (39.6)
12月調査	2007年度修正率	2002	0.260	-1.514	-0.002	760 (38.0)	377 (18.8)	865 (43.2)
2008年								
3月調査	2007年度修正率	2004	0.373	-0.669	-0.003	611 (30.5)	600 (29.9)	793 (39.6)
	2008年度前年比	1977	7.769	-0.437	0.008	936 (47.3)	394 (19.9)	647 (32.7)
6月調査	2007年度修正率	1977	0.304	-0.466	-0.001	717 (36.3)	600 (30.3)	660 (33.4)
	2008年度修正率	1977	0.412	-0.720	-0.004	538 (27.2)	642 (32.5)	797 (40.3)
9月調査	2008年度修正率	1957	0.580	-2.997	-0.005	525 (26.8)	564 (28.8)	868 (44.4)
12月調査	2008年度修正率	1945	0.927	-0.828	-0.008	620 (31.9)	335 (17.2)	990 (50.9)
2009年								
3月調査	2008年度修正率	1931	0.436	-1.852	-0.018	406 (21.0)	426 (22.1)	1099 (56.9)
	2009年度前年比	1909	3.594	-4.152	-0.017	697 (36.5)	311 (16.3)	901 (47.2)
6月調査	2008年度修正率	1913	1.088	-0.904	-0.001	705 (36.9)	545 (28.5)	663 (34.7)
	2009年度修正率	1913	2.081	-13.56	-0.019	605 (31.6)	528 (27.6)	780 (40.8)

(図表 6 - 4)

母集団推計値への影響度合い y_{it} の分布 (4)

(1) 経常利益：大企業・非製造業

		データ数	最大値	最小値	平均値	正数の数、割合(%)	0の数、割合(%)	負数の数、割合(%)
2007年								
3月調査	2007年度前年比	1221	0.226	-0.807	-0.001	474 (38.8)	412 (33.7)	335 (27.4)
	2006年度修正率	1222	0.398	-0.180	0.003	687 (56.2)	200 (16.4)	335 (27.4)
6月調査	2007年度修正率	1222	0.451	-0.250	0.003	540 (44.2)	297 (24.3)	385 (31.5)
	2007年度修正率	1216	0.433	-1.633	0.000	268 (22.0)	711 (58.5)	237 (19.5)
12月調査	2007年度修正率	1211	0.775	-0.662	0.001	475 (39.2)	289 (23.9)	447 (36.9)
2008年								
3月調査	2007年度修正率	1210	0.258	-0.683	-0.003	295 (24.4)	519 (42.9)	396 (32.7)
	2008年度前年比	1198	0.160	-0.539	0.001	474 (39.6)	394 (32.9)	330 (27.5)
6月調査	2007年度修正率	1198	0.990	-0.131	0.003	587 (49.0)	195 (16.3)	416 (34.7)
	2008年度修正率	1198	0.983	-1.265	-0.001	469 (39.1)	290 (24.2)	439 (36.6)
9月調査	2008年度修正率	1188	0.989	-1.232	-0.004	187 (15.7)	662 (55.7)	339 (28.5)
12月調査	2008年度修正率	1185	0.647	-1.223	-0.008	386 (32.6)	263 (22.2)	536 (45.2)
2009年								
3月調査	2008年度修正率	1181	1.991	-0.704	-0.007	239 (20.2)	375 (31.8)	567 (48.0)
	2009年度前年比	1172	1.239	-2.429	-0.005	384 (32.8)	405 (34.6)	383 (32.7)
6月調査	2008年度修正率	1173	0.649	-1.465	-0.005	584 (49.8)	178 (15.2)	411 (35.0)
	2009年度修正率	1173	1.979	-2.532	-0.007	416 (35.5)	269 (22.9)	488 (41.6)

(2) 経常利益：中堅企業・非製造業

		データ数	最大値	最小値	平均値	正数の数、割合(%)	0の数、割合(%)	負数の数、割合(%)
2007年								
3月調査	2007年度前年比	1689	0.472	-0.582	0.005	878 (52.0)	254 (15.0)	557 (33.0)
	2006年度修正率	1691	0.660	-0.298	0.003	810 (47.9)	365 (21.6)	516 (30.5)
6月調査	2007年度修正率	1691	0.448	-0.414	0.000	572 (33.8)	568 (33.6)	551 (32.6)
	2007年度修正率	1674	0.238	-0.290	-0.001	512 (30.6)	562 (33.6)	600 (35.8)
12月調査	2007年度修正率	1659	1.400	-0.523	0.000	673 (40.6)	253 (15.3)	733 (44.2)
2008年								
3月調査	2007年度修正率	1651	0.403	-0.592	-0.001	534 (32.3)	430 (26.0)	687 (41.6)
	2008年度前年比	1635	1.070	-0.676	0.005	835 (51.1)	222 (13.6)	578 (35.4)
6月調査	2007年度修正率	1635	0.175	-0.794	0.001	752 (46.0)	329 (20.1)	554 (33.9)
	2008年度修正率	1635	0.337	-1.049	-0.002	493 (30.2)	553 (33.8)	589 (36.0)
9月調査	2008年度修正率	1621	0.684	-1.930	-0.005	410 (25.3)	527 (32.5)	684 (42.2)
12月調査	2008年度修正率	1617	0.858	-1.687	-0.005	576 (35.6)	260 (16.1)	781 (48.3)
2009年								
3月調査	2008年度修正率	1609	0.377	-0.455	-0.011	425 (26.4)	345 (21.4)	839 (52.1)
	2009年度前年比	1585	1.081	-0.930	0.000	692 (43.7)	260 (16.4)	633 (39.9)
6月調査	2008年度修正率	1586	1.900	-3.085	-0.001	742 (46.8)	339 (21.4)	505 (31.8)
	2009年度修正率	1586	3.971	-0.480	0.002	529 (33.4)	516 (32.5)	541 (34.1)

(3) 経常利益：中小企業・非製造業

		データ数	最大値	最小値	平均値	正数の数、割合(%)	0の数、割合(%)	負数の数、割合(%)
2007年								
3月調査	2007年度前年比	3448	1.982	-0.687	0.002	1494 (43.3)	935 (27.1)	1019 (29.6)
	2006年度修正率	3449	0.821	-0.410	0.001	1236 (35.8)	1231 (35.7)	982 (28.5)
6月調査	2007年度修正率	3449	0.769	-1.117	-0.001	964 (28.0)	1423 (41.3)	1062 (30.8)
	2007年度修正率	3411	1.101	-0.363	0.000	908 (26.6)	1316 (38.6)	1187 (34.8)
12月調査	2007年度修正率	3362	0.368	-0.343	0.000	1147 (34.1)	877 (26.1)	1338 (39.8)
2008年								
3月調査	2007年度修正率	3399	0.388	-0.458	-0.001	969 (28.5)	1139 (33.5)	1291 (38.0)
	2008年度前年比	3351	0.723	-1.114	0.001	1438 (42.9)	912 (27.2)	1001 (29.9)
6月調査	2007年度修正率	3355	0.357	-0.536	0.001	1154 (34.4)	1144 (34.1)	1057 (31.5)
	2008年度修正率	3355	0.377	-0.801	-0.001	898 (26.8)	1300 (38.7)	1157 (34.5)
9月調査	2008年度修正率	3308	0.830	-0.300	-0.001	801 (24.2)	1237 (37.4)	1270 (38.4)
12月調査	2008年度修正率	3265	0.454	-0.869	-0.003	1009 (30.9)	780 (23.9)	1476 (45.2)
2009年								
3月調査	2008年度修正率	3324	0.772	-0.953	-0.005	834 (25.1)	981 (29.5)	1509 (45.4)
	2009年度前年比	3276	1.082	-0.804	0.002	1257 (38.4)	795 (24.3)	1224 (37.4)
6月調査	2008年度修正率	3277	0.792	-0.430	0.001	1191 (36.3)	1109 (33.8)	977 (29.8)
	2009年度修正率	3277	0.666	-0.640	-0.001	999 (30.5)	1193 (36.4)	1085 (33.1)

(図表 6 - 5)

母集団推計値への影響度合い y_{it} の分布 (5)

(1) 当期純利益 : 大企業・製造業

		データ数	最大値	最小値	平均値	正数の数、割合(%)	0の数、割合(%)	負数の数、割合(%)
2007年								
3月調査	2007年度前年比	1245	3.060	-2.278	0.002	528 (42.4)	361 (29.0)	356 (28.6)
	2006年度修正率	1245	0.567	-1.471	-0.002	625 (50.2)	168 (13.5)	452 (36.3)
6月調査	2007年度修正率	1245	0.904	-1.434	0.004	554 (44.5)	234 (18.8)	457 (36.7)
	2007年度修正率	1239	0.591	-0.694	0.002	311 (25.1)	685 (55.3)	243 (19.6)
12月調査	2007年度修正率	1235	1.652	-0.244	0.002	476 (38.5)	258 (20.9)	501 (40.6)
2008年								
3月調査	2007年度修正率	1233	0.359	-0.955	-0.003	264 (21.4)	588 (47.7)	381 (30.9)
	2008年度前年比	1226	0.992	-1.202	0.002	492 (40.1)	387 (31.6)	347 (28.3)
6月調査	2007年度修正率	1226	1.203	-0.953	-0.005	439 (35.8)	173 (14.1)	614 (50.1)
	2008年度修正率	1226	1.176	-2.985	-0.010	414 (33.8)	259 (21.1)	553 (45.1)
9月調査	2008年度修正率	1219	0.947	-0.306	0.000	218 (17.9)	669 (54.9)	332 (27.2)
12月調査	2008年度修正率	1215	0.653	-2.452	-0.016	300 (24.7)	199 (16.4)	716 (58.9)
2009年								
3月調査	2008年度修正率	1214	0.519	-6.155	-0.082	110 (9.1)	247 (20.3)	857 (70.6)
	2009年度前年比	1207	2.291	-615.2	6.506	460 (38.1)	366 (30.3)	381 (31.6)
6月調査	2008年度修正率	1208	499.0	-776.0	-11.30	439 (36.3)	147 (12.2)	622 (51.5)
	2009年度修正率	1208	13.05	-11.62	-0.017	521 (43.1)	196 (16.2)	491 (40.6)

(2) 当期純利益 : 中堅企業・製造業

		データ数	最大値	最小値	平均値	正数の数、割合(%)	0の数、割合(%)	負数の数、割合(%)
2007年								
3月調査	2007年度前年比	1226	1.620	-1.149	0.004	643 (52.4)	139 (11.3)	444 (36.2)
	2006年度修正率	1228	2.302	-0.470	0.002	532 (43.3)	254 (20.7)	442 (36.0)
6月調査	2007年度修正率	1228	0.450	-0.705	0.000	404 (32.9)	423 (34.4)	401 (32.7)
	2007年度修正率	1222	1.233	-0.426	0.001	413 (33.8)	349 (28.6)	460 (37.6)
12月調査	2007年度修正率	1213	0.750	-0.337	-0.001	526 (43.4)	150 (12.4)	537 (44.3)
2008年								
3月調査	2007年度修正率	1215	1.221	-0.732	-0.004	399 (32.8)	289 (23.8)	527 (43.4)
	2008年度前年比	1195	1.218	-1.030	0.009	643 (53.8)	132 (11.0)	420 (35.1)
6月調査	2007年度修正率	1198	0.598	-2.271	-0.006	476 (39.7)	236 (19.7)	486 (40.6)
	2008年度修正率	1198	0.472	-0.862	-0.003	374 (31.2)	381 (31.8)	443 (37.0)
9月調査	2008年度修正率	1193	0.480	-0.653	-0.005	350 (29.3)	332 (27.8)	511 (42.8)
12月調査	2008年度修正率	1188	0.454	-0.931	-0.016	375 (31.6)	129 (10.9)	684 (57.6)
2009年								
3月調査	2008年度修正率	1185	0.781	-2.591	-0.053	222 (18.7)	197 (16.6)	766 (64.6)
	2009年度前年比	1174	11.88	-8.219	0.047	550 (46.8)	123 (10.5)	501 (42.7)
6月調査	2008年度修正率	1174	5.490	-26.73	-0.102	445 (37.9)	216 (18.4)	513 (43.7)
	2009年度修正率	1174	3.612	-2.774	-0.002	432 (36.8)	326 (27.8)	416 (35.4)

(3) 当期純利益 : 中小企業・製造業

		データ数	最大値	最小値	平均値	正数の数、割合(%)	0の数、割合(%)	負数の数、割合(%)
2007年								
3月調査	2007年度前年比	2026	0.995	-2.595	0.007	973 (48.0)	412 (20.3)	641 (31.6)
	2006年度修正率	2027	0.733	-4.533	-0.002	747 (36.9)	658 (32.5)	622 (30.7)
6月調査	2007年度修正率	2027	0.731	-0.387	0.000	616 (30.4)	739 (36.5)	672 (33.2)
	2007年度修正率	2010	0.366	-0.612	-0.002	577 (28.7)	667 (33.2)	766 (38.1)
12月調査	2007年度修正率	1998	1.310	-1.413	-0.003	704 (35.2)	423 (21.2)	871 (43.6)
2008年								
3月調査	2007年度修正率	2000	0.431	-1.006	-0.006	583 (29.2)	611 (30.6)	806 (40.3)
	2008年度前年比	1972	7.097	-0.921	0.013	940 (47.7)	419 (21.2)	613 (31.1)
6月調査	2007年度修正率	1972	0.548	-0.670	-0.003	650 (33.0)	633 (32.1)	689 (34.9)
	2008年度修正率	1972	0.619	-1.172	-0.005	530 (26.9)	674 (34.2)	768 (38.9)
9月調査	2008年度修正率	1954	0.980	-1.823	-0.006	518 (26.5)	575 (29.4)	861 (44.1)
12月調査	2008年度修正率	1942	0.446	-0.981	-0.013	601 (30.9)	354 (18.2)	987 (50.8)
2009年								
3月調査	2008年度修正率	1928	0.647	-3.322	-0.033	385 (20.0)	440 (22.8)	1103 (57.2)
	2009年度前年比	1905	14.87	-14.93	-0.013	748 (39.3)	320 (16.8)	837 (43.9)
6月調査	2008年度修正率	1909	10.35	-7.823	-0.012	671 (35.1)	564 (29.5)	674 (35.3)
	2009年度修正率	1909	5.531	-32.67	-0.060	593 (31.1)	569 (29.8)	747 (39.1)

