

# 事前不正確率の違いによるロジスティック回帰を用いた 不正会計の検知モデルの検知精度

佐 藤 夏 輝 ・ 小 村 亜 唯 子  
平 井 裕 久

## 要 旨

ロジスティック回帰を利用した不正会計企業を検知するためのモデルを構築する研究が進められている。構築された不正会計検知モデルの精度比較には、Correct Classification、Sensitivity、Type I Error、Type II Error等の指標が用いられている。本稿では、ロジスティック回帰によって構築した不正会計検知モデルの精度が、①トレーニングデータの事前不正確率の違いによってどのように変化するのか、②ロジスティック回帰によって推定された不正会計実施確率のカットオフ値に、事前不正確率を加味した場合と加味しない場合で、その精度はどのように異なるかを明らかにすることを目的としている。

実証においては、2011年から2019年までの期間で東京証券取引所に上場している企業を対象とする。不正会計検知モデルを構築し、その精度比較を行った結果、①事前不正確率が2.5%から5%の間では、事前不正確率が高まるほどPrecisionの値が改善するが、5%から50%にかけてその値が悪化することが明らかになった。また、②事前不正確率を加味しない場合に比較して、事前不正確率を加味した場合にはCorrect Classificationは低い、Sensitivityが高いことが明らかになった。

## 1. はじめに

不正会計とは、Beasley et al. (2010) によれば「財務諸表・財務開示（財務諸表の注記、米国証券取引委員会（Securities and Exchange Commission：SEC）への提出書類）における意図的で重大な虚偽記載、財務諸表・財務開示に重要な直接的影響を与える違法行為の実行」（p.5）とされている<sup>1</sup>。不正会計が行われると、投資家等の企業のステークホルダーが影響を受けるだけでなく、財務報告プロセス自体に信頼性があるのかという懸念を抱かせる（Beasley et al., 2010, p.5）。そのため、取締役会、社外監査役、監査人、規制当局、アナリスト等を含む主要なステークホルダーが、不正会計を防止・発見するための研究が求められている（Beasley et al., 2010, p.2）。

---

1 不正会計の定義は、未だ定まっていないといわれている（大城, 2014）。

そこで、財務諸表から得られる会計数値のデータ、Form 10-K等の報告書における記述、市場に関するデータ等を利用して、不正会計を行っている企業を検知するためのモデル（以下、不正会計の検知モデル）を構築する研究が進められている。不正会計の検知モデルの構築にあたり、代表的な研究とされているのがDechow et al. (2011)の研究である。Dechow et al. (2011)は、それまでの会計分野における先行研究の理論や実証結果を基に、不正会計企業の特徴を示す要因を変数化し、ロジスティック回帰による不正会計の検知モデルを構築した。構築された不正会計検知モデルの意義は、監査プロセスにおいてさらなる追加的な調査を必要とする企業を特定するためのスクリーニングとして使用できることとされている。Dechow et al. (2011)の研究を受けて、日本企業を対象としたロジスティック回帰を用いた不正会計の検知モデルの構築には、首藤他(2016)や佐藤他(2022)等が取り組んでいる。これらの研究では、ロジスティック回帰モデルで推定した不正実施確率(*Probability*)を、事前不正確率(*Prior Fraud Probability*)で除した値をFスコアと呼び、この値が1.00以上の時に不正会計企業、1.00未満の時に非不正会計企業と判定している<sup>2</sup>。

Dechow et al. (2011)以前にも、Fanning and Cogger (1998)、Lin et al. (2003)、Kotsiantis et al. (2006)等がロジスティック回帰を用いて不正会計の検知モデルの構築が行われている。これらの研究では、ロジスティック回帰モデルで推定された不正実施確率(*Probability*)が0.500以上の時に不正会計企業、0.500未満の時に非不正会計企業と判定している。すなわち、Dechow et al. (2011)以降の研究は、不正実施確率(*Probability*)に事前不正確率を加味しているのに対し、Dechow et al. (2011)よりも前の研究は事前不正確率を加味していないという違いがある。

Dechow et al. (2011)ではFスコアを0.00から4.00まで、首藤他(2016)では0.00から5.00までの範囲で、検知精度がどのように異なるのかを明らかにしている。すなわち、Dechow et al. (2011)では不正実施確率(*Probability*)が0.000から0.016の範囲、また首藤他(2016)では0.000から0.024の範囲における検知精度の推移を明らかにしている。しかし、Fanning and Cogger (1998)、Lin et al. (2003)、Kotsiantis et al. (2006)等では、不正会計と非不正会計の判定に用いた不正実施確率(*Probability*)を0.500として検知精度が検討されている。そのために、Fスコアを用いているDechow et al. (2011)、首藤他(2016)、佐藤他(2022)等と、Fスコアを用いていないFanning and Cogger (1998)、Lin et al. (2003)、Kotsiantis et al. (2006)等の不正会計の検知モデルの精度について、単純に比較できないという問題が残っている。

加えて、近年の研究では、不正会計企業と非不正会計企業の観測数に偏りがあり、そのようなデータで不正会計検知モデルを構築することに対して、アンダーサンプリング

2 ロジスティック回帰モデルで推定された不正実施確率(*Probability*)と事前不正確率とが一致するとき、Fスコアは1.00である。これをDechow et al. (2011)は、「その企業が無条件の期待値(The Unconditional Expectation)と同じ確率で不正会計を実施していること」(p.61)と解釈している。

が有効であるという指摘(Perols et al., 2017 ; Bao et al., 2020)がなされている。アンダーサンプリングとは、観測値の多い非不正会計企業のサンプルからランダムに観測値を削除し、データの不均衡を解消する手法である。すなわち、アンダーサンプリングによって、事前不正確率の値が変化する。また、データセットから求める事前不正確率は、現在、不正会計企業として摘発された企業とそれ以外の企業の比率であり、不正会計を実施しているがまだ摘発されていない企業が含まれている可能性もある (Dechow et al., 2011)。そこで、複数のパターンの事前不正確率をアンダーサンプリングによって設定し、不正会計の検知モデルの精度を比較、検討する必要がある。このようなアプローチをとった研究として、事前不正確率に10%、20%、33%、50%という4パターンを設定して検知精度の比較を行ったFeroz et al. (2000) がある。しかし、検知精度の比較に際して各パターンにおける試行回数が1回であるために、ランダムに削除された非不正会計企業のサンプルによる影響を排除しきれない。

まとめると、前者は、構築した不正会計の検知モデルによって不正会計と非不正会計を判定するにあたり、事前不正確率を変化させたときに検知精度がどのように変化するのか、という問題であり、後者は、不正会計の検知モデルを構築するにあたり、アンダーサンプリングによって事前不正確率を変化させたときに検知精度がどのように変化するのか、という問題である。

そこで、本稿は、事前不正確率の違いによるロジスティック回帰を用いた不正会計検知モデルの検知精度の違いを明らかにすることを目的とする。具体的には、①Dechow et al. (2011) に従って、不正実施確率 (*Probability*) のカットオフ値に事前不正確率を加味した場合と加味しない場合で、その精度はどのように異なるのかを検討する。続いて、②構築した不正会計検知モデルの精度が、トレーニングデータの事前不正確率の違いによってどのように変化するのかを、Feroz et al. (2000) に倣って複数のパターンの事前不正確率を設定し、精度を比較する。

本稿の構成は次の通りである。第2節では、ロジスティック回帰モデルによって不正会計の検知モデルを構築したFanning and Cogger (1998)、Feroz et al. (2000)、Lin et al. (2003)、Kotsiantis et al. (2006)、Dechow et al. (2011)、首藤他 (2016)、佐藤他 (2022) をレビューする。第3節では、先行研究レビューを踏まえて、本稿のリサーチデザインを示す。第4節では、不正会計検知モデルの構築結果を示し、それぞれの検知モデルの精度比較を行い、結果の考察をする。第5節では本稿の要約、意義、課題を示して結びとする。

## 2. 先行研究

まず、Dechow et al. (2011) とDechow et al. (2011) をベースとして日本企業を対象にした不正会計検知モデルを構築した首藤他 (2016) と佐藤他 (2022) をレビューす