(図表 6 - 6)

母集団推計値への影響度合い y_{it} の分布 (6)

(1) 当期純利益 : 大企業・非製造業

		データ数	最大値	最小値	平均値	正数の数、割合(%)	0の数、割合(%)	負数の数、割合(%)
2007年								
3月調査	2007年度前年比	1220	2.666	-1.386	0.003	480 (39.3)	396 (32.5)	344 (28.2)
	2006年度修正率	1221	0.629	-0.276	0.001	573 (46.9)	200 (16.4)	448 (36.7)
6月調査	2007年度修正率	1221	0.738	-0.329	0.004	504 (41.3)	296 (24.2)	421 (34.5)
	2007年度修正率	1216	0.521	-2.395	-0.001	263 (21.6)	697 (57.3)	256 (21.1)
12月調査	2007年度修正率	1211	0.931	-1.574	-0.002	478 (39.5)	270 (22.3)	463 (38.2)
2008年								
3月調査	2007年度修正率	1210	0.911	-0.653	-0.005	274 (22.6)	519 (42.9)	417 (34.5)
	2008年度前年比	1198	0.690	-1.087	0.004	505 (42.2)	381 (31.8)	312 (26.0)
6月調査	2007年度修正率	1198	1.476	-0.917	-0.003	500 (41.7)	190 (15.9)	508 (42.4)
	2008年度修正率	1198	1.413	-1.053	0.000	461 (38.5)	300 (25.0)	437 (36.5)
9月調査	2008年度修正率	1188	0.883	-1.301	-0.005	187 (15.7)	673 (56.6)	328 (27.6)
12月調査	2008年度修正率	1185	0.657	-1.272	-0.013	363 (30.6)	264 (22.3)	558 (47.1)
2009年								
3月調査	2008年度修正率	1181	22.15	-21.57	-0.018	213 (18.0)	369 (31.2)	599 (50.7)
	2009年度前年比	1172	26.84	-26.44	0.003	439 (37.5)	388 (33.1)	345 (29.4)
6月調査	2008年度修正率	1173	8.357	-2.488	-0.023	425 (36.2)	181 (15.4)	567 (48.3)
	2009年度修正率	1173	2.405	-4.349	-0.005	442 (37.7)	265 (22.6)	466 (39.7)

(2) 当期純利益 : 中堅企業・非製造業

		データ数	最大値	最小値	平均値	正数の数、割合(%)	0の数、割合(%)	負数の数、割合(%)
2007年								
3月調査	2007年度前年比	1684	2.183	-2.266	0.009	857 (50.9)	282 (16.7)	545 (32.4)
	2006年度修正率	1686	1.321	-3.085	-0.005	714 (42.3)	369 (21.9)	603 (35.8)
6月調査	2007年度修正率	1686	0.792	-0.596	0.002	556 (33.0)	575 (34.1)	555 (32.9)
	2007年度修正率	1671	0.368	-0.825	-0.002	511 (30.6)	588 (35.2)	572 (34.2)
12月調査	2007年度修正率	1656	1.647	-0.920	0.000	646 (39.0)	288 (17.4)	722 (43.6)
2008年								
3月調査	2007年度修正率	1649	1.164	-0.768	-0.003	508 (30.8)	463 (28.1)	678 (41.1)
	2008年度前年比	1632	3.885	-1.768	0.011	839 (51.4)	240 (14.7)	553 (33.9)
6月調査	2007年度修正率	1633	0.427	-2.490	-0.007	680 (41.6)	345 (21.1)	608 (37.2)
	2008年度修正率	1633	0.758	-0.972	-0.002	473 (29.0)	571 (35.0)	589 (36.1)
9月調査	2008年度修正率	1620	2.211	-3.590	-0.006	400 (24.7)	540 (33.3)	680 (42.0)
12月調査	2008年度修正率	1616	2.226	-2.351	-0.010	562 (34.8)	277 (17.1)	777 (48.1)
2009年								
3月調査	2008年度修正率	1608	4.086	-2.759	-0.019	405 (25.2)	356 (22.1)	847 (52.7)
	2009年度前年比	1584	5.590	-5.509	0.013	720 (45.5)	269 (17.0)	595 (37.6)
6月調査	2008年度修正率	1585	3.460	-9.088	-0.021	639 (40.3)	333 (21.0)	613 (38.7)
	2009年度修正率	1585	7.893	-2.200	0.003	517 (32.6)	533 (33.6)	535 (33.8)

(3) 当期純利益 : 中小企業・非製造業

		データ数	最大値	最小値	平均値	正数の数、割合(%)	0の数、割合(%)	負数の数、割合(%)
2007年								
3月調査	2007年度前年比	3439	3.869	-1.272	0.005	1478 (43.0)	990 (28.8)	971 (28.2)
	2006年度修正率	3440	0.652	-0.442	-0.001	1073 (31.2)	1296 (37.7)	1071 (31.1)
6月調査	2007年度修正率	3440	1.434	-1.785	-0.001	934 (27.2)	1469 (42.7)	1037 (30.1)
	2007年度修正率	3403	1.841	-3.604	-0.002	870 (25.6)	1378 (40.5)	1155 (33.9)
12月調査	2007年度修正率	3353	2.085	-1.028	-0.001	1091 (32.5)	939 (28.0)	1323 (39.5)
2008年								
3月調査	2007年度修正率	3389	0.851	-0.841	-0.002	899 (26.5)	1215 (35.9)	1275 (37.6)
	2008年度前年比	3341	2.209	-1.547	0.006	1436 (43.0)	957 (28.6)	948 (28.4)
6月調査	2007年度修正率	3344	0.752	-0.866	-0.001	1065 (31.8)	1195 (35.7)	1084 (32.4)
	2008年度修正率	3344	0.957	-1.330	-0.002	869 (26.0)	1340 (40.1)	1135 (33.9)
9月調査	2008年度修正率	3299	1.000	-0.606	-0.002	790 (23.9)	1278 (38.7)	1231 (37.3)
12月調査	2008年度修正率	3257	0.863	-1.879	-0.007	950 (29.2)	856 (26.3)	1451 (44.6)
2009年								
3月調査	2008年度修正率	3316	1.609	-2.197	-0.009	790 (23.8)	1051 (31.7)	1475 (44.5)
	2009年度前年比	3268	8.076	-1.607	0.016	1285 (39.3)	839 (25.7)	1144 (35.0)
6月調査	2008年度修正率	3270	4.210	-4.453	-0.001	1082 (33.1)	1135 (34.7)	1053 (32.2)
	2009年度修正率	3270	0.782	-2.994	-0.002	994 (30.4)	1218 (37.2)	1058 (32.4)

(図表 6 - 7)

母集団推計値への影響度合い y_{it} の分布 (7)

(1) 設備投資額：大企業・製造業

		データ数	最大値	最小値	平均値	正数の数、割合(%)	0の数、割合(%)	負数の数、割合(%)
2007年								
3月調査	2007年度前年比	1244	0.605	-0.566	0.002	474 (38.1)	363 (29.2)	407 (32.7)
	2006年度修正率	1245	0.378	-0.327	-0.003	400 (32.1)	225 (18.1)	620 (49.8)
6月調査	2007年度修正率	1245	0.464	-0.728	0.004	482 (38.7)	385 (30.9)	378 (30.4)
	2007年度修正率	1239	0.460	-0.151	0.001	199 (16.1)	806 (65.1)	234 (18.9)
12月調査	2007年度修正率	1235	0.458	-0.185	-0.001	422 (34.2)	275 (22.3)	538 (43.6)
2008年								
3月調査	2007年度修正率	1233	0.410	-0.176	-0.001	206 (16.7)	699 (56.7)	328 (26.6)
	2008年度前年比	1223	0.341	-1.173	-0.003	428 (35.0)	378 (30.9)	417 (34.1)
6月調査	2007年度修正率	1225	0.321	-0.699	-0.004	365 (29.8)	225 (18.4)	635 (51.8)
	2008年度修正率	1225	0.429	-0.460	0.004	474 (38.7)	386 (31.5)	365 (29.8)
9月調査	2008年度修正率	1218	0.381	-0.379	-0.001	175 (14.4)	769 (63.1)	274 (22.5)
12月調査	2008年度修正率	1214	0.183	-0.321	-0.003	352 (29.0)	272 (22.4)	590 (48.6)
2009年								
3月調査	2008年度修正率							
	2009年度前年比	1184	0.298	-1.115	-0.011	244 (20.6)	360 (30.4)	580 (49.0)
6月調査	2008年度修正率	1201	0.320	-0.380	-0.003	443 (36.9)	193 (16.1)	565 (47.0)
	2009年度修正率	1201	0.536	-2.130	-0.014	325 (27.1)	347 (28.9)	529 (44.0)

(2) 設備投資額：中堅企業・製造業

		データ数	最大値	最小値	平均値	正数の数、割合(%)	0の数、割合(%)	負数の数、割合(%)
2007年								
3月調査	2007年度前年比	1234	3.721	-0.831	0.001	497 (40.3)	232 (18.8)	505 (40.9)
	2006年度修正率	1235	1.192	-0.827	0.001	391 (31.7)	446 (36.1)	398 (32.2)
6月調査	2007年度修正率	1235	0.327	-0.697	0.003	288 (23.3)	729 (59.0)	218 (17.7)
	2007年度修正率	1227	0.263	-1.047	-0.001	291 (23.7)	634 (51.7)	302 (24.6)
12月調査	2007年度修正率	1218	0.410	-0.389	0.000	418 (34.3)	387 (31.8)	413 (33.9)
2008年								
3月調査	2007年度修正率	1219	0.362	-0.479	-0.001	293 (24.0)	577 (47.3)	349 (28.6)
	2008年度前年比	1200	0.752	-3.437	-0.003	457 (38.1)	232 (19.3)	511 (42.6)
6月調査	2007年度修正率	1202	0.329	-0.584	-0.001	361 (30.0)	444 (36.9)	397 (33.0)
	2008年度修正率	1202	0.612	-0.391	0.003	275 (22.9)	699 (58.2)	228 (19.0)
9月調査	2008年度修正率	1197	0.491	-0.569	0.000	263 (22.0)	622 (52.0)	312 (26.1)
12月調査	2008年度修正率	1192	0.604	-0.714	-0.003	349 (29.3)	366 (30.7)	477 (40.0)
2009年								
3月調査	2008年度修正率							
	2009年度前年比	1167	0.666	-1.003	-0.027	270 (23.1)	205 (17.6)	692 (59.3)
6月調査	2008年度修正率	1172	1.331	-0.363	0.002	382 (32.6)	442 (37.7)	348 (29.7)
	2009年度修正率	1172	1.108	-0.473	0.002	254 (21.7)	650 (55.5)	268 (22.9)

(3) 設備投資額：中小企業・製造業

		データ数	最大値	最小値	平均値	正数の数、割合(%)	0の数、割合(%)	負数の数、割合(%)
2007年								
3月調査	2007年度前年比	2030	0.758	-2.050	-0.008	519 (25.6)	740 (36.5)	771 (38.0)
	2006年度修正率	2031	1.149	-1.811	0.006	430 (21.2)	1261 (62.1)	340 (16.7)
6月調査	2007年度修正率	2031	1.015	-0.700	0.004	398 (19.6)	1372 (67.6)	261 (12.9)
	2007年度修正率	2014	1.760	-0.490	0.003	423 (21.0)	1261 (62.6)	330 (16.4)
12月調査	2007年度修正率	2002	0.760	-0.422	0.004	532 (26.6)	1067 (53.3)	403 (20.1)
2008年								
3月調査	2007年度修正率	2004	0.995	-1.984	0.001	421 (21.0)	1195 (59.6)	388 (19.4)
	2008年度前年比	1977	1.725	-2.779	-0.011	452 (22.9)	762 (38.5)	763 (38.6)
6月調査	2007年度修正率	1977	0.698	-0.703	0.001	383 (19.4)	1235 (62.5)	359 (18.2)
	2008年度修正率	1977	0.945	-0.887	0.004	398 (20.1)	1304 (66.0)	275 (13.9)
9月調査	2008年度修正率	1957	9.080	-0.546	0.006	379 (19.4)	1212 (61.9)	366 (18.7)
12月調査	2008年度修正率	1945	1.188	-0.810	0.000	432 (22.2)	1040 (53.5)	473 (24.3)
2009年								
3月調査	2008年度修正率							
	2009年度前年比	1885	1.656	-2.144	-0.022	266 (14.1)	744 (39.5)	875 (46.4)
6月調査	2008年度修正率	1896	0.412	-0.410	0.001	358 (18.9)	1224 (64.6)	314 (16.6)
	2009年度修正率	1896	1.134	-0.711	0.002	318 (16.8)	1320 (69.6)	258 (13.6)

(注) 設備投資額・ソフトウェア投資額は、リース会計基準の変更に伴い調査項目の定義を変更したため、09年3月調査の08年度修正率は算出していません。

(図表 6 - 8)

母集団推計値への影響度合い y_{it} の分布 (8)

(1) 設備投資額：大企業・非製造業

		データ数	最大値	最小値	平均値	正数の数、割合(%)	0の数、割合(%)	負数の数、割合(%)
2007年								
3月調査	2007年度前年比	1216	2.293	-1.830	0.003	329 (27.1)	472 (38.8)	415 (34.1)
	2006年度修正率	1218	0.569	-0.290	0.000	415 (34.1)	370 (30.4)	433 (35.6)
6月調査	2007年度修正率	1218	1.000	-1.722	0.002	403 (33.1)	524 (43.0)	291 (23.9)
	2007年度修正率	1212	0.566	-0.318	0.001	207 (17.1)	822 (67.8)	183 (15.1)
12月調査	2007年度修正率	1208	1.459	-0.370	0.003	416 (34.4)	383 (31.7)	409 (33.9)
2008年								
3月調査	2007年度修正率	1207	0.341	-1.644	-0.004	252 (20.9)	689 (57.1)	266 (22.0)
	2008年度前年比	1194	0.678	-0.569	-0.001	331 (27.7)	448 (37.5)	415 (34.8)
6月調査	2007年度修正率	1195	0.702	-1.060	0.000	424 (35.5)	352 (29.5)	419 (35.1)
	2008年度修正率	1195	1.325	-0.907	0.000	383 (32.1)	514 (43.0)	298 (24.9)
9月調査	2008年度修正率	1185	0.466	-0.743	0.000	177 (14.9)	805 (67.9)	203 (17.1)
12月調査	2008年度修正率	1182	0.933	-1.176	-0.001	389 (32.9)	377 (31.9)	416 (35.2)
2009年								
3月調査	2008年度修正率							
	2009年度前年比	1139	0.822	-0.768	-0.002	244 (21.4)	438 (38.5)	457 (40.1)
6月調査	2008年度修正率	1157	1.480	-1.521	-0.003	419 (36.2)	356 (30.8)	382 (33.0)
	2009年度修正率	1157	0.430	-0.604	-0.001	355 (30.7)	461 (39.8)	341 (29.5)

(2) 設備投資額：中堅企業・非製造業

		データ数	最大値	最小値	平均値	正数の数、割合(%)	0の数、割合(%)	負数の数、割合(%)
2007年								
3月調査	2007年度前年比	1687	6.880	-1.751	0.002	443 (26.3)	645 (38.2)	599 (35.5)
	2006年度修正率	1689	0.496	-1.342	0.001	468 (27.7)	895 (53.0)	326 (19.3)
6月調査	2007年度修正率	1689	0.696	-1.464	0.001	350 (20.7)	1116 (66.1)	223 (13.2)
	2007年度修正率	1672	1.319	-0.941	0.001	305 (18.2)	1099 (65.7)	268 (16.0)
12月調査	2007年度修正率	1657	0.271	-1.019	-0.001	436 (26.3)	824 (49.7)	397 (24.0)
2008年								
3月調査	2007年度修正率	1649	0.390	-3.374	-0.003	308 (18.7)	985 (59.7)	356 (21.6)
	2008年度前年比	1633	2.790	-1.332	-0.001	415 (25.4)	610 (37.4)	608 (37.2)
6月調査	2007年度修正率	1633	0.747	-0.359	0.001	433 (26.5)	885 (54.2)	315 (19.3)
	2008年度修正率	1633	1.685	-0.305	0.003	297 (18.2)	1107 (67.8)	229 (14.0)
9月調査	2008年度修正率	1620	1.740	-4.128	-0.003	265 (16.4)	1066 (65.8)	289 (17.8)
12月調査	2008年度修正率	1616	1.592	-2.770	-0.001	405 (25.1)	811 (50.2)	400 (24.8)
2009年								
3月調査	2008年度修正率							
	2009年度前年比	1550	0.823	-8.159	-0.021	296 (19.1)	581 (37.5)	673 (43.4)
6月調査	2008年度修正率	1562	4.990	-0.764	0.005	449 (28.7)	827 (52.9)	286 (18.3)
	2009年度修正率	1562	2.503	-0.982	0.003	286 (18.3)	1020 (65.3)	256 (16.4)

(3) 設備投資額：中小企業・非製造業

		データ数	最大値	最小値	平均値	正数の数、割合(%)	0の数、割合(%)	負数の数、割合(%)
2007年								
3月調査	2007年度前年比	3447	5.319	-4.386	-0.005	433 (12.6)	2116 (61.4)	898 (26.1)
	2006年度修正率	3447	1.773	-0.515	0.001	441 (12.8)	2670 (77.5)	336 (9.7)
6月調査	2007年度修正率	3447	1.877	-4.343	0.002	445 (12.9)	2739 (79.5)	263 (7.6)
	2007年度修正率	3408	3.506	-0.410	0.002	454 (13.3)	2657 (78.0)	297 (8.7)
12月調査	2007年度修正率	3359	2.856	-0.206	0.002	530 (15.8)	2443 (72.7)	386 (11.5)
2008年								
3月調査	2007年度修正率	3400	1.110	-0.342	0.001	461 (13.6)	2505 (73.7)	434 (12.8)
	2008年度前年比	3353	0.651	-3.141	-0.008	420 (12.5)	2082 (62.1)	851 (25.4)
6月調査	2007年度修正率	3357	0.808	-0.459	0.001	474 (14.1)	2559 (76.2)	324 (9.7)
	2008年度修正率	3357	1.150	-0.426	0.003	463 (13.8)	2626 (78.2)	268 (8.0)
9月調査	2008年度修正率	3310	2.826	-0.501	0.002	442 (13.4)	2557 (77.3)	311 (9.4)
12月調査	2008年度修正率	3265	0.523	-0.562	0.001	492 (15.1)	2385 (73.0)	388 (11.9)
2009年								
3月調査	2008年度修正率							
	2009年度前年比	3236	2.909	-3.145	-0.014	281 (8.7)	2001 (61.8)	954 (29.5)
6月調査	2008年度修正率	3245	2.393	-1.281	0.002	423 (13.0)	2548 (78.5)	274 (8.4)
	2009年度修正率	3245	0.671	-2.254	0.001	395 (12.2)	2620 (80.7)	230 (7.1)

(注) 設備投資額・ソフトウェア投資額は、リース会計基準の変更に伴い調査項目の定義を変更したため、09年3月調査の08年度修正率は算出してない。

(図表 6 - 9)

母集団推計値への影響度合い y_{it} の分布 (9)

(1) ソフトウェア投資額：大企業・製造業

		データ数	最大値	最小値	平均値	正数の数、割合(%)	0の数、割合(%)	負数の数、割合(%)
2007年								
3月調査	2007年度前年比	1231	1.282	-1.050	-0.003	267 (21.7)	625 (50.8)	339 (27.5)
	2006年度修正率	1234	2.213	-0.875	0.002	351 (28.4)	541 (43.8)	342 (27.7)
6月調査	2007年度修正率	1234	0.668	-0.906	0.002	291 (23.6)	715 (57.9)	228 (18.5)
	2007年度修正率	1228	0.732	-0.514	0.002	152 (12.4)	954 (77.7)	122 (9.9)
12月調査	2007年度修正率	1224	1.717	-0.199	0.007	362 (29.6)	536 (43.8)	326 (26.6)
2008年								
3月調査	2007年度修正率	1222	0.839	-0.820	0.000	163 (13.3)	871 (71.3)	188 (15.4)
	2008年度前年比	1211	2.066	-1.153	-0.001	273 (22.5)	623 (51.4)	315 (26.0)
6月調査	2007年度修正率	1214	0.715	-0.573	0.000	381 (31.4)	504 (41.5)	329 (27.1)
	2008年度修正率	1214	1.855	-0.672	0.007	301 (24.8)	679 (55.9)	234 (19.3)
9月調査	2008年度修正率	1209	0.785	-0.517	0.000	144 (11.9)	924 (76.4)	141 (11.7)
12月調査	2008年度修正率	1205	1.237	-0.580	0.000	334 (27.7)	522 (43.3)	349 (29.0)
2009年								
3月調査	2008年度修正率							
	2009年度前年比	1177	0.296	-2.228	-0.010	192 (16.3)	559 (47.5)	426 (36.2)
6月調査	2008年度修正率	1192	0.942	-0.897	0.001	429 (36.0)	435 (36.5)	328 (27.5)
	2009年度修正率	1192	0.988	-1.499	-0.004	280 (23.5)	602 (50.5)	310 (26.0)

(2) ソフトウェア投資額：中堅企業・製造業

		データ数	最大値	最小値	平均値	正数の数、割合(%)	0の数、割合(%)	負数の数、割合(%)
2007年								
3月調査	2007年度前年比	1235	4.456	-5.113	-0.006	222 (18.0)	770 (62.3)	243 (19.7)
	2006年度修正率	1235	1.414	-0.914	0.000	164 (13.3)	936 (75.8)	135 (10.9)
6月調査	2007年度修正率	1235	1.404	-0.928	0.007	118 (9.6)	1046 (84.7)	71 (5.7)
	2007年度修正率	1227	0.451	-1.228	-0.005	134 (10.9)	993 (80.9)	100 (8.1)
12月調査	2007年度修正率	1218	1.180	-0.911	0.010	176 (14.4)	877 (72.0)	165 (13.5)
2008年								
3月調査	2007年度修正率	1219	0.633	-1.736	-0.005	105 (8.6)	978 (80.2)	136 (11.2)
	2008年度前年比	1200	4.776	-11.27	0.009	219 (18.3)	753 (62.8)	228 (19.0)
6月調査	2007年度修正率	1201	0.712	-0.972	-0.001	173 (14.4)	883 (73.5)	145 (12.1)
	2008年度修正率	1201	0.921	-1.172	-0.001	86 (7.2)	1036 (86.3)	79 (6.6)
9月調査	2008年度修正率	1196	0.808	-1.042	-0.001	116 (9.7)	980 (81.9)	100 (8.4)
12月調査	2008年度修正率	1191	0.703	-3.497	-0.008	157 (13.2)	862 (72.4)	172 (14.4)
2009年								
3月調査	2008年度修正率							
	2009年度前年比	1166	1.180	-6.628	-0.019	147 (12.6)	712 (61.1)	307 (26.3)
6月調査	2008年度修正率	1171	2.454	-2.025	0.000	159 (13.6)	863 (73.7)	149 (12.7)
	2009年度修正率	1171	15.86	-2.470	0.017	113 (9.6)	982 (83.9)	76 (6.5)

(3) ソフトウェア投資額：中小企業・製造業

		データ数	最大値	最小値	平均値	正数の数、割合(%)	0の数、割合(%)	負数の数、割合(%)
2007年								
3月調査	2007年度前年比	2030	2.395	-2.732	-0.002	171 (8.4)	1620 (79.8)	239 (11.8)
	2006年度修正率	2031	1.163	-2.732	-0.003	126 (6.2)	1815 (89.4)	90 (4.4)
6月調査	2007年度修正率	2031	1.780	-2.657	0.003	140 (6.9)	1820 (89.6)	71 (3.5)
	2007年度修正率	2014	1.818	-1.221	0.001	128 (6.4)	1776 (88.2)	110 (5.5)
12月調査	2007年度修正率	2002	1.685	-1.293	0.001	141 (7.0)	1746 (87.2)	115 (5.7)
2008年								
3月調査	2007年度修正率	2004	1.107	-1.687	-0.004	106 (5.3)	1769 (88.3)	129 (6.4)
	2008年度前年比	1977	4.260	-2.099	-0.001	172 (8.7)	1573 (79.6)	232 (11.7)
6月調査	2007年度修正率	1977	2.197	-1.316	-0.001	104 (5.3)	1788 (90.4)	85 (4.3)
	2008年度修正率	1977	7.383	-1.734	0.009	130 (6.6)	1776 (89.8)	71 (3.6)
9月調査	2008年度修正率	1957	2.891	-6.184	0.000	114 (5.8)	1733 (88.6)	110 (5.6)
12月調査	2008年度修正率	1945	1.774	-1.663	-0.002	131 (6.7)	1694 (87.1)	120 (6.2)
2009年								
3月調査	2008年度修正率							
	2009年度前年比	1885	4.830	-3.985	-0.016	88 (4.7)	1493 (79.2)	304 (16.1)
6月調査	2008年度修正率	1896	3.974	-1.993	0.000	89 (4.7)	1709 (90.1)	98 (5.2)
	2009年度修正率	1896	1.757	-4.532	0.004	99 (5.2)	1723 (90.9)	74 (3.9)

(注) 設備投資額・ソフトウェア投資額は、リース会計基準の変更に伴い調査項目の定義を変更したため、09年3月調査の08年度修正率は算出してない。

(図表 6 - 1 0)

母集団推計値への影響度合い y_{it} の分布 (1 0)

(1) ソフトウェア投資額：大企業・非製造業

		データ数	最大値	最小値	平均値	正数の数、割合(%)	0の数、割合(%)	負数の数、割合(%)
2007年								
3月調査	2007年度前年比	1203	0.922	-3.014	-0.004	235 (19.5)	694 (57.7)	274 (22.8)
	2006年度修正率	1205	1.075	-0.806	0.000	290 (24.1)	617 (51.2)	298 (24.7)
6月調査	2007年度修正率	1205	0.626	-1.701	-0.001	245 (20.3)	744 (61.7)	216 (17.9)
	2007年度修正率	1199	0.448	-0.553	0.000	140 (11.7)	944 (78.7)	115 (9.6)
12月調査	2007年度修正率	1194	0.759	-0.741	-0.002	285 (23.9)	638 (53.4)	271 (22.7)
2008年								
3月調査	2007年度修正率	1195	1.867	-0.196	0.002	161 (13.5)	867 (72.6)	167 (14.0)
	2008年度前年比	1183	1.833	-0.639	0.003	231 (19.5)	659 (55.7)	293 (24.8)
6月調査	2007年度修正率	1184	3.079	-1.368	0.006	326 (27.5)	598 (50.5)	260 (22.0)
	2008年度修正率	1184	1.932	-0.804	0.007	273 (23.1)	739 (62.4)	172 (14.5)
9月調査	2008年度修正率	1175	0.464	-0.718	0.000	132 (11.2)	916 (78.0)	127 (10.8)
12月調査	2008年度修正率	1172	1.170	-2.810	-0.008	299 (25.5)	617 (52.6)	256 (21.8)
2009年								
3月調査	2008年度修正率							
	2009年度前年比	1132	0.478	-0.699	-0.005	182 (16.1)	620 (54.8)	330 (29.2)
6月調査	2008年度修正率	1149	4.683	-0.918	0.006	307 (26.7)	562 (48.9)	280 (24.4)
	2009年度修正率	1149	3.774	-1.626	0.003	226 (19.7)	675 (58.7)	248 (21.6)