る。これらの研究では、不正会計企業と非不正会計企業の判別に、Fスコアを用いている。次に、不正会計企業と非不正会計企業の判別の際、不正実施確率 (*Probability*) のカットオフ値を0.500としているFanning and Cogger (1998)、Lin et al. (2003)、Kotsiantis et al. (2006) をレビューする。最後に、複数のパターンの事前不正確率を設定し、不正会計検知モデルの精度比較を行ったFeroz et al. (2000) をレビューする。

## 2.1. Dechow et al. (2011) の研究概要

Dechow et al. (2011) は、米国企業を対象に不正会計の検知モデルを構築した代表的な研究である。1982年から2005年までのAAER (Accounting and Auditing Enforcement Release) の中で企業名が記載されている2,190件のうち、利益捻出型の不正を行い、かつ不正会計検知モデルに使用するための財務データと非財務データを収集可能な354企業・年が不正会計企業として選択されている。非不正会計企業は、88,032企業・年である。

不正会計の検知モデルの構築にあたり、次の3つのステップを踏んでいる。1) 検知モデルに加える変数の選定、2) ロジスティック回帰モデルをベースとした不正会計検知モデルの構築、3) 構築した不正会計検知モデルの推定結果からFスコアと呼ばれる不正会計企業を判別する値の算出と検知モデルの性能評価のテストである。

まず1つ目のステップでは、①会計発生高の質に関する変数、②パフォーマンスに関する変数、③非財務情報に関する変数、④オフ・バランス・シートの変数、⑤市場に関する変数を、不正会計検知モデルの説明変数の候補として取り上げている。これらの候補となる計28変数について、統計的に有意で、かつ、その差の大小方向が先行研究の理論に従っている変数のみをロジスティック回帰モデルの推定に用いている。

2つ目のステップとして、選択された変数を基に、ロジスティック回帰モデルを推定している。不正会計検知モデル構築の際には、制約付きBackward方式で変数選択を行い、結果的には11の変数が選択された。

3つ目のステップでは、不正実施確率 (*Probability*)、すなわち各企業が不正会計を実施する確率を(1)式の通り算出する。(1)式において*Predicted Value*とは構築したモデルの係数を用いて推計される不正会計の程度を表す値である。

$$Probability = \frac{e^{(Predicted Value)}}{1 + e^{(Predicted Value)}} \quad (1)$$

次に(2)式により事前不正確率 (*Prior Fraud Probability*)<sup>3</sup> を、不正会計実施観測値 (企業・年) を全観測値 (企業・年) で除した値として計算する。Dechow et al. (2011) では、不正会計企業・年数の354を、非不正会計企業・年数の88,032で除して、0.004を事前不正確率として求めている。

3 Dechow et al. (2011) は、事前不正確率ではなく、不正会計実施の無条件確率 (*Prior Fraud Probability*) と表記しているが、その算出式では事前不正確率を算出している。そのため、本稿のレビューでは事前不正確率として表記している。

$$\text{Prior Fraud Probability} = \frac{\text{不正会計観測値}}{(\text{非不正会計観測値} + \text{不正会計観測値})} \quad (2)$$

さらに、不正会計実施確率を事前不正確率で除すことで(3)式の通りFスコア（*Fscore*）を算出する。

$$\text{Fscore} = \frac{\text{Probability}}{\text{Prior Fraud Probability}} \quad (3)$$

Fスコアのカットオフ値を1.00とした場合の不正会計検知モデルの検知精度を評価するために、Correct Classification（正答率）、Sensitivity（真陽性率）、Type I Error（第1種の過誤）、Type II Error（第2種の過誤）の4つの指標を用いている。結果として、Correct Classificationは63.74%、Sensitivityは67.23%、Type I Errorは36.28%、Type II Errorは32.77%という結果であった。

## 2.2. 首藤他（2016）と佐藤他（2022）の研究概要

首藤他（2016）では、Dechow et al.（2011）に倣い、2000年から2014年の日本の上場企業を分析対象とした不正会計の検知モデルを構築している。首藤他（2016）はDechow et al.（2011）と次の3点が異なっている。第1に、会計発生高の質に関する変数として裁量的発生高（Discretionary Accruals）を利用している点である。第2に、実体的裁量行動にも注目し、Roychowdhury（2006）に基づく実体的裁量行動に関する変数について分析を行っている点である。第3に、保守主義の程度を表す変数と日本特有の要因を考慮した変数について分析を行っている点である。

構築したモデルの検知精度に係る検証として、Fスコアのカットオフ値を1.00とした場合の検知モデルの検知精度を評価したところ、不正会計実施観測値229観測値のうち212観測値が正しく分類されている。検知精度は、Correct Classificationは76.34%、Sensitivityは92.58%、Type I Errorは23.74%、Type II Errorは7.42%という結果であった。

佐藤他（2022）は2011年から2017年の日本の上場企業を対象にして、ロジスティック回帰を用いた不正会計検知モデルを構築した研究である。佐藤他（2022）が、Dechow et al.（2011）と首藤他（2016）と異なるのは、利益捻出型の不正会計を不正会計のタイプ別に分類し、そのタイプ別に不正会計検知モデルを構築している点である。具体的には、収益に関する不正、費用に関する不正、資産に関する不正、負債に関する不正を不正会計のタイプとして、タイプ別に不正会計検知モデルを構築した。さらに、トレーニングデータを用いた学習にはLasso回帰を用いた変数選択が行われている点も異なっている。

不正会計検知モデルの検知精度に係る検証では、Fスコアのカットオフ値を1.00とした場合の検知モデルの検知精度をCorrect Classification、Sensitivity、Type I Error、



Type II Errorで評価している。利益捻出型の不正会計を対象にしたモデルをベンチマークモデルとし、不正会計のタイプ別に構築した計4個のモデルを比較すると、負債に関する不正会計を対象にしたModel 5のCorrect Classificationが最も高い結果(66.99%)となっている。また、Sensitivityを見ると、収益に関する不正を対象にしたModel 2(51.02%)が他のモデルよりも不正会計企業を正しく分類できていることが明らかになった。

### 2.3. Fanning and Cogger (1998)、Lin et al. (2003)、Kotsiantis et al. (2006)の研究概要

Fanning and Cogger (1998) は、1980年から1993年の米国企業を対象に、ロジスティック回帰を用いた不正会計検知モデルを構築している。不正会計検知モデルを構築するにあたり、3つのステップを踏んでいる。1) 検知モデルに加える変数の選定、2) ロジスティック回帰をベースとした不正会計検知モデルの構築、3) 不正会計検知モデルの性能評価である。

1つ目のステップでは、①コーポレート・ガバナンスに関する変数、②会計監査人に関する変数、③企業と経営者(取締役)とのエージェンシー問題に関する変数、④子会社に関する変数、⑤人事異動を含む人的要因に関する変数、⑥財務諸表の勘定科目と比率を不正会計検知モデルに投入する変数の候補に挙げ、計62変数について単変量解析を行い、有意な変数(計20変数)を不正会計検知モデルに投入している。2つ目のステップとして、トレーニングデータを用いてロジスティック回帰を用いた不正会計検知モデルを構築している。この際、ステップワイズ法により変数選択が行われ、8変数が選択されている。3つ目のステップとして、2つ目のステップで構築した不正会計検知モデルをテストデータに適用し、推定された不正実施確率(*Probability*)のカットオフ値を0.500とした場合のモデルの精度指標を算出している。モデルの精度は、Correct Classification、Sensitivity、非不正会計企業の的中率<sup>4</sup>を使用しており、結果として、Correct Classificationが50%、Sensitivityが67%、非不正会計企業の的中率が33%であった。

Lin et al. (2003) は、1980年から1995年の米国企業を対象に、ロジスティック回帰を用いて、不正会計検知モデルを構築している。不正会計検知モデルの構築手順は、Fanning and Cogger (1998)と同様であり、売上高、貸倒引当金、売掛金に関わる変数を説明変数とした不正会計検知モデルを構築した。結果として、不正実施確率(*Probability*)のカットオフ値を0.500とした場合の検知精度は、Correct Classificationが79.0%、Sensitivityが5.0%、非不正会計企業の的中率が97.5%であった。