(2) ソフトウェア投資額：中堅企業・非製造業

		データ数	最大値	最小値	平均値	正数の数、割合(%)	0の数、割合(%)	負数の数、割合(%)
2007年								
3月調査	2007年度前年比	1687	9.260	-2.052	0.010	227 (13.5)	1139 (67.5)	321 (19.0)
	2006年度修正率	1689	1.597	-5.481	0.001	222 (13.1)	1303 (77.1)	164 (9.7)
6月調査	2007年度修正率	1689	1.555	-11.00	-0.004	165 (9.8)	1416 (83.8)	108 (6.4)
	2007年度修正率	1672	1.348	-0.761	-0.001	153 (9.2)	1403 (83.9)	116 (6.9)
12月調査	2007年度修正率	1657	2.058	-1.081	0.001	196 (11.8)	1269 (76.6)	192 (11.6)
2008年								
3月調査	2007年度修正率	1649	0.898	-1.576	-0.001	153 (9.3)	1329 (80.6)	167 (10.1)
	2008年度前年比	1634	2.136	-2.973	-0.002	240 (14.7)	1079 (66.0)	315 (19.3)
6月調査	2007年度修正率	1634	3.344	-1.814	0.001	205 (12.5)	1248 (76.4)	181 (11.1)
	2008年度修正率	1634	1.987	-1.279	0.002	144 (8.8)	1382 (84.6)	108 (6.6)
9月調査	2008年度修正率	1620	0.531	-0.264	0.002	147 (9.1)	1350 (83.3)	123 (7.6)
12月調査	2008年度修正率	1616	0.961	-0.659	0.000	242 (15.0)	1175 (72.7)	199 (12.3)
2009年								
3月調査	2008年度修正率							
	2009年度前年比	1551	3.479	-3.325	-0.008	168 (10.8)	1022 (65.9)	361 (23.3)
6月調査	2008年度修正率	1562	2.565	-0.707	0.005	207 (13.3)	1185 (75.9)	170 (10.9)
	2009年度修正率	1562	1.351	-0.951	0.002	132 (8.5)	1302 (83.4)	128 (8.2)

(3) ソフトウェア投資額：中小企業・非製造業

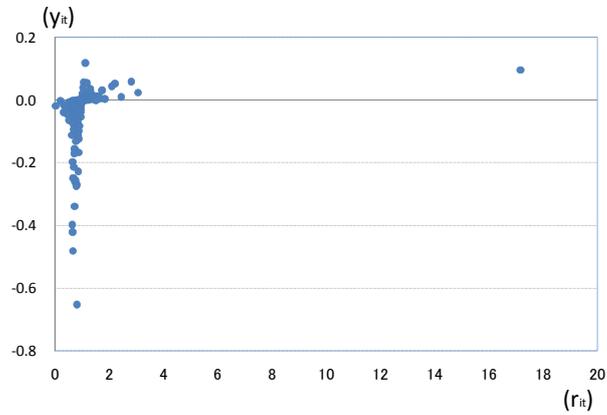
		データ数	最大値	最小値	平均値	正数の数、割合(%)	0の数、割合(%)	負数の数、割合(%)
2007年								
3月調査	2007年度前年比	3447	3.649	-1.275	-0.002	194 (5.6)	2975 (86.3)	278 (8.1)
	2006年度修正率	3447	1.313	-0.681	0.000	144 (4.2)	3165 (91.8)	138 (4.0)
6月調査	2007年度修正率	3447	1.174	-2.772	0.002	175 (5.1)	3190 (92.5)	82 (2.4)
	2007年度修正率	3408	0.974	-0.704	0.002	177 (5.2)	3111 (91.3)	120 (3.5)
12月調査	2007年度修正率	3359	4.799	-1.509	0.001	168 (5.0)	3055 (90.9)	136 (4.0)
2008年								
3月調査	2007年度修正率	3400	1.440	-1.692	-0.002	146 (4.3)	3037 (89.3)	217 (6.4)
	2008年度前年比	3353	1.381	-5.376	-0.004	171 (5.1)	2908 (86.7)	274 (8.2)
6月調査	2007年度修正率	3357	2.085	-1.171	0.000	147 (4.4)	3082 (91.8)	128 (3.8)
	2008年度修正率	3357	3.007	-0.614	0.006	168 (5.0)	3119 (92.9)	70 (2.1)
9月調査	2008年度修正率	3311	2.478	-1.684	0.001	159 (4.8)	3035 (91.7)	117 (3.5)
12月調査	2008年度修正率	3265	2.041	-1.911	0.001	142 (4.3)	2997 (91.8)	126 (3.9)
2009年								
3月調査	2008年度修正率							
	2009年度前年比	3235	4.658	-3.275	-0.011	143 (4.4)	2760 (85.3)	332 (10.3)
6月調査	2008年度修正率	3244	2.342	-3.156	0.000	129 (4.0)	3008 (92.7)	107 (3.3)
	2009年度修正率	3244	4.876	-3.276	0.008	155 (4.8)	3021 (93.1)	68 (2.1)

(注) 設備投資額・ソフトウェア投資額は、リース会計基準の変更に伴い調査項目の定義を変更したため、09年3月調査の08年度修正率は算出してない。

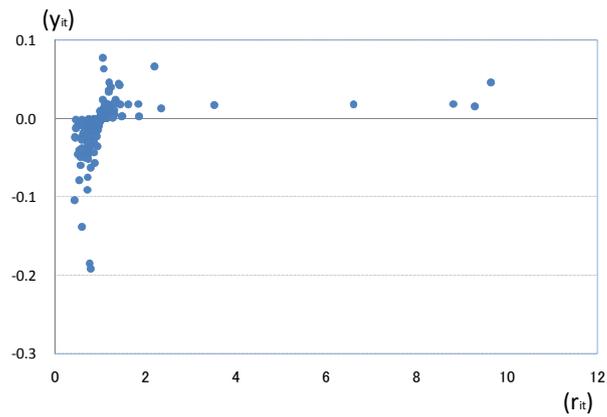
(図表 7 - 1)

比率 r_i と影響度合い y_{it} の散布図 (1)

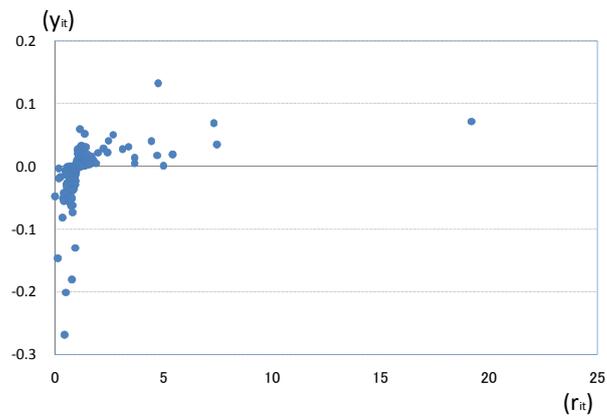
(1) 大企業・製造業 (2009 年 6 月調査、売上高 2009 年度修正率)



(2) 中堅企業・製造業 (2009 年 6 月調査、売上高 2009 年度修正率)



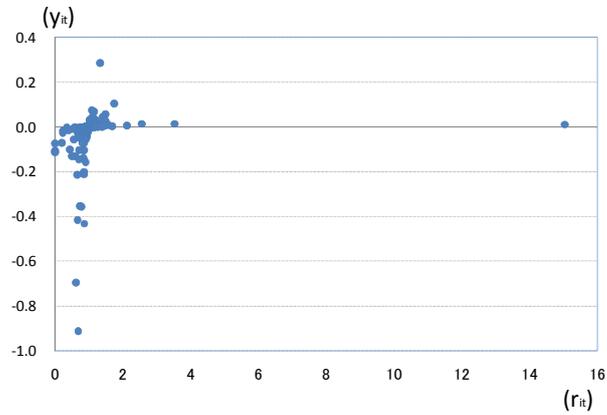
(3) 中小企業・製造業 (2009 年 6 月調査、売上高 2009 年度修正率)



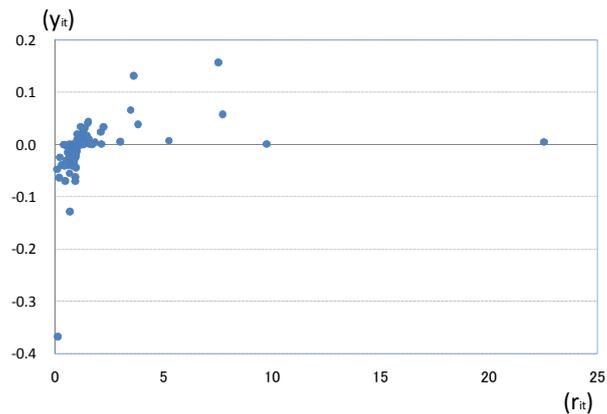
(図表 7 - 2)

比率 r_i と影響度合い y_{it} の散布図 (2)

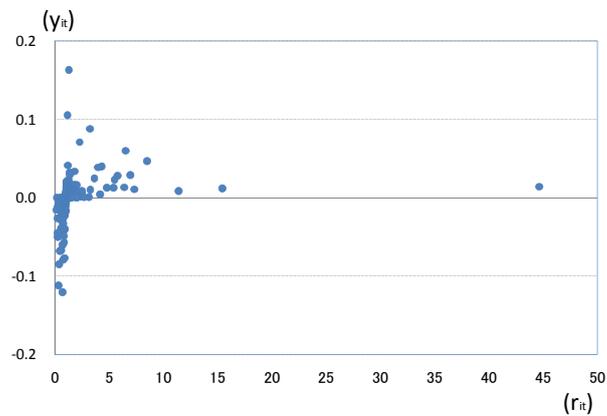
(1) 大企業・非製造業 (2009 年 6 月調査、売上高 2009 年度修正率)



(2) 中堅企業・非製造業 (2009 年 6 月調査、売上高 2009 年度修正率)¹⁵



(3) 中小企業・非製造業 (2009 年 6 月調査、売上高 2009 年度修正率)



¹⁵ 図表中のデータ以外に、 $(r_{it}, y_{it}) = (2950, 0.12)$ というデータがあるが、散布図からは省略している。

(図表 8)

比率を用いる方法を適用した場合の外れ値検出結果

(2009 年 6 月調査、売上高 2009 年度修正率)

(6 区分ベースのデータセットのうち外れ値が占める割合)

(1) 上限値 T_U を超えたもの

		$T_U=2$	$T_U=3$	$T_U=4$
製造業				
	大企業	0.50%	0.17%	0.08%
	中堅企業	0.59%	0.42%	0.34%
	中小企業	0.84%	0.63%	0.42%
非製造業				
	大企業	0.34%	0.17%	0.09%
	中堅企業	0.82%	0.63%	0.38%
	中小企業	0.91%	0.58%	0.43%

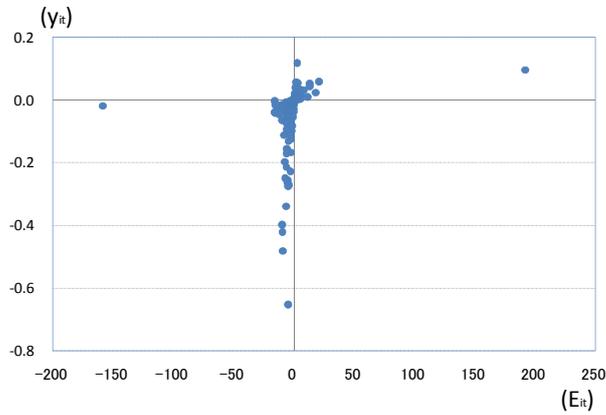
(2) 下限値 T_L を超えたもの

		$T_L=0.3$	$T_L=0.2$	$T_L=0.1$
製造業				
	大企業	0.17%	0.08%	0.08%
	中堅企業	0.00%	0.00%	0.00%
	中小企業	0.26%	0.21%	0.05%
非製造業				
	大企業	0.60%	0.43%	0.34%
	中堅企業	0.25%	0.19%	0.06%
	中小企業	0.21%	0.06%	0.00%

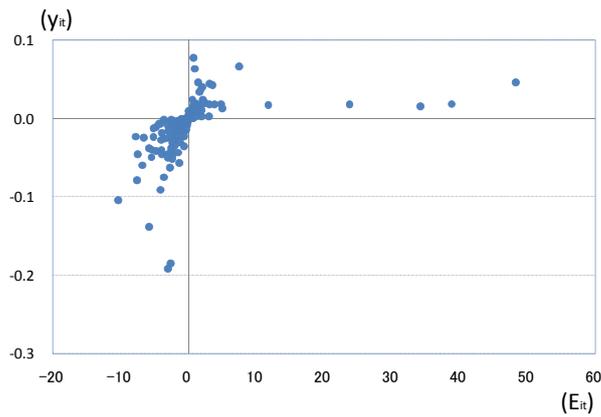
(注) 比率 r_{it} は、(今回調査の回答値) / (前回調査の回答値) により算出。この比率 r_{it} が、上限値 T_U または下限値 T_L を超えた場合、外れ値として検出。

E_{it} と影響度合い y_{it} の散布図 (1)

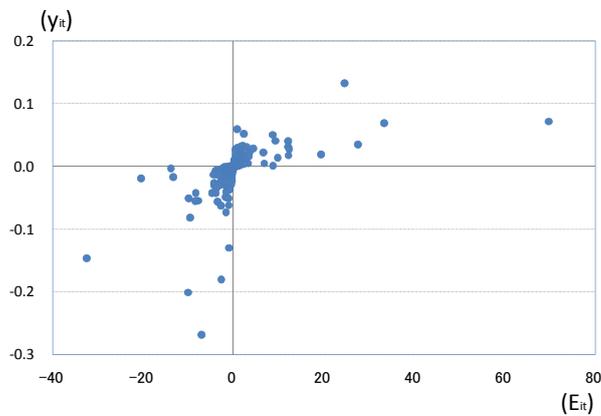
(1) 大企業・製造業 (2009 年 6 月調査、売上高 2009 年度修正率)



(2) 中堅企業・製造業 (2009 年 6 月調査、売上高 2009 年度修正率)



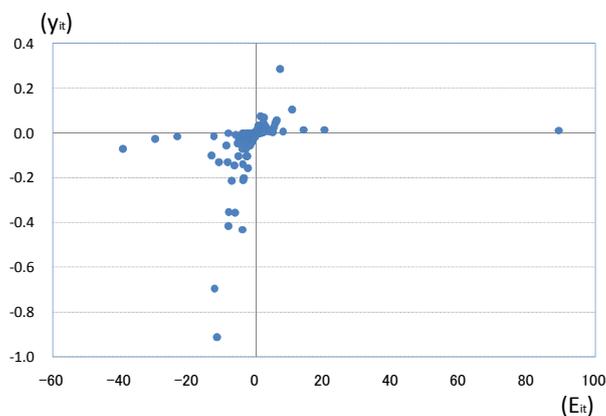
(3) 中小企業・製造業 (2009 年 6 月調査、売上高 2009 年度修正率)



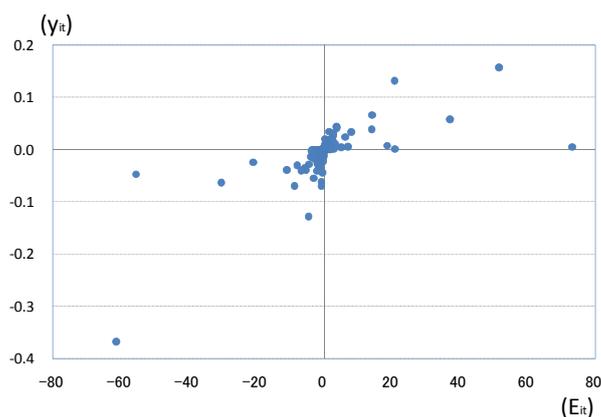
(図表 9 - 2)

E_{it} と影響度合い y_{it} の散布図 (2)

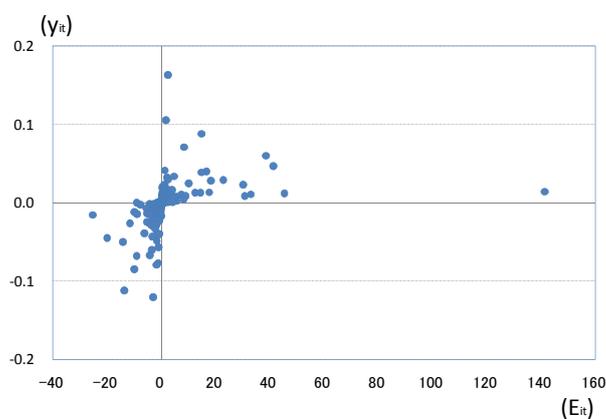
(1) 大企業・非製造業 (2009 年 6 月調査、売上高 2009 年度修正率)



(2) 中堅企業・非製造業 (2009 年 6 月調査、売上高 2009 年度修正率)¹⁶



(3) 中小企業・非製造業 (2009 年 6 月調査、売上高 2009 年度修正率)



¹⁶ 図表中のデータ以外に、 $(E_{it}, y_{it}) = (26535, 0.12)$ というデータがあるが、散布図からは省略している。

(図表 1 0)

Hidiroglou-Berthelot method を適用した場合の外れ値検出結果

(2009 年 6 月調査、売上高 2009 年度修正率)

(1) E_{it} の分布

	データ数	最小値	第1四分位数	メジアン	第3四分位数	最大値
製造業						
大企業	1,208	-159	-0.86	4.65E-05	0.13	192
中堅企業	1,178	-10	-0.27	0.00	0.00	48
中小企業	1,915	-32	-0.31	0.00	0.00	70
非製造業						
大企業	1,165	-39	-0.40	0.00	2.99E-03	89
中堅企業	1,585	-61	-0.13	0.00	0.00	26535
中小企業	3,283	-25	-0.14	0.00	4.12E-03	141

(2) 散らばりの尺度 D_L 、 D_U および許容区間

	散らばりの尺度		許容区間					
			c=20		c=30		c=40	
	D_L	D_U	下限	上限	下限	上限	下限	上限
製造業								
大企業	0.86	0.13	-17.15	2.68	-25.72	4.02	-34.30	5.37
中堅企業	0.27	0.00	-5.33	0.00	-8.00	0.00	-10.66	0.00
中小企業	0.31	0.00	-6.11	0.00	-9.17	0.00	-12.23	0.00
非製造業								
大企業	0.40	2.99E-03	-8.00	0.06	-12.00	0.09	-16.00	0.12
中堅企業	0.13	0.00	-2.63	0.00	-3.94	0.00	-5.25	0.00
中小企業	0.14	4.12E-03	-2.87	0.08	-4.31	0.12	-5.75	0.16

(3) 6 区分ベースのデータセットのうち外れ値が占める割合

	c=20		c=30		c=40	
	下限	上限	下限	上限	下限	上限
製造業						
大企業	0.08%	1.82%	0.08%	1.16%	0.08%	0.66%
中堅企業	0.76%	22.75%	0.08%	22.75%	0.00%	22.75%
中小企業	0.57%	23.24%	0.37%	23.24%	0.21%	23.24%
非製造業						
大企業	1.12%	20.60%	0.52%	18.80%	0.26%	17.51%
中堅企業	1.32%	24.10%	0.76%	24.10%	0.63%	24.10%
中小企業	0.91%	20.07%	0.58%	17.85%	0.37%	15.32%

(図表 1 1 - 1)

単位距離D_j (1)

(1) 売上高

		大・製造	中堅・製造	中小・製造	大・非製	中堅・非製	中小・非製
2007年							
3月調査	2007年度前年比	0.044	0.066	0.042	0.051	0.050	0.033
	2006年度修正率	0.021	0.023	0.024	0.027	0.025	0.016
6月調査	2007年度修正率	0.057	0.049	0.034	0.051	0.035	0.020
	2006年度修正率	0.032	0.049	0.027	0.024	0.039	0.026
9月調査	2007年度修正率	0.039	0.026	0.022	0.047	0.031	0.020
12月調査	2007年度修正率						
2008年							
3月調査	2007年度修正率	0.025	0.034	0.037	0.025	0.029	0.028
	2008年度前年比	0.034	0.078	0.045	0.037	0.072	0.031
6月調査	2007年度修正率	0.025	0.027	0.020	0.047	0.023	0.017
	2008年度修正率	0.081	0.045	0.042	0.090	0.040	0.024
9月調査	2008年度修正率	0.039	0.059	0.038	0.045	0.038	0.024
12月調査	2008年度修正率	0.064	0.043	0.035	0.081	0.026	0.023
2009年							
3月調査	2008年度修正率	0.131	0.092	0.045	0.088	0.050	0.035
	2009年度前年比	0.099	0.128	0.083	0.154	0.099	0.048
6月調査	2008年度修正率	0.046	0.036	0.022	0.053	0.030	0.018
	2009年度修正率	0.237	0.071	0.065	0.171	0.053	0.034

(2) 経常利益

		大・製造	中堅・製造	中小・製造	大・非製	中堅・非製	中小・非製
2007年							
3月調査	2007年度前年比	0.142	0.264	0.259	0.155	0.275	0.172
	2006年度修正率	0.092	0.129	0.169	0.106	0.170	0.120
6月調査	2007年度修正率	0.209	0.155	0.172	0.185	0.199	0.127
	2006年度修正率	0.104	0.191	0.183	0.096	0.135	0.128
9月調査	2007年度修正率	0.138	0.155	0.185	0.219	0.169	0.132
12月調査	2007年度修正率						
2008年							
3月調査	2007年度修正率	0.096	0.162	0.180	0.156	0.153	0.117
	2008年度前年比	0.191	0.278	0.314	0.151	0.339	0.207
6月調査	2007年度修正率	0.182	0.171	0.195	0.181	0.156	0.136
	2008年度修正率	0.329	0.218	0.222	0.218	0.157	0.136
9月調査	2008年度修正率	0.146	0.229	0.256	0.203	0.254	0.144
12月調査	2008年度修正率	0.313	0.291	0.284	0.329	0.187	0.183
2009年							
3月調査	2008年度修正率	0.980	0.451	0.348	0.540	0.267	0.195
	2009年度前年比	1.191	0.953	0.950	0.326	0.593	0.295
6月調査	2008年度修正率	0.702	0.302	0.492	0.357	0.288	0.177
	2009年度修正率	1.370	0.809	0.973	0.710	0.280	0.187

(3) 当期純利益

		大・製造	中堅・製造	中小・製造	大・非製	中堅・非製	中小・非製
2007年							
3月調査	2007年度前年比	0.248	0.659	0.496	0.494	0.677	0.313
	2006年度修正率	0.278	0.324	0.317	0.246	0.532	0.209
6月調査	2007年度修正率	0.280	0.291	0.256	0.347	0.310	0.165
	2006年度修正率	0.145	0.273	0.227	0.167	0.234	0.200
9月調査	2007年度修正率	0.194	0.246	0.312	0.307	0.370	0.255
12月調査	2007年度修正率						
2008年							
3月調査	2007年度修正率	0.166	0.265	0.287	0.377	0.326	0.214
	2008年度前年比	0.278	0.463	0.497	0.522	0.816	0.398
6月調査	2007年度修正率	0.238	0.359	0.317	0.356	0.419	0.261
	2008年度修正率	0.378	0.329	0.275	0.374	0.229	0.198
9月調査	2008年度修正率	0.143	0.358	0.336	0.273	0.362	0.208
12月調査	2008年度修正率	0.411	0.474	0.422	0.489	0.425	0.334
2009年							
3月調査	2008年度修正率	1.372	0.860	0.647	0.880	0.492	0.398
	2009年度前年比	343.7	4.946	3.302	0.884	1.246	0.735
6月調査	2008年度修正率	318.0	4.015	1.880	0.951	0.923	0.451
	2009年度修正率	4.244	1.656	2.720	1.202	0.495	0.309

(図表 1 1 - 2)

単位距離D_j (2)

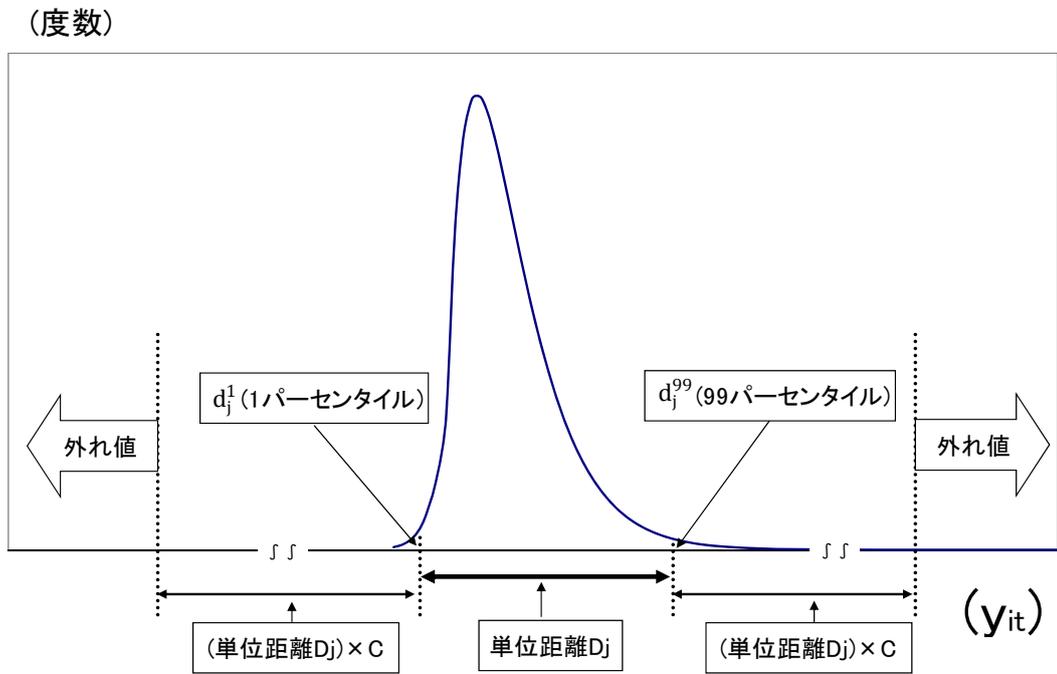
(4) 設備投資額

		大・製造	中堅・製造	中小・製造	大・非製	中堅・非製	中小・非製
2007年							
3月調査	2007年度前年比	0.269	0.519	0.660	0.225	0.413	0.339
	2006年度修正率	0.194	0.185	0.109	0.173	0.135	0.072
6月調査	2007年度修正率	0.289	0.222	0.211	0.202	0.150	0.155
	2007年度修正率	0.069	0.151	0.158	0.048	0.113	0.105
12月調査	2007年度修正率	0.132	0.194	0.219	0.133	0.158	0.114
2008年							
3月調査	2007年度修正率	0.064	0.197	0.190	0.164	0.162	0.104
	2008年度前年比	0.281	0.629	0.684	0.267	0.406	0.398
6月調査	2007年度修正率	0.173	0.177	0.131	0.179	0.110	0.080
	2008年度修正率	0.247	0.273	0.277	0.264	0.157	0.151
9月調査	2008年度修正率	0.084	0.203	0.194	0.069	0.116	0.136
12月調査	2008年度修正率	0.138	0.244	0.156	0.191	0.144	0.103
2009年							
3月調査	2008年度修正率						
	2009年度前年比	0.293	0.546	0.633	0.353	0.649	0.417
6月調査	2008年度修正率	0.224	0.177	0.165	0.224	0.145	0.091
	2009年度修正率	0.452	0.354	0.326	0.288	0.273	0.201