Kotsiantis et al. (2006) は、2001年から2002年の間に、アテネ証券取引所(Athens Stock Exchange)に上場しているギリシャの製造業企業を対象にして、ロジスティッ

4 非不正会計企業の的中率は、 $\text{True Negative} / (\text{True Negative} + \text{False Positive})$ で求める。すなわち、不正会計検知モデルで正しく非不正会計企業と判別された観測値を非不正会計企業と判別された観測値で除したものである。

ク回帰による不正会計検知モデルを構築している。検知モデルの説明変数の候補である、①収益性に関する変数、②レバレッジに関する変数、③流動性に関する変数、④効率性に関する変数、⑤キャッシュ・フローに関する変数、⑥倒産確率の変数の計25変数について、統計尺度であるReliefFに従って各変数の影響度をランキングし、上位8変数を用いて不正会計検知モデルに投入している。

続いて、層別10分割交差検証により、その結果を平均した値で不正会計検知モデルの精度を評価している。推定された不正実施確率（*Probability*）のカットオフ値を0.500とした場合のモデルの精度指標の結果は、Correct Classificationは75.3%、Sensitivityは36.6%、非不正会計企業の的中率は88.9%であった。

#### 2.4. Feroz et al. (2000) の研究概要

Feroz et al. (2000) は、米国企業を対象にして、ロジスティック回帰モデルによる不正会計検知モデルの結果をレビューする。不正会計企業として、SECが報告したAAER (No.1 ~ No.309) のうち、CompustatのデータとMoody's Manualsの非財務データが入手可能な42社を選択している。そして、非不正会計企業として、不正会計企業と同じ産業に属している同規模の企業である90社を選択している。

不正会計検知モデルの構築にあたり、3つのステップを踏んでいる。1) 検知モデルに加える変数の選定、2) ロジスティック回帰をベースとした不正会計検知モデルの構築と、3) 不正会計検知モデルの性能評価である。

1つ目のステップでは、①財務に関する変数と②人事異動に関する変数を挙げている。①に該当する4つの変数として、収益性、金利感応度、監査が困難な取引、ゴーイングコンサーンを取り上げている<sup>5</sup>。これらの4つの変数については、前期から当期にかけての変化率も併せて算出している。②に該当する変数として、CEOの交代回数、CFOの交代回数、監査役の交代回数を用いている。

2つ目のステップとして、不正会計企業と非不正会計企業のサンプルサイズをコントロールし、不正会計企業と非不正会計企業の比率を変えて、不正会計検知モデルの構築を行っている。まず、不正会計企業と非不正会計企業を全体のデータ数からランダムに3分の2（不正会計企業：28社、非不正会計企業：60社）を抽出し、それをトレーニングデータとした。残りの3分の1（不正会計企業：14社、非不正会計企業：30社）をテストデータとした。次に、非不正会計企業に対する不正会計企業の割合<sup>6</sup>を4つのパターン（10%、20%、33%、50%）に分け、これらをトレーニングデータに適用している。具体的には、不正会計企業の割合10%のトレーニングデータとして、不正会計企業6社、非不正会計企業60社、20%のトレーニングデータとして、不正会計企業12社、非不正会

5 Feroz et al.(2000)において監査が困難な取引は、売上高に占める売掛金の割合で測定されている。ゴーイングコンサーンは、アルトマンのZスコアで測定されている。

6 ここでの非不正会計企業に対する不正会計企業の割合は、分析サンプルにおける非不正会計企業と不正会計企業の割合を意味しており、母集団における不正会計企業の割合を意味する事前不正確率とは厳密には意味が異なっている。

計企業60社、33%のトレーニングデータとして、不正会計企業20社、非不正会計企業60社、50%のトレーニングデータとして、不正会計企業28社、非不正会計企業56社を用いている。そして、これらの4つのトレーニングデータを用いて、不正会計検知モデルを構築している。なお、不正会計検知モデルの構築の際、変数選択は行われていない。

3つ目のステップとして、2つ目のステップで構築した4つのモデルをテストデータで検証し、検知精度を計算している。テストデータに対しても、不正会計企業の割合について4つのパターン（10%、20%、33%、50%）を設定している。不正実施確率（*Probability*）が0.500以上の場合を不正会計企業、0.500未満の場合を非不正会計企業と判定し、トレーニングデータの4パターンとテストデータの4パターンの組み合わせである計16パターンのCorrect Classificationを算出し、その値を比較している。

16パターンの検知精度の結果は表1に示す通りである。不正会計検知モデルのCorrect Classificationが51.9%から88.0%の範囲である。これらの結果はトレーニングデータとテストデータの組み合わせに影響されている。例えば、トレーニングデータが50%、テストデータが50%の組み合わせにおけるCorrect Classificationは51.9%と最も低い精度となっている。一方で、トレーニングデータが10%、テストデータが10%の組み合わせにおけるCorrect Classificationは最も高い結果88.0%となっている。他の組み合わせについても見ると、トレーニングデータとテストデータの非不正会計企業に対する不正会計企業に割合が小さい組み合わせになるにつれ（表1の左上になるにつれ）、検知精度が向上していることが確認できる。

表1：Feroz et al. (2000) のモデルの評価

テスト \ トレーニング		10%	20%	33%	50%
	10%	88.0%	82.8%	79.0%	73.3%
	20%	81.0%	77.0%	71.7%	65.3%
	33%	71.4%	68.3%	63.7%	62.5%
	50%	64.1%	63.9%	66.0%	51.9%

### 3. リサーチデザイン

#### 3.1. データ

本稿では2011年から2019年<sup>7</sup>を対象として、東京証券取引所（市場第一部、市場第二部、マザーズ）に上場する企業（銀行・証券・保険・その他金融の業種を除く）の中で、日本の会計基準を採用している不正会計企業と非不正会計企業のデータを用いる。これら

7 なお、過去3年間の値の変動に関わる変数を作成する際には、2008年～2010年のデータも併せて使用している。



の企業における財務・非財務データと株価データは、「日経NEEDS-FinancialQUEST」と「eol」から収集している。

表2に年別に不正会計企業・非不正会計企業の観測値と不正会計企業割合を示す。2011年から2019年までの非不正会計企業の総サンプルは、18,627観測値（企業・年）である。これに対し、不正会計企業のサンプルは、東京商工リサーチから取得した「不適切会計リスト」にある全ての不正会計企業について有価証券報告書の訂正報告書を確認し、財務諸表の値に修正があったケースのみを選択した<sup>8</sup>。財務諸表の値に修正があったケースの中には、訂正報告書によって税引後当期純利益が訂正前よりも減額修正されたケース（利益捻出型の不正、65.88%）、増額修正されたケース（利益圧縮型の不正、24.41%）、修正が無かったケース（9.71%）が存在している。本稿では、この内、利益捻出型の224観測値（企業・年）を対象とする。結果として、フルサンプルにおける事前不正確率は1.19%である。

表2：年別の不正と非不正のデータ数

年	不正会計 企業	非不正会計 企業	企業 合計	不正会計 企業割合
2011年	21	1,936	1,957	1.07%
2012年	26	1,975	2,001	1.30%
2013年	26	1,991	2,017	1.29%
2014年	38	2,031	2,069	1.84%
2015年	36	2,072	2,108	1.71%
2016年	31	2,120	2,151	1.44%
2017年	27	2,177	2,204	1.23%
2018年	13	2,263	2,276	0.57%
2019年	6	2,062	2,068	0.29%
計	224	18,627	18,851	1.19%

### 3.2. 説明変数

ロジスティック回帰モデルの説明変数は、Dechow et al. (2011) と首藤他 (2016) に倣って、①会計発生高の質に関する変数、②パフォーマンスに関する変数、③非財務情報に関する変数、④オフ・バランス・シートの変数、⑤市場に関する変数、⑥保守主義に関する変数、⑦実体的裁量行動に関する変数、⑧日本特有の要因に関する変数とし

8 適切会計リストに記載されている企業でも、横領目的による不正で財務諸表への影響が僅少だった場合や、有価証券報告書を提出する前に倒産・上場廃止した場合には、必ずしも有価証券報告書の訂正報告書において財務諸表の修正が開示されていない。