(5) ソフトウェア投資額

		大・製造	中堅・製造	中小・製造	大・非製	中堅・非製	中小・非製
2007年							
3月調査	2007年度前年比	0.371	1.153	0.926	0.274	0.359	0.478
	2006年度修正率	0.395	0.410	0.285	0.294	0.188	0.148
6月調査	2007年度修正率	0.498	0.514	0.449	0.266	0.133	0.174
	2007年度修正率	0.111	0.394	0.302	0.084	0.132	0.154
12月調査	2007年度修正率	0.252	0.452	0.413	0.295	0.150	0.169
2008年							
3月調査	2007年度修正率	0.169	0.274	0.375	0.166	0.194	0.141
	2008年度前年比	0.307	1.069	1.139	0.279	0.389	0.428
6月調査	2007年度修正率	0.349	0.474	0.354	0.297	0.210	0.134
	2008年度修正率	0.375	0.397	0.587	0.440	0.201	0.187
9月調査	2008年度修正率	0.129	0.297	0.349	0.097	0.121	0.182
12月調査	2008年度修正率	0.229	0.494	0.384	0.291	0.169	0.155
2009年							
3月調査	2008年度修正率						
	2009年度前年比	0.389	1.191	0.912	0.371	0.680	0.593
6月調査	2008年度修正率	0.383	0.457	0.283	0.481	0.417	0.189
	2009年度修正率	0.505	0.743	0.549	0.629	0.306	0.306

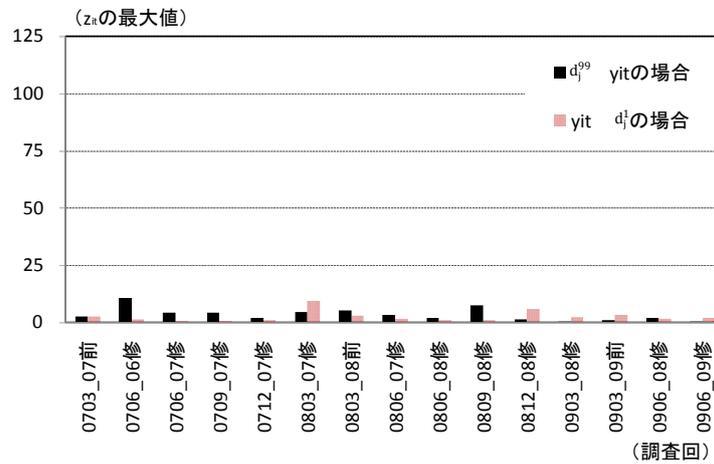
外れ値の検出方法



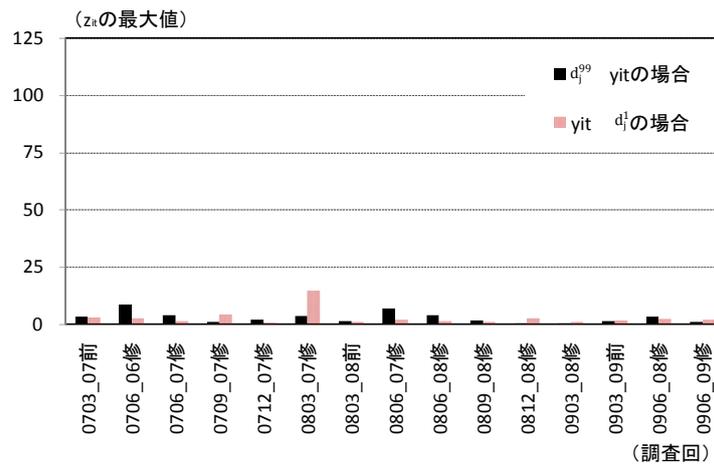
(図表 1 3 - 1)

外れ度合いの指標 z_{it} の最大値 (1)

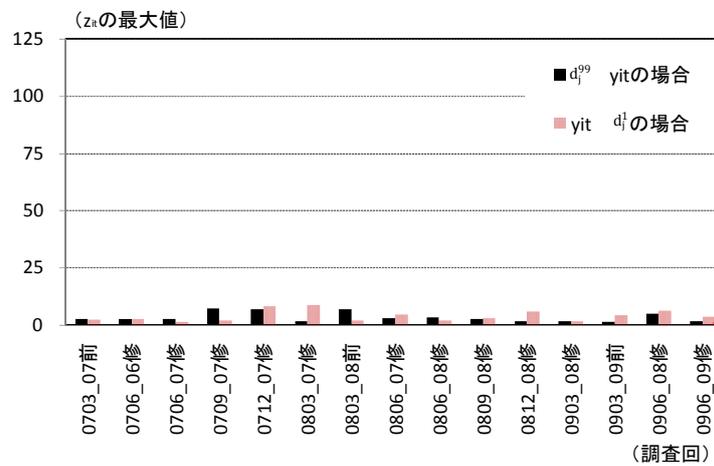
(1) 売上高：大企業・製造業



(2) 売上高：中堅企業・製造業



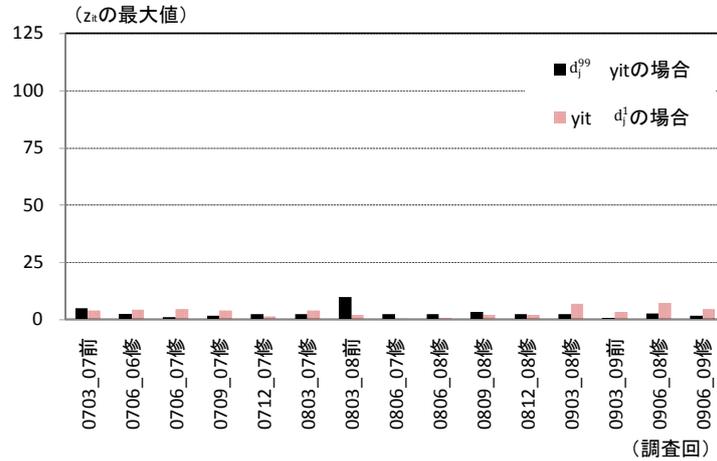
(3) 売上高：中小企業・製造業



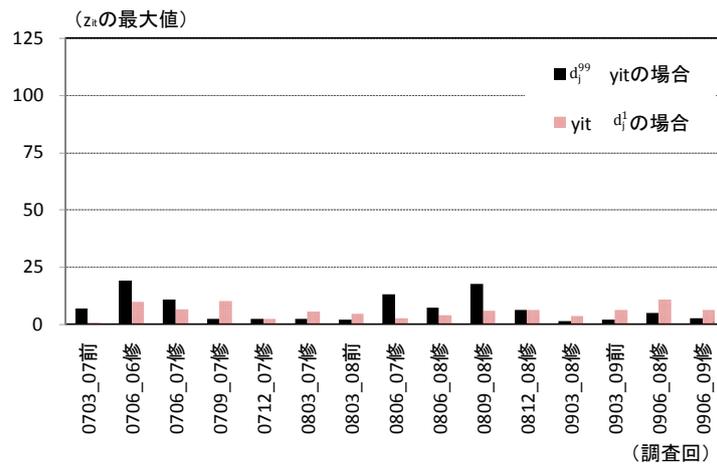
(図表 1 3 - 2)

外れ度合いの指標 z_{it} の最大値 (2)

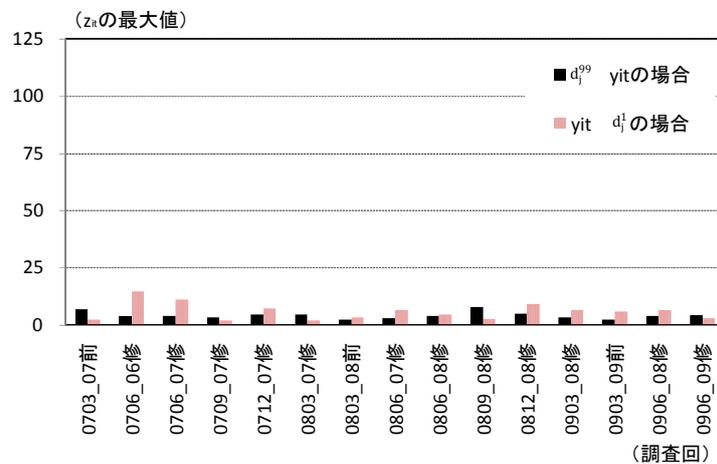
(1) 売上高：大企業・非製造業



(2) 売上高：中堅企業・非製造業



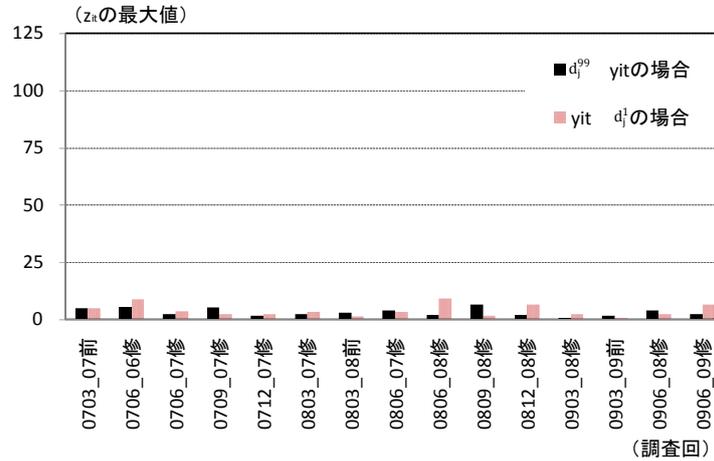
(3) 売上高：中小企業・非製造業



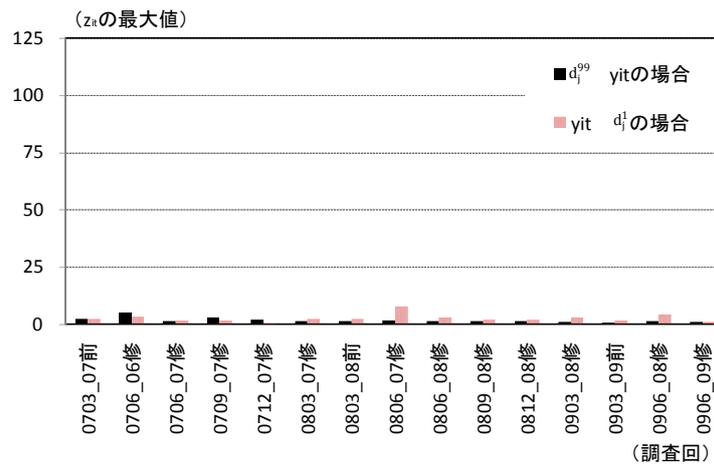
(図表 1 3 - 3)

外れ度合いの指標 z_{it} の最大値 (3)

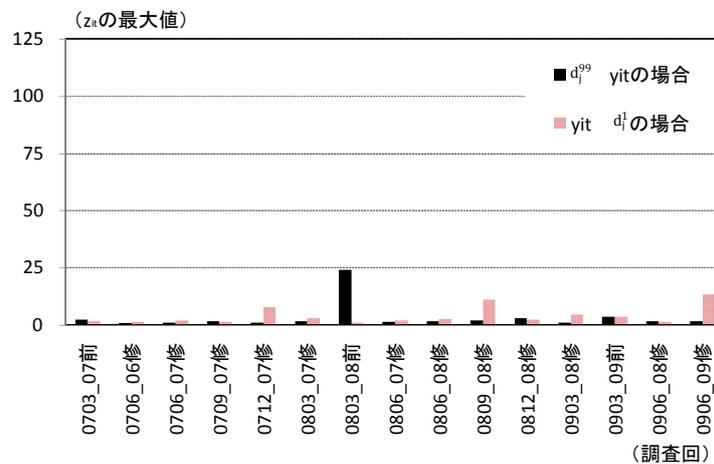
(1) 経常利益：大企業・製造業



(2) 経常利益：中堅企業・製造業



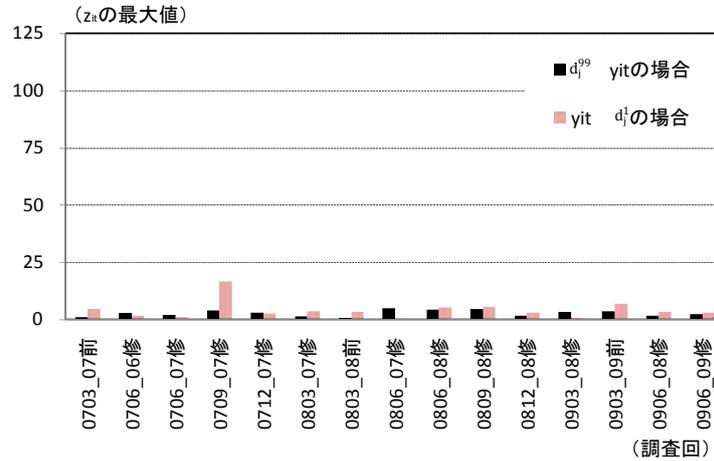
(3) 経常利益：中小企業・製造業



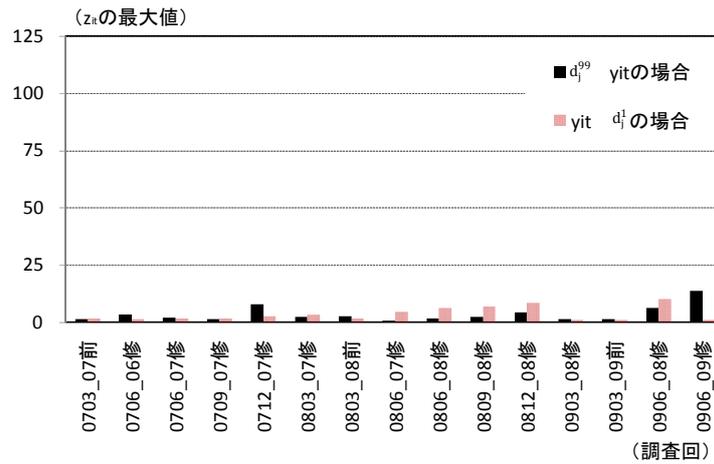
(図表 1 3 - 4)