て、計38変数を用いる。表3に、目的変数と38の説明変数について、各変数の変数名、説明変数の区分、測定方法を示している。

表3：説明変数の一覧

変数名	測定方法
目的変数: 不正会計企業ダミー	不正会計企業(企業・年)であれば1、そうでなければ0を示すダミー変数
①会計発生高の質に関する変数	
<i>WC accruals</i>	$((\Delta \text{流動資産} - \Delta \text{現金預金及び有価証券等}) - (\Delta \text{流動負債} - (\Delta \text{短期借入金} + \Delta 1 \text{年以内返済予定長期借入金} + \Delta 1 \text{年以内償還予定社債} + \Delta \text{コマーシャル・ペーパー} + \Delta 1 \text{年以内返済リース債務}) - \Delta \text{未払税務})) / \text{総資産の期中平均値}$
<i>RSST accruals</i>	$RSST = (\Delta WC + \Delta NCO + \Delta FIN) / \text{総資産の期中平均値}$ $WC = (\text{流動資産} - \text{現金預金及び有価証券等}) - (\text{流動負債} - \text{短期有利子負債}),$ $NCO = (\text{総資産} - \text{流動資産} - \text{投資その他の資産}) -$ $(\text{負債合計} - \text{流動負債} - \text{長期有利子負債}),$ $FIN = (\text{有価証券等} + \text{投資その他の資産}) - (\text{長期有利子負債} + \text{短期有利子負債})$
<i>Change in receivables</i>	$(\Delta \text{売上債権} + \Delta \text{未収入金}) / \text{総資産の期中平均値}$
<i>Change in inventory</i>	$\Delta \text{棚卸資産} / \text{総資産の期中平均}$
<i>% Soft assets</i>	$(\text{総資産} - \text{有形固定資産} - \text{現金預金及び有価証券等}) / \text{総資産}$
<i>MJ discretionary accruals</i>	Dechow et al. (1995)に基づく裁量的発生高 $WC \text{ accrualst} = \alpha + \beta_1 (\Delta \text{Sales} - \Delta \text{Receivablest}) + \beta_2 \text{PPEt} + \epsilon t$ $WC \text{ accruals} = \text{運転資本発生高}$ $Sales = \text{売上高}$ $Receivables = \text{売上債権} + \text{未収入金}$ $PPE = \text{有形固定資産}$
<i>MJ discretionary accruals SD</i>	<i>MJ discretionary accruals</i> の過去3年間の標準偏差
<i>MJ discretionary accruals AB</i>	<i>MJ discretionary accruals</i> の絶対値
<i>PM discretionary accruals</i>	Kothari et al. (2005)に基づく裁量的発生高 $WC \text{ accrualst} = \alpha + \beta_1 (\Delta \text{Sales} - \Delta \text{Receivablest}) + \beta_2 \text{PPEt} + \beta_3 \text{ROA} + \epsilon t$ $ROA = \text{税引後当期純利益}$
<i>PM discretionary accruals SD</i>	<i>PM discretionary accruals</i> の過去3年間の標準偏差
<i>PM discretionary accruals AB</i>	<i>PM discretionary accruals</i> の絶対値
<i>CFO discretionary accruals</i>	Kasznik (1999)に基づく裁量的発生高 $WC \text{ accrualst} = \alpha + \beta_1 (\Delta \text{Sales} - \Delta \text{Receivablest}) + \beta_2 \text{PPEt} + \beta_3 \Delta \text{CFO} + \epsilon t$ $Total \text{ accruals} = \text{会計発生高 (税引後経常利益} - \text{営業活動によるキャッシュ・フロー)}$ $CFO = \text{営業活動によるキャッシュ・フロー}$
<i>CFO discretionary accruals SD</i>	<i>CFO discretionary accruals</i> の過去3年間の標準偏差
<i>CFO discretionary accruals AB</i>	<i>CFO discretionary accruals</i> の絶対値
②パフォーマンスに関する変数	
<i>Change in cash sales</i>	現金売上高の変化率 = $\text{現金売上高} / \text{前期現金売上高} - 1$ 現金売上高 = $\text{売上高} - \Delta \text{貸倒引当金控除前売上債権及び未収入金}$
<i>Change in cash margin</i>	キャッシュ・フロー・マージンの変化率 = $\text{キャッシュ・フロー・マージン} / \text{前期キャッシュ・フロー・マージン} - 1$ $\text{キャッシュ・フロー・マージン} = 1 - (\text{売上原価} - \Delta \text{棚卸資産} + \Delta \text{買掛金}) / (\text{売上高} - \Delta \text{貸倒引当金控除前売上債権及び未収入金})$
<i>Change in free cash flows</i>	フリー・キャッシュ・フローの変化額 = $(\text{フリー・キャッシュ・フロー} - \text{前期フリー・キャッシュ・フロー}) / \text{総資産の期中平均}$ フリー・キャッシュ・フロー = $\text{税引後当期純利益} - RSST \text{ accruals}$
<i>Change in return on assets</i>	$ROA$ の変化額 = $\text{当期} ROA - \text{前期} ROA$ $ROA = \text{税引後当期純利益} / \text{総資産の期中平均}$
<i>Deferred tax expense</i>	法人税等調整額 / 総資産の期中平均

表3：説明変数の一覧（続き）

変数名	測定方法
③非財務情報に関する変数	
<i>Abnormal change in employees</i>	(期末従業員数 / 前期末従業員数 - 1) - (総資産 / 前期末総資産 - 1)
④オフ・バランス・シートの変数	
<i>Existence of operating leases</i>	オペレーティング・リース未経過リース料 > 0 なら 1, それ以外は 0 を示すダミー変数
⑤市場に関する変数	
<i>Ex ante financing need</i>	(営業キャッシュ・フロー - 過去3年の平均設備投資額) / 流動資産 < -0.5 なら 1, それ以外は 0 を示すダミー変数
<i>Actual issuance</i>	普通株式及び優先株式の売却 > 0 または有利子負債の増加 > 0 のとき 1, それ以外は 0 を示すダミー変数
<i>CFF</i>	財務活動によるキャッシュ・フロー / 総資産の期中平均
<i>Leverage</i>	長期有利子負債 / 総資産
<i>Market adjusted stock-return</i>	(当期末時価総額 - 前期末時価総額) / 前期末時価総額 - (当期末 <i>TOPIX</i> - 前期末 <i>TOPIX</i> ) / 前期末 <i>TOPIX</i>
<i>Lagged market adjusted</i>	1期前の <i>Market adjusted stock - return</i> の値
<i>Book-to-market</i>	自己資本 / 時価総額
<i>Earnings-to-price</i>	税引後当期純利益 / 時価総額
⑥保守主義に関する変数	
<i>C score RR</i>	Khan and Watts (2009) に基づく会計保守主義スコア (Cscore). 推計モデルにロウ・リターンを利用する. $X_t = \beta_1 + \beta_2 D_t + \beta_3 R_t + \beta_4 D_t R_t + \varepsilon_t$ $X = \text{税引後当期純利益} / \text{期首時価総額}$ $R = \text{株価の年率リターン}$ $D = R \text{ が負のとき } 1, \text{ それ以外は } 0 \text{ とするダミー変数}$ $Gscore = \beta_3 = \mu_1 + \mu_2 Size + \mu_3 M/B_t + \mu_4 Lev_t$ $Cscore = \beta_4 = \lambda_1 + \lambda_2 Size + \lambda_3 M/B_t + \lambda_4 Lev_t$
<i>C score RR SD</i>	<i>C score RR</i> の過去3年の標準偏差
<i>C score AR</i>	Khan and Watts (2009) に基づく会計保守主義スコア (Cscore). 推計モデルに市場調整済みリターンを利用する. $X_t = \beta_1 + \beta_2 D_t + \beta_3 R_t + \beta_4 D_t R_t + \varepsilon_t$ $X = \text{税引後当期純利益} / \text{期首時価総額}$ $R = \text{株価の年率リターン}$ $D = R \text{ が負のとき } 1, \text{ それ以外は } 0 \text{ とするダミー変数}$ $Gscore = \beta_3 = \mu_1 + \mu_2 Size + \mu_3 M/B_t + \mu_4 Lev_t$ $Cscore = \beta_4 = \lambda_1 + \lambda_2 Size + \lambda_3 M/B_t + \lambda_4 Lev_t$
<i>C score AR SD</i>	<i>C score AR</i> の過去3年の標準偏差
⑦実体的裁量行動に関する変数	
<i>AB cash flow</i>	Roychowdhury (2006) に基づく異常営業キャッシュ・フロー $CFO_t / A_{t-1} = \alpha + \beta_1 (I_t / A_{t-1}) + \beta_2 (Sales_t / A_{t-1}) + \beta_3 (\Delta Sales_t / A_{t-1}) + \varepsilon_t$ $CFO = \text{営業活動によるキャッシュ・フロー}$ $A = \text{総資産}$ $Sales = \text{売上高}$
<i>Discretionary expense</i>	Roychowdhury (2006) に基づく異常裁量費用 $DE_t / A_{t-1} = \alpha + \beta_1 (I_t / A_{t-1}) + \beta_2 (Sales_t / A_{t-1}) + \varepsilon_t$ $DE = \text{裁量的費用 (研究開発費 + 広告宣伝費 + 払販費・その他販売費 + 役員報酬・賞与 + 人件費・福利厚生費)}$
<i>AB product cost</i>	Roychowdhury (2006) に基づく異常製造原価 $PD_t / A_{t-1} = \alpha + \beta_1 (I_t / A_{t-1}) + \beta_2 (Sales_t / A_{t-1}) + \beta_3 (\Delta Sales_t / A_{t-1}) + \beta_4 (\Delta Sales_{t-1} / A_{t-1}) + \varepsilon_t$ $PD = \text{製造原価 (売上原価 + 期末棚卸資産 - 期首棚卸資産)}$
⑧日本特有の要因に関する変数	
<i>FIN</i>	金融機関所有株数 / 期末発行済株式総数
<i>CORP</i>	その他法人所有株数 / 期末発行済株式総数

※非不正会計企業の説明変数に関して、各説明変数の異常値の影響を軽減するために、上下1%の範囲でウインザライズ処理を行っている。

### 3.3. 不正会計検知モデルの精度の指標

モデルの評価指標は、先行研究で用いられているCorrect Classification、Sensitivity、Type I Error、Type II Errorに対して、Precision（適合率）を新たに加え、5つの指標を用いる（表4）。

表4：評価指標

推定 実測	不正会計企業	非不正会計企業
不正会計企業	True Positive (TP)	False Negative (FN)
非不正会計企業	False Positive (FP)	True Negative (TN)

$$\text{Correct Classification} = (\text{TP} + \text{TN}) / (\text{TP} + \text{FN} + \text{FP} + \text{TN})$$

$$\text{Sensitivity} = (\text{TP}) / (\text{TP} + \text{FN})$$

$$\text{The rate of Type I Error} = (\text{FP}) / (\text{TN} + \text{FP})$$

$$\text{The rate of Type II Error} = (\text{FN}) / (\text{TP} + \text{FN})$$

$$\text{Precision} = (\text{TP}) / (\text{TP} + \text{FP})$$

### 3.4. カットオフ値における事前不正確率の違いと検知精度の変化の検証手順

第1のステップでは、データをトレーニングデータとテストデータに分割する。2011年から2016年までのデータをトレーニングデータ、2017年から2019年までのデータをテストデータとする。トレーニングデータにおける事前不正確率は1.45%である。

第2のステップでは、トレーニングデータを使用して、ロジスティック回帰モデルに基づく不正会計検知モデルを構築する。その際の変数選択には、Lasso回帰を用いる。

第3のステップでは、構築された不正会計検知モデルを用いて、テストデータに対する推定を行う。Fスコアのカットオフ値を0.00から34.56までの範囲における精度を比較する。その理由は、不正実施確率（*Probability*）のカットオフ値に事前不正確率を加味した場合、すなわちFスコアの値を変化させた場合に、その精度はどのように異なるのかを明らかにすると共に、不正実施確率（*Probability*）のカットオフ値を0.500とした場合の精度と併せて比較を行うためである。Fスコアのカットオフ値の0.00から34.56は、不正実施確率（*Probability*）の0.000から0.500に対応している。

### 3.5. トレーニングデータにおける事前不正確率の違いと検知精度の変化の検証手順

第1のステップでは、データをトレーニングデータとテストデータに分割する。2011年から2019年までのデータを層別に10分割し、9割のデータをトレーニングデータ、1割のデータをテストデータとする（層別10分割交差検証法）。10分割交差検証法を用いる理由は、アンダーサンプリングによってトレーニングデータの事前不正確率を変化さ

せた場合に、どのように精度が異なるのかを明らかにする上で、ホールドアウト法よりも10分割交差検証法を用いる方が頑健性の高い結果を得ることができると考えられるためである。層別にデータを分割している理由は、トレーニングデータとテストデータに含まれる不正会計企業と非不正会計企業の比率を一定に保つ必要があるためである（Perols et al., 2017）。

第2のステップでは、トレーニングデータの事前不正確率が2.5%、5%、10%、20%、33%、50%となるように非不正会計企業のサンプルをアンダーサンプリングする。

第3のステップでは、アンダーサンプリングされたトレーニングデータを使用して、ロジスティック回帰モデルに基づく不正会計検知モデルを構築する。その際の変数選択には、Lasso回帰を用いる。

第4のステップでは、構築された不正会計検知モデルによってテストデータに対する推定を行い、その精度の平均値を求め、精度比較を行う。カットオフ値は、トレーニングデータの事前不正確率に基づいて算出するFスコアの値が1.00とする。例えば、事前不正確率が2.5%となるようにアンダーサンプリングされた10個のトレーニングデータセットを使用して不正会計検知モデルの構築を行い、そのモデルを用いてテストデータに対する推定を行うと、Correct Classification、Sensitivity、Type I Error、Type II Error、Precisionのそれぞれについて、10個の値を得ることができる。これを平均化して、事前不正確率が2.5%の場合の検知モデルの精度とする。

## 4. 結果と考察

### 4.1. カットオフ値における事前不正確率の違いと検知精度の変化の検証結果

2011年から2016年までのトレーニングデータに対してロジスティック回帰（Lasso回帰による変数選択）によって構築されたモデルは、(4)式の通りである。モデル構築のために投入した計38変数のうち、計20変数が説明変数として採択されている。

$$\begin{aligned}
 Y(0,1) = & -5.869 - 2.347 \times (\text{Change in inventory}) + 1.887 \times (\% \text{ Soft assets}) \\
 & + 0.021 \times (\text{MJ discretionary accruals AB}) + 2.791 \times (\text{PM discretionary accruals AB}) \\
 & + 0.931 \times (\text{CFO discretionary accruals}) + 0.046 \times (\text{Change in cash sales}) \\
 & - 0.082 \times (\text{Change in cash margin}) + 1.704 \times (\text{Change in return on assets}) \\
 & + 0.562 \times (\text{Abnormal change in employees}) - 0.043 \times (\text{Existence of operating leases}) \\
 & + 0.132 \times (\text{Ex ante financing need}) + 0.096 \times (\text{Actual issuance}) + 1.749 \times (\text{Leverage}) \\
 & + 0.146 \times (\text{Book-to-market}) + 0.274 \times (\text{Earnings-to-price}) + 0.207 \times (\text{C score RR}) \\
 & + 0.284 \times (\text{C score AR}) - 2.093 \times (\text{AB cash flow}) \\
 & - 0.668 \times (\text{AB product cost}) + 4.952 \times (\text{CORP})
 \end{aligned} \tag{4}$$