外れ度合いの指標 z_{it} の最大値 (4)

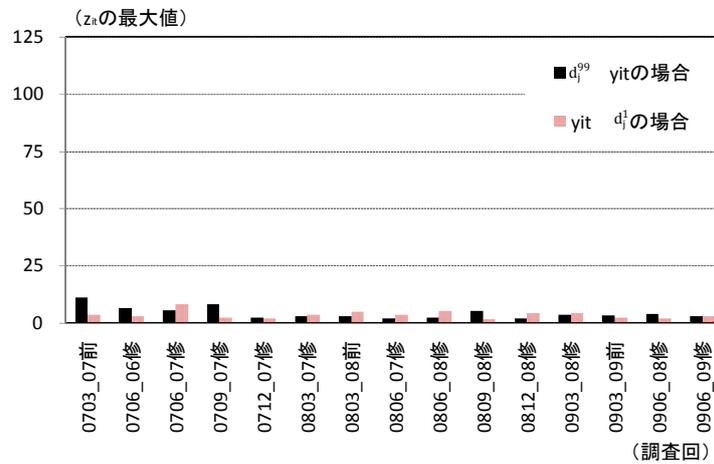
(1) 経常利益：大企業・非製造業



(2) 経常利益：中堅企業・非製造業



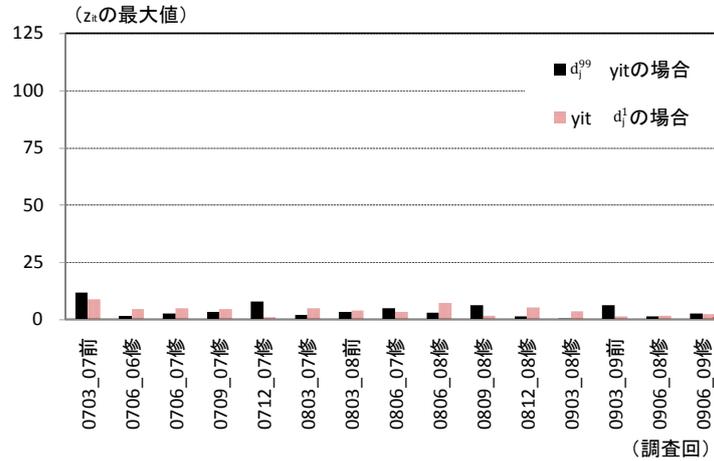
(3) 経常利益：中小企業・非製造業



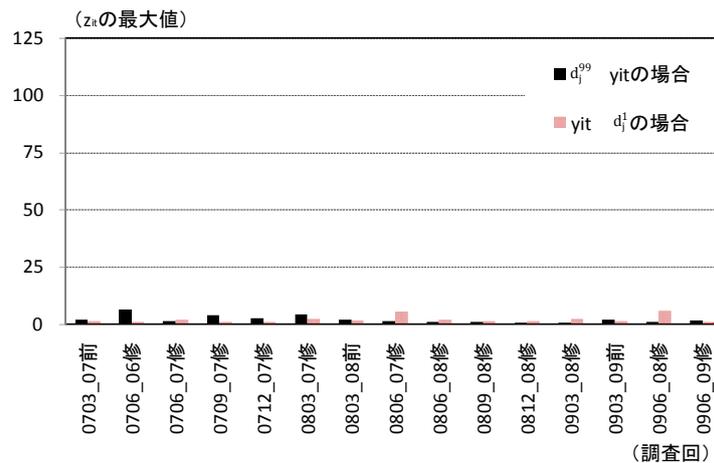
(図表 1 3 - 5)

外れ度合いの指標 z_{it} の最大値 (5)

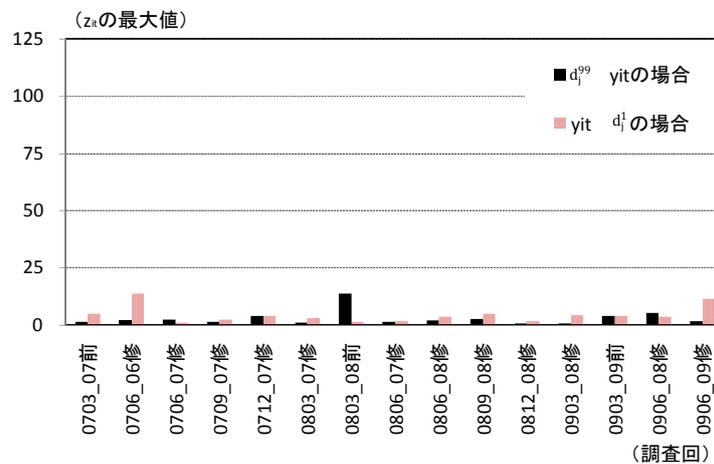
(1) 当期純利益 : 大企業・製造業



(2) 当期純利益 : 中堅企業・製造業



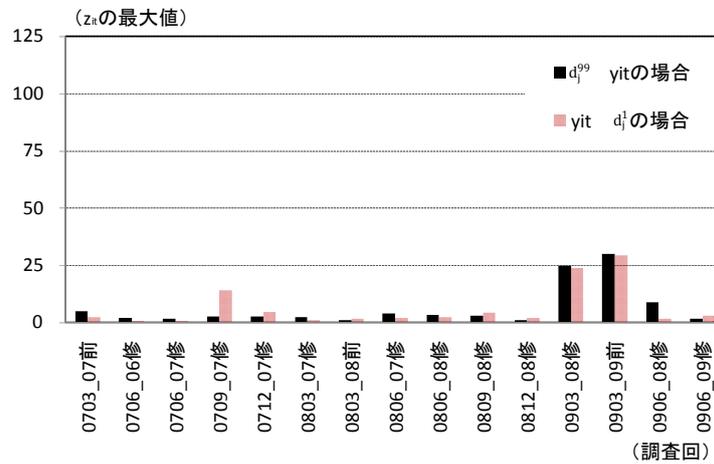
(3) 当期純利益 : 中小企業・製造業



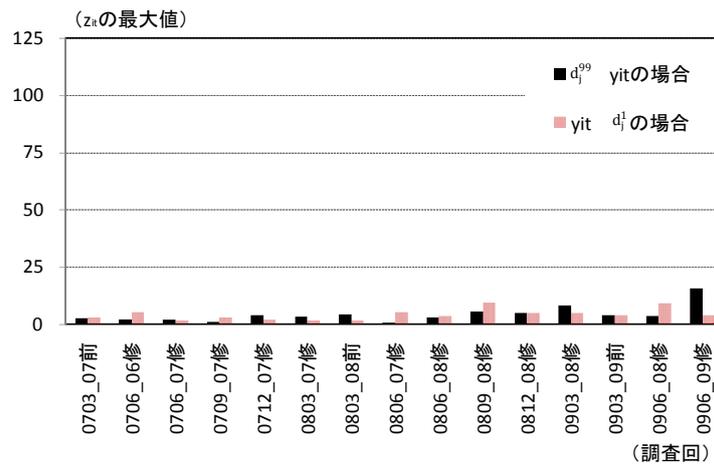
(図表 1 3 - 6)

外れ度合いの指標 z_{it} の最大値 (6)

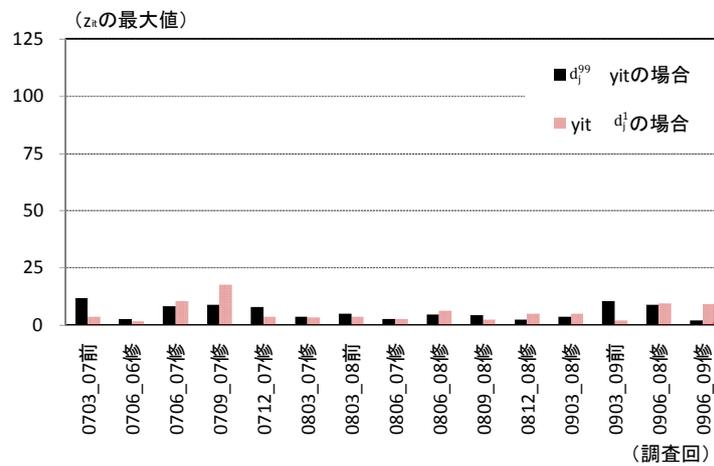
(1) 当期純利益 : 大企業・非製造業



(2) 当期純利益 : 中堅企業・非製造業



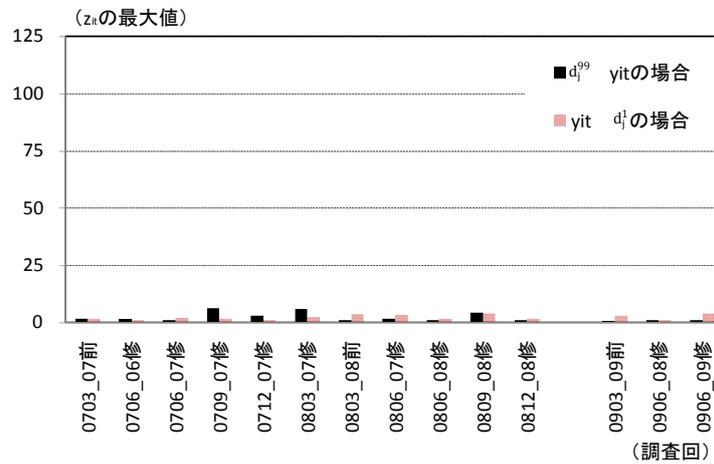
(3) 当期純利益 : 中小企業・非製造業



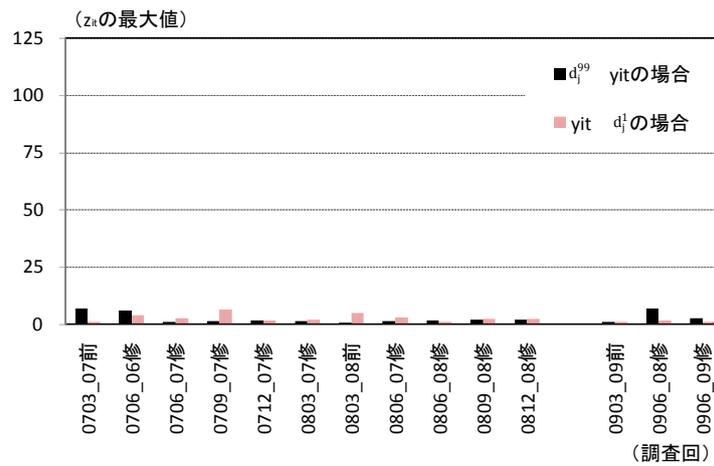
(図表 1 3 - 7)

外れ度合いの指標 z_{it} の最大値 (7)