$Y(0, 1)$  : 不正会計企業 (企業・年) であれば 1、そうでなければ 0 を示すダミー変数

*Change in inventory* : 棚卸資産の変化額

% *Soft assets* : ソフト資産比率

*MJ discretionary accruals AB* : Dechow et al. (1995) に基づく裁量的発生高の絶対値

*PM discretionary accruals AB* : Kothari et al. (2005) に基づく裁量的発生高の絶対値

*CFO discretionary accruals* : Kasznik (1999) に基づく裁量的発生高

*Change in cash sales* : 現金売上高の変化率

*Change in cash margin* : キャッシュ・フロー・マージンの変化率

*Change in return on assets* : ROA の変化額

*Abnormal change in employees* : 従業員の変化率

*Existence of operating leases* : オペレーティング・リース未経過リース料に関するダミー変数

*Ex ante financing need* : 事前の資金調達必要性

*Actual issuance* : 株式と負債に関する資金調達を行っているか否かを示すダミー変数

*Leverage* : レバレッジ

*Book-to-market* : 簿価対時価比率

*Earnings-to-price* : 時価総額に対する税引後当期純利益の割合

*C score RR* : Khan and watts (2009) に基づく保守主義。推計にはロウ・リターンを利用する。

*C score AR* : Khan and watts (2009) に基づく保守主義。推計には市場調整済みリターンを利用する。

*AB cash flow* : Roychowdhury (2006) に基づく異常営業キャッシュ・フロー

*AB product cost* : Roychowdhury (2006) に基づく異常製造原価

*CORP* : 一般事業法人持株比率

この(4)式の説明変数とその係数の推定値をもって、2017年から2019年のテストデータに対する推定を行い、不正実施確率 (*Probability*) を算出した。この後、不正実施確率 (*Probability*) を事前不正確率の1.45%で除してFスコアを算出し、Fスコアが0.25から35.00までの範囲における検知精度としてCorrect Classification、Sensitivity、Type I Error、Type II Error、Precisionの値を表5に示している。

表5 カットオフ値における事前不正確率と検知精度

Fスコア	0.25	0.50	0.75	<b>1.00</b>	1.25	<b>1.50</b>	1.75	2.00
不正実施確率(Probability)	0.0036	0.0072	0.0109	0.0145	0.0181	0.0217	0.0253	0.0289
Correct Classification	1.16%	12.28%	38.85%	<b>65.93%</b>	81.63%	<b>90.27%</b>	94.52%	96.35%
Sensitivity	100.00%	91.30%	67.39%	<b>43.48%</b>	26.09%	<b>21.74%</b>	10.87%	8.70%
Type I Error	99.54%	<b>88.28%</b>	61.35%	<b>33.91%</b>	17.98%	<b>9.24%</b>	4.89%	3.03%
Type II Error	0.00%	8.70%	32.61%	<b>56.52%</b>	73.91%	<b>78.26%</b>	89.13%	91.30%
Precision	0.71%	0.73%	0.77%	<b>0.90%</b>	1.02%	<b>1.64%</b>	1.55%	1.99%
Fスコア	3.00	4.00	5.00	6.00	7.00	8.00	9.00	10.00
不正実施確率(Probability)	0.0434	0.0579	0.0723	0.0868	0.1013	0.1157	0.1302	0.1447
Correct Classification	98.61%	99.10%	99.21%	99.24%	99.25%	99.27%	99.27%	99.27%
Sensitivity	4.35%	2.17%	2.17%	2.17%	2.17%	2.17%	0.00%	0.00%
Type I Error	0.72%	0.22%	0.11%	0.08%	0.06%	0.05%	0.03%	0.03%
Type II Error	95.65%	97.83%	97.83%	97.83%	97.83%	97.83%	100.00%	100.00%
Precision	4.08%	6.67%	12.50%	16.67%	20.00%	25.00%	0.00%	0.00%

表5 カットオフ値における事前不正確率と検知精度（続き）

Fスコア	11.00	12.00	13.00	14.00	15.00	25.00	34.56	35.00
不正実施確率(Probability)	0.1591	0.1736	0.1881	0.2026	0.2170	0.3617	<b>0.5000</b>	0.5064
Correct Classification	99.28%	99.28%	99.30%	99.30%	99.30%	99.30%	<b>99.30%</b>	99.30%
Sensitivity	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	<b>0.00%</b>	0.00%
Type I Error	0.02%	0.02%	0.00%	0.00%	0.00%	0.00%	<b>0.00%</b>	0.00%
Type II Error	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	<b>100.00%</b>	100.00%
Precision	0.00%	0.00%	-	-	-	-	-	-

表5を見ると、不正会計検知モデルの全体の正しい判別率を表すCorrect Classificationは、Fスコアが1.00のときに65.93%である。Fスコアが高まるにつれてCorrect Classificationも改善するという関係になっているが、Fスコアが13以上におけるCorrect Classificationの値は、99.30%からは改善されないという結果となっている。これは、Fスコアが13以上では、実際に不正を実施した企業を推定不正会計企業として判別しない（Type II Errorが100.00%）ことに起因している。

不正会計企業に対する評価指標であるSensitivityは、Fスコアが1.00のときに43.48%である。これは、不正会計企業のうち、43.48%を構築した不正会計検知モデルで正しく不正会計企業として判別できていることを示している。SensitivityはFスコアが高まるにつれて減少する傾向であり、Fスコアが9以上ではSensitivityは0.00%となる。すなわち、不正会計企業を、検知モデルで不正会計企業としてより多く判別したいと考える場合には、事前不正確率によるカットオフ値を低く設定（すなわち、Fスコアを低い値とする）必要がある。例えば、Fスコアを0.50とした場合には不正会計企業のうち、91.30%を不正会計企業として判別することができる。

Type I Errorは推定不正会計企業における誤判別の割合（非不正会計企業を誤って不正会計企業と推定した割合）を、Type II Errorは推定非不正会計企業における誤判別の割合（不正会計企業を誤って非不正会計と推定した割合）を示している。表10において、このエラーの合計が最小となるのは、Fスコアが1.50の場合であった。Fスコアが1.50を不正会計企業と非不正会計企業の判定に用いると、Correct Classificationは90.27%、Type I ErrorとType II Errorの合計が87.50%であり、良好な精度と考えられる。

Fスコアが1.00、1.50、34.56を基準とした場合の不正会計企業・非不正会計企業、推定不正会計企業・推定非不正会計企業の件数と、精度に関する5つの指標の結果を表6に示す。この3つのケースを比較することで、トレーニングデータにおける事前不正確率で不正会計企業と非不正会計企業の判別を行う場合（Fスコアが1.00）と、Type I ErrorとType II Errorの合計が最小となる場合（Fスコアが1.50）と、不正実施確率（Probability）が0.50の場合（Fスコアが34.56）の精度を比較することができる。

表6を見ると、Fスコアが1.00の場合と34.56の場合では、Fスコアが34.56の場合のほうがCorrect Classificationが高まっている（Fスコアが1.00の場合のCorrect

表6 Fスコアが1.00、1.50、34.56をカットオフ値とした場合の検知精度

	Fスコア : 1.00			Fスコア : 1.50			Fスコア : 34.56		
	不正実施確率 : 0.0145			不正実施確率 : 0.0217			不正実施確率 : 0.5000		
	推定不正 会計企業	推定非不正 会計企業	計	推定不正 会計企業	推定非不正 会計企業	計	推定不正 会計企業	推定非不正 会計企業	計
不正会計企業	20	26	46	10	36	46	0	46	46
非不正会計企業	2,205	4,297	6,502	601	5,901	6,502	0	6,502	6,502
計	2,225	4,323	6,548	611	5,937	6,548	0	6,548	6,548
Correct Classification		65.93%			90.27%			99.30%	
Sensitivity		43.48%			21.74%			0.00%	
Type I Error		33.91%			9.24%			0.00%	
Type II Error		56.52%			78.26%			100.00%	
Precision		0.90%			1.64%			-	

Classification = 65.93%、Fスコアが34.56の場合のCorrect Classification = 99.30%)。しかし、Sensitivityを見ると、Fスコアが34.56の場合ではSensitivityは0.00%と、不正会計企業を1件も不正会計企業として判定することができていないことが明らかになっている。

Fanning and Cogger (1998)、Lin et al. (2003)、Kotsiantis et al. (2006) 等は、不正会計企業の検知モデルを構築するにあたり、ロジスティック回帰とその他の機械学習手法の精度を比較し、機械学習手法のほうが精度が高いと結論している。しかし、これらの研究では、ロジスティック回帰に基づく不正会計の検知モデルによって不正会計企業と非不正会計企業を判定する際に、不正実施確率 (*Probability*) を0.500としており、これは本稿におけるFスコアが34.56と同じ意味である。本稿の結果からは、Fスコアが34.56、すなわち不正実施確率 (*Probability*) が0.500をカットオフ値としてしまうと、不正会計企業を正しく判定することができていないという結果であり、これが先行研究においてロジスティック回帰と機械学習の精度を比較した際に、ロジスティック回帰による不正会計の検知モデルの精度が低いという結論に繋がった要因の1つと考えられる。従って、Fスコアが1.00や1.50等をカットオフ値として、機械学習手法との精度比較を行うことが求められるだろう。

続いて、Fスコアが1.00の場合とFスコアが1.50の場合のPrecisionの結果を見る。推定不正会計企業における正しい判別率を示すPrecisionは、Fスコアが1.00の場合は0.90%であるのに対し、Fスコアが1.50の場合は1.64%であり、Fスコアが1.50の場合のほうが良い値である。Fスコアが1.00の場合は、Fスコアが1.50の場合に比較してSensitivityが高く (Fスコアが1.00の場合のSensitivity = 43.48%、Fスコアが1.50の場合のSensitivity = 21.74%)、不正会計企業を見逃してしまう可能性が低いが、公認会計士や監査法人が、構築された不正会計検知モデルで推定された結果に基づいて2,205社に対する追加的な監査を行ったとしても、それが無駄になってしまうことを示している。Sensitivityの精度をどの程度高めたいか、反対にPrecisionの精度の悪化をどの程度まで許容するかは、個々の公認会計士や監査法人の判断によると考えられるため、不正会計実施企業を見逃

すコストと、非不正会計企業に対して追加的な監査を行うコストを比較・検討する必要があるだろう。これは、Type I ErrorとType II Errorの重みづけをどの程度とするかという課題に繋がる。

#### 4.2. トレーニングデータにおける事前不正確率の違いと検知精度の変化の検証結果

次に、トレーニングデータの事前不正確率が2.5%、5%、10%、20%、33%、50%となるように非不正会計企業のサンプルをアンダーサンプリングした場合の検知精度を比較し、トレーニングデータの事前不正確率の違いによって、不正会計検知モデルの精度はどのように変化するかを確認する。事前不正確率と精度の値は、表7に示す。

表7 トレーニングデータの事前不正確率の違いによる検知精度

事前不正確率	2.5%	5%	10%	20%	33.33%	50%
Correct Classification	62.18%	61.85%	61.34%	61.25%	59.32%	57.33%
Sensitivity	56.13%	57.04%	57.49%	57.02%	58.79%	59.33%
Type I Error	37.75%	38.09%	38.62%	38.70%	40.67%	42.70%
Type II Error	43.87%	42.96%	42.51%	42.98%	41.21%	40.67%
Precision	1.76%	1.77%	1.76%	1.75%	1.72%	1.64%

不正会計検知モデルの全体の正しい判別率を示すCorrect Classificationは、事前不正確率が2.5%から50%と高くなるようにアンダーサンプリングするにつれて、その値が悪化する（62.18%→57.33%）。すなわち、アンダーサンプリングするほど、構築した不正会計検知モデルは全体の的中率が落ちるという結果である。Feroz et al. (2000) では、事前不正確率を10%、20%、33%、50%とした場合の不正会計検知モデルの精度比較を行った際に、事前不正確率が高いほどCorrect Classificationの値が悪化するという結果を得ている。本稿の結果においてもCorrect Classificationの値が悪化しており、Feroz et al. (2000) と同様の結果である。さらに、2.5%から10%の範囲については、Feroz et al. (2000) で検証されておらず、本稿ではこの範囲でもCorrect Classificationの値が悪化することが明らかとなった。

不正会計企業に対する評価指標であるSensitivityは、事前不正確率が20%の場合を除き、2.5%から50%と高くなるにつれてその値が改善している傾向が明らかになった（56.13%→59.33%）。すなわち、事前不正確率を高く設定したアンダーサンプリングを行うほど、構築された不正会計検知モデルによって不正会計企業を非不正会計企業と推定してしまい、不正を見逃すリスクが低減する。

推定不正会計企業における正しい判別率を示すPrecisionは、事前不正確率が5%となるようにアンダーサンプリングした場合の値が1.77%で他のケースのPrecisionの値よりも高いことが明らかになった。このような結果となった理由は、事前不正確率が2.5%

から5%となるにつれて、Type I Errorが悪化するものの(37.75%→38.09%:0.34%の悪化)、それ以上にSensitivityが改善している(56.13%→57.04%:0.91%の改善)ことに起因していると考えられる。すなわち、事前不正確率が5.0%となるようにアンダーサンプリングすることで、追加的な監査プロセスを行わなければならない企業件数を少なくすることができる。

このように、Precisionの値を見ると、事前不正確率が5%程度となるようにアンダーサンプリングすることが有効であると考えられるが、Sensitivityの精度をどの程度高めたいかによって、さらに事前不正確率を高めるようにアンダーサンプリングすることもあるだろう。ただ、Type I ErrorとType II Errorの合計値も、事前不正確率が5%の場合が他のケースよりも低い値を示しており、精度の良いモデルを構築することができると思われる。

不正会計企業と非不正会計企業の観測値に偏りがあることが不正会計検知モデルを構築する上での問題であることは、Perols et al. (2017) やBao et al. (2020) 等で指摘されている。Bao et al. (2020) はこの問題に対して、アンダーサンプリングが有効であると述べている。本稿の結果は、事前不正確率が5%程度になるようアンダーサンプリングを行うことは、不正会計企業と非不正会計企業の観測値の不均衡の解消につながり、検知精度(特に、PrecisionとType I ErrorとType II Errorの合計)の改善に寄与していることを示している。一方で、アンダーサンプリングを行い事前不正確率が10%以上のトレーニングデータとすると、不正会計企業と非不正会計企業を正しく判別できるモデルを構築することが難しくなることを示している。Dechow et al. (2011) は、SECにより見つかっていない不正会計企業がサンプルの中に含まれている可能性に言及している。本稿の結果から、事前不正確率を現実的な値(本稿では1.19%)とはあまりにも乖離している値(10%~50%)になるようにトレーニングデータを設定するのではなく、約4倍から5倍(4.76%~5.95%)を想定し、不正会計検知モデルを構築することが有効であるかもしれない。

## 5. 結び

本稿は、2011年から2019年における東京証券取引所の上場企業を対象として、ロジスティック回帰を用いた不正会計検知モデルを構築し、事前不正確率の違いによって検知モデルの精度がどのように変化するのかを明らかにした。結果として、不正実施確率(Probability)のカットオフ値に事前不正確率を加味して算出するFスコアの値が大きくなるほど、全体の分類精度が高まるが、不正会計実施企業を不正会計企業として推定できなくなることが明らかになった。加えて、トレーニングデータに対してアンダーサンプリングを行い、事前不正確率を2.5%から50%までの複数のパターンを設定して不正会計検知モデルを構築すると、事前不正確率が高くなるほど、全体の分類精度は低ま



るがSensitivityは改善すること、Type I ErrorとType II Error の合計とPrecision で見る検知精度が最も高くなる事前不正確率（本稿の場合は5%）があることを示した。

本稿の貢献は次の3点である。

第1に、不正会計企業と非不正会計企業の観測値に偏りがあるという問題に対して、アンダーサンプリングが有効であること、そしてアンダーサンプリングを行うべき適切な範囲があることを示したという貢献がある。

第2に、ロジスティック回帰モデルによる不正会計検知モデルを構築するにあたり、Lasso回帰を用いている点は佐藤他（2022）を除く先行研究とは異なっている。佐藤他（2022）は、Lasso回帰を用いた変数選択を行った理由として、不正会計検知モデルの構築に当たり、先行研究で用いられてきた変数選択の方法（例えば制約付きBackward方式）では、潜在的に有用な変数を選択せずに捨ててしまうことがあり、Lasso回帰を用いることでそのような問題に対処することが出来るためと述べている。本稿でも佐藤他（2022）に倣ってLasso回帰による変数選択を行っている。