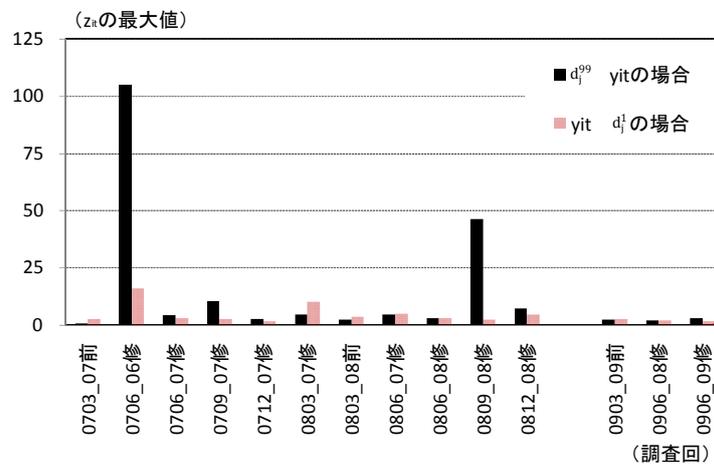
(1) 設備投資額：大企業・製造業



(2) 設備投資額：中堅企業・製造業



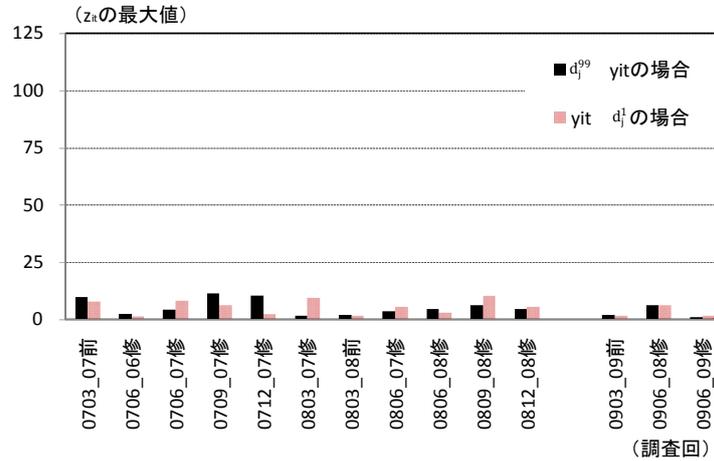
(3) 設備投資額：中小企業・製造業



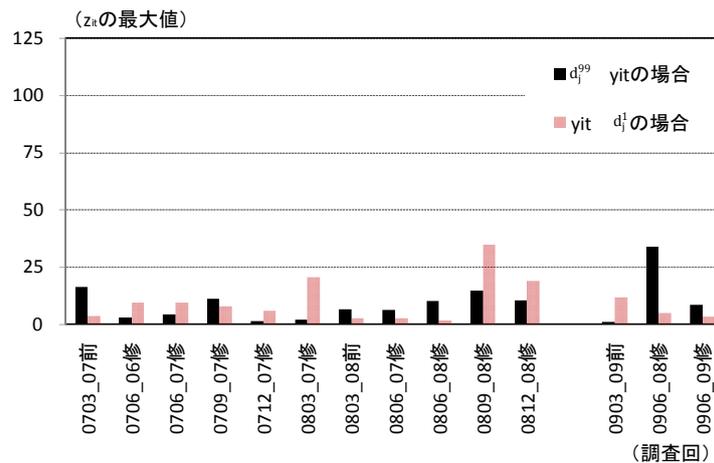
(図表 1 3 - 8)

外れ度合いの指標 z_{it} の最大値 (8)

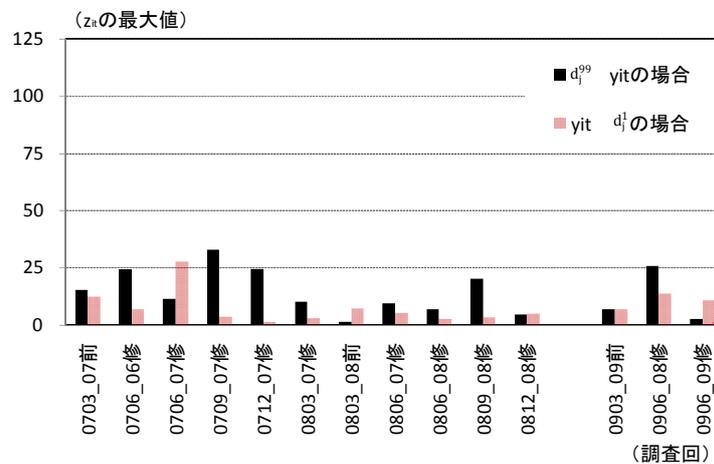
(1) 設備投資額：大企業・非製造業



(2) 設備投資額：中堅企業・非製造業



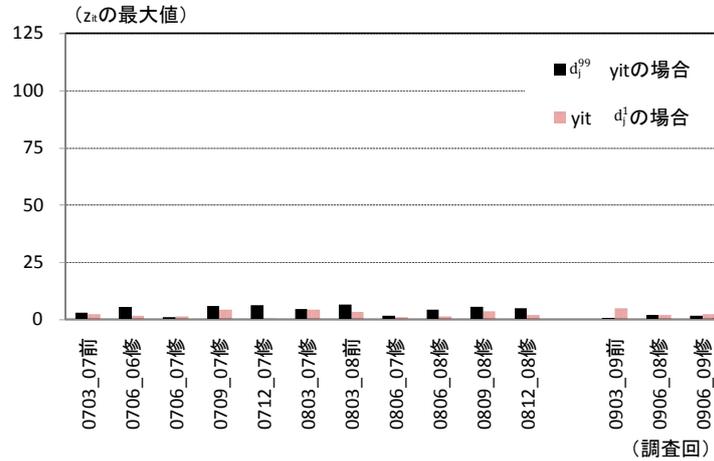
(3) 設備投資額：中小企業・非製造業



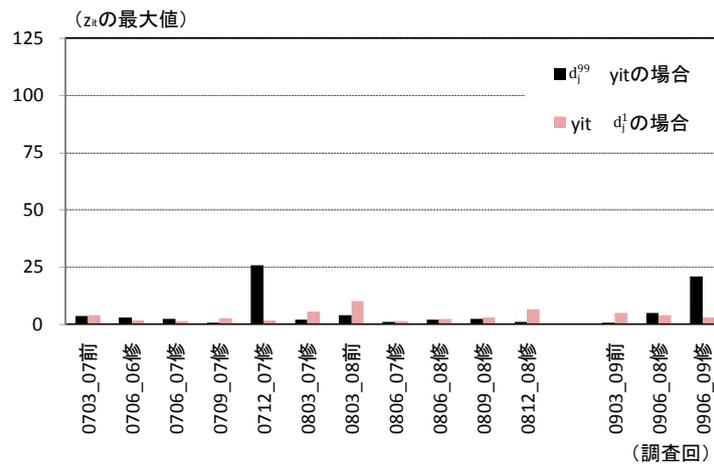
(図表 1 3 - 9)

外れ度合いの指標 z_{it} の最大値 (9)

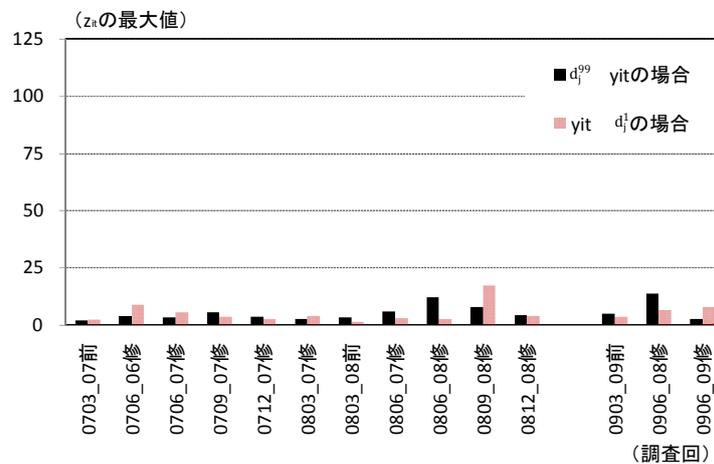
(1) ソフトウェア投資額：大企業・製造業



(2) ソフトウェア投資額：中堅企業・製造業



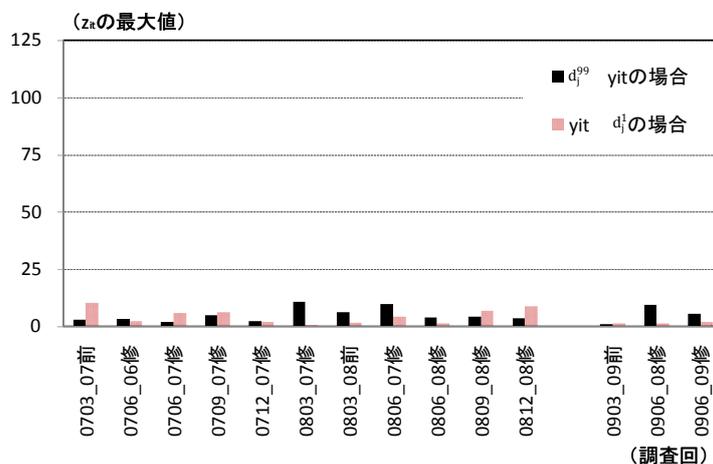
(3) ソフトウェア投資額：中小企業・製造業



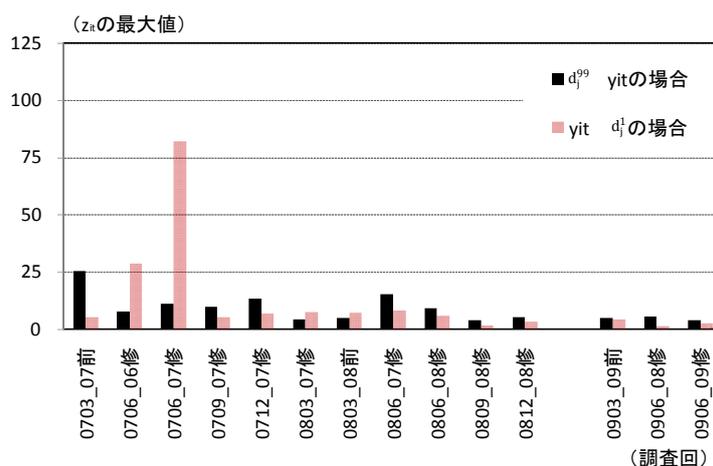
(図表 1 3 - 1 0)

外れ度合いの指標 z_{it} の最大値 (1 0)

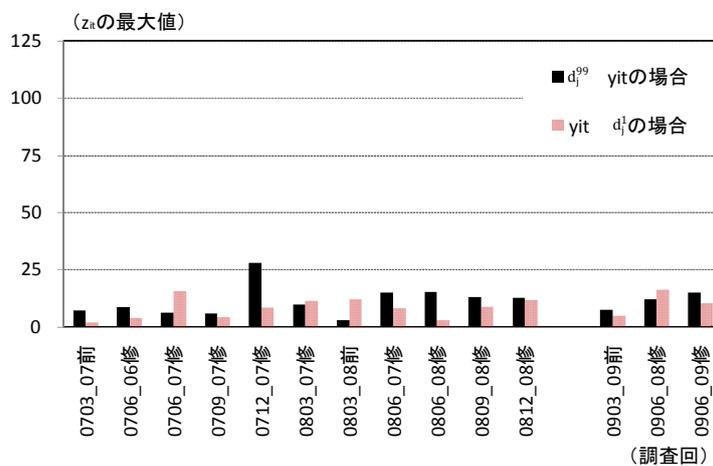
(1) ソフトウェア投資額：大企業・非製造業



(2) ソフトウェア投資額：中堅企業・非製造業



(3) ソフトウェア投資額：中小企業・非製造業



(図表 1 4)

基準 C の設定値と外れ値の検出数

	C=25	C=50	C=75	C=100	C=125
5項目合計	14	2	2	1	0
売上高	0	0	0	0	0
製造業					
大企業	0	0	0	0	0
中堅企業	0	0	0	0	0
中小企業	0	0	0	0	0
非製造業					
大企業	0	0	0	0	0
中堅企業	0	0	0	0	0
中小企業	0	0	0	0	0
経常利益	0	0	0	0	0
製造業					
大企業	0	0	0	0	0
中堅企業	0	0	0	0	0
中小企業	0	0	0	0	0
非製造業					
大企業	0	0	0	0	0
中堅企業	0	0	0	0	0
中小企業	0	0	0	0	0
当期純利益	2	0	0	0	0
製造業					
大企業	0	0	0	0	0
中堅企業	0	0	0	0	0
中小企業	0	0	0	0	0
非製造業					
大企業	2	0	0	0	0
中堅企業	0	0	0	0	0
中小企業	0	0	0	0	0
設備投資額	7	1	1	1	0
製造業					
大企業	0	0	0	0	0
中堅企業	0	0	0	0	0
中小企業	2	1	1	1	0
非製造業					
大企業	0	0	0	0	0
中堅企業	2	0	0	0	0
中小企業	3	0	0	0	0
ソフトウェア投資額	5	1	1	0	0
製造業					
大企業	0	0	0	0	0
中堅企業	1	0	0	0	0
中小企業	0	0	0	0	0
非製造業					
大企業	0	0	0	0	0
中堅企業	3	1	1	0	0
中小企業	1	0	0	0	0

外れ値処理の影響試算

(1) 外れ値処理を行うデータの y_{it} と z_{it}

	母集団推計値への影響度合い y_{it}	外れ度合いの指標 z_{it}
2007年6月調査 中小企業・製造業 2006年度設備投資額	+ 11.49	+ 104.81
2007年6月調査 中堅企業・非製造業 2007年度ソフトウェア投資額	11.00	+ 82.18
2008年9月調査 中小企業・製造業 2008年度設備投資額	+ 9.08	+ 46.11

(2) 外れ値処理後の y_{it} と6区分母集団推計値への影響試算

		外れ値 処理前	外れ値処理後		
			(a) 横置き 補完	(b) 平均値 補完	(c) 伸び率 補完
2007年6月調査 中小企業・製造業 2006年度設備投資額	6区分 修正率	+ 11.51	+ 0.02	0.87	0.05
	外れ値 の y_{it}	+ 11.49	0.00	0.89	0.07
2007年6月調査 中堅企業・非製造業 2007年度ソフトウェア投資額	6区分 修正率	6.32	+ 4.68	23.15	+ 12.06
	外れ値 の y_{it}	11.00	0.00	27.83	+ 7.38
2008年9月調査 中小企業・製造業 2008年度設備投資額	6区分 修正率	+ 11.34	+ 2.16	+ 0.20	+ 2.20
	外れ値 の y_{it}	+ 9.08	0.00	1.94	+ 0.04

(a) 横置き補完：回答値に前回調査と同値を代入する方法。前回調査からの変化幅が除外される。

(b) 平均値補完：外れ値が属する層の平均値を代入する方法。当該回答値のウェイトを「0」として集計から外すのと同義。

(c) 伸び率補完：外れ値を除いた層合計により層全体の修正率を算出。当修正率を外れ値の前回回答値に乗じた計数を用いる方法。