第3に、本稿の結果から、ロジスティック回帰による不正会計の検知モデルを構築する際、次の2段階の手順を踏むという方向性が考えられる。第1のステップでは、層別10分割交差検証法に基づいて、データセットに対する事前不正確率を複数設定し、アンダーサンプリングをすることで、不正会計企業と非不正会計企業の観測数に偏りがあることによる検知精度の悪化を改善する。本稿の結果でいえば、事前不正確率が5%となるようにアンダーサンプリングする。第2のステップでは、アンダーサンプリングされたトレーニングデータに対して、ロジスティック回帰による不正会計の検知モデルを構築し、選択された変数とその推定値の値を得る。推定された検知モデルからFスコアを算出し、得られたFスコアの値に対し、精度を比較してFスコアのカットオフ値をどの値とするかを決定し、不正会計企業と非不正会計企業の判定を行う。

本稿の限界と将来の課題は次の4点である。

第1に、不正会計検知モデルの構築にあたりモデルに加える説明変数の候補をDechow et al.(2011)と首藤他(2016)に基づいて設定しているために、先行研究(Fanning and Cogger, 1998 ; Lin et al., 2003 ; Kotsiantis et al., 2006 ; Feroz et al., 2000) で用いられた説明変数やその他の研究（例えば、Perols, 2011）で使用された説明変数についての検討を十分には行っていない点である。例えば、ガバナンスに関する変数や倒産確率に関する変数を不正会計検知モデルに加えることによって検知精度が改善するかについても検討する必要がある。

第2に、Sensitivityの精度とPrecisionの精度のトレードオフについて、公認会計士や監査法人に対するアンケート調査やインタビュー調査を実施して、その重みづけを行う必要がある。ヒアリングして得た加重値を中心にシミュレーションを行い、加重値を変化させた場合に、構築される不正会計検知モデルがどのように異なるか、また、その検知精度がどのように変化するかを明らかにする必要がある。

第3に、本稿で得られた結果の頑健性のテストのため、トレーニングデータの期間を3年、テストデータの期間をその後の1年とするような組み合わせをローリングして、不正会計検知モデルの精度比較を行うことがある。ローリングによる精度比較を行う理由は、構築した不正会計検知モデルを監査の現場で使用する際に、過去の長期間のデータを収集するほどコストがかかること、構築した検知モデルによって将来の中長期間における不正会計よりも、次期の不正会計を検知したいというニーズがあると考えられるためである。

第4に、本稿は、アンダーサンプリングによるサンプル選択バイアスが予測確率に与える影響については取り上げていない。例えば、Dal Pozzolo et al. (2015) は、この問題に対してバイアスの影響を除去する（キャリブレーション）手法を提示しており、この手法を参考にして、日本企業における不正会計の検知モデルを構築することを今後の課題とする。

（さとう なつき・神奈川大学大学院工学研究科博士前期課程）

（こむら あゆこ・神奈川大学工学部経営工学科特別助手）

（ひらい ひろひさ・神奈川大学工学部経営工学科教授）

## 参考文献

- Bao, Y., Ke, B., Li, B., Julia Yu, Y., and J. Zhang (2020) "Detecting Accounting Fraud in Publicly Traded U.S. Firms Using a Machine Learning Approach", *Journal of Accounting Research*, Vol.58, pp.199-235.
- Beasley, M. S., Carcello, J. V., Hermanson, D. R., and T. L. Neal (2010) "Fraudulent Financial Reporting: 1998-2007. An Analysis of US Public Companies", *American Institute of Certified Public Accountants (AICPA) Historical Collection*, pp.1-85.
- Dal Pozzolo, A., Caelen, O., Johnson, R. A., and G. Bontempi (2015) "Calibrating Probability with Undersampling for Unbalanced Classification", In *Computational Intelligence, 2015 IEEE Symposium Series on*, pp.159-166.
- Dechow, P. M., Sloan, R. G., and A. P. Sweeney (1995) "Detecting Earnings Management", *The Accounting Review*, Vol.70, No.2, pp.193-226.
- Dechow, P. M., Ge, W., Larson, C. R., and R. G. Sloan (2011) "Predicting Material Accounting Misstatements", *Contemporary Accounting Research*, Vol.28, No.1, pp.17-82.
- Fanning, K., and K. Cogger (1998) "Neural Network Detection of Management Fraud using Published Financial Data", *International Journal of Intelligent Systems in Accounting, Finance and Management*, Vol.7, No. 1, pp.21-41.
- Feroz, E., Kwon, T., Pastena, V., and K. Park (2000) "The Efficacy of Red Flags in Predicting the SEC's Targets: An Artificial Neural Networks Approach", *International Journal of Intelligent Systems in Accounting, Finance and Management*, Vol.9, No.3, pp.145-157.
- Kaszniak, R. (1999) "On the Association between Voluntary Disclosure and Earnings Management", *Journal of Accounting Research*, Vol.37, No.1, pp.57-81.
- Khan, M. N., and R. L. Watts (2009) "Estimation and Empirical Properties of a Firm-Year Measure of Accounting Conservatism", *Journal of Accounting & Economics*, Vol.48, No.2-3, pp.132-150.
- Kothari, S. P., Leone, A., and C. Wasley (2005) "Performance-Matched Discretionary Accrual Measures", *Journal of Accounting and Economics*, Vol.39, No.1, pp.163-197.

- Kotsiantis, S., Koumanakos, E., Tzelepis, D., and V. Tampakas (2006) "Forecasting Fraudulent Financial Statements using Data Mining", *International Journal of Computational Intelligence*, Vol.3, No.2, pp.104-110.
- Lin, J., Hwang, M., and J. Becker (2003) "A Fuzzy Neural Network for Assessing the Risk of Fraudulent Financial Reporting", *Managerial Auditing Journal*, Vol.18, No.8, pp.657-665.
- 大城直人 (2014)「不正会計の早期発見に関する海外調査・研究報告書」, 金融庁金融研究センターディスカッションペーパー, DP2014-6, pp.1-63.
- Perols, J. L. (2011) "Financial Statement Fraud Detection: An Analysis of Statistical and Machine Learning Algorithms", *A Journal of Practice & Theory*, Vol.30, No.2, pp.19-50.
- Perols, J. L., Bowen, R. M., Zimmermann, C., and B. Samba (2017) "Finding Needles in a Haystack: Using Data Analytics to Improve Fraud Prediction", *The Accounting Review*, Vol.92, No.2, pp.221-245.
- Roychowdhury, S. (2006) "Earnings Management through Real Activities Manipulation", *Journal of Accounting and Economics*, Vol.42, No.3, pp.335-370.
- 佐藤夏輝, 小村亜唯子, 平井裕久 (2022)「日本企業における不正会計のタイプ別検知モデルの構築」, 産業経理, Vol.81, No.4, pp.126-140.
- Song, M., Oshiro, N., and A. Shuto (2016) "Predicting Accounting Fraud: Evidence from Japan", *The Japanese Accounting Review*, Vol.6, pp.17-63.
- 首藤昭信, 大城直人, 宋明子 (2016)「不正会計予測モデルの構築」, CARF-SERIES, CARF-J-103, pp.1-52.

## Detection Accuracy in a Model Detecting Accounting Fraud depending on Prior Fraud Probability using Logistic Regression

SATO Natsuki · KOMURA Ayuko · HIRAI Hirohisa

### Abstract

Construction of a model detecting accounting fraud using logistic regression is being studied. The indices such as Correct Classification, Sensitivity, Type I error and Type II error are used to compare the accuracy of models detecting accounting fraud. This paper aims to clarify (1) changes in the accuracy of models detecting accounting fraud constructed using logistic regression depending on differences in prior fraud probability of training data and (2) differences in the accuracy when prior fraud probability is added or is not added to the cutoff value for the execution probability of accounting fraud.

Demonstration experiments were held in companies listed on the Tokyo Stock Exchange during the period from 2011 to 2019. We constructed a model detecting accounting fraud and compared the accuracy. The results showed (1) when prior fraud probability is between 0.0172 and 0.05, the higher probability led to improvement in the value of Correct Classification and when it is between 0.05 and 0.5, it made the value worse and also showed (2) the value of Type I error was worse when prior fraud probability is added compared with no addition of the probability.