

自然言語処理を用いた株主提案の要領に関する分析

久多里桐子

北九州市立大学 経済学部 経営情報学科

kirikokudari@kitakyu-u.ac.jp

概要

本論文は、株主提案の要領に関する基礎研究を行うことを目的とする。株主提案の要領に含まれる3つのテキスト(提案内容、提案理由、取締役会の意見)を独立した文書とみなし、各文書をトピックモデル(LDA)を用いて分類した。その上で、各トピックにおける賛成率にどのような傾向が見られるのかを確認し、さらに、提案内容のトピック別に文書の特徴量(文字数および文字数に占める数字の割合)と賛成率の関係を分析した。2011年～2019年の定時株主総会における株主提案議案1,352件を対象に分析を行った結果、トピックによって得られる賛成率が異なること、そしてトピックによって文書の特徴量と賛成率との関係が異なる可能性があることを確認した。本論文で得られた結果は、株主提案の要領が他の株主の議決権行使に影響を及ぼしており、さらに、株主提案議案の賛成率は要領に記載された3つの文書のトピック、および文書特徴量などの複合的な要因に影響される可能性があることを示している。

1 本論文の目的と意義

本論文は株主提案の要領に関する基礎研究を行うことを目的とする。株主提案の要領は、主に3つの文書から構成される。提案内容、提案理由、そして株主提案議案に対する企業側の意見である。本論文では各株主提案議案から上記3つの文書に関するテキストを収集し、機械学習の手法を用いて議案を分類した上で、賛成率にどのような差異が見られるのかを分析する。また、文書の特徴量として文字数および文字数に占める数字の割合という2つの尺度を抽出し、それらの尺度と賛成率の間にどのような関係が見られるのかを明らかにする。

2021年3月に施行された改正会社法では、株主提案権の濫用的行使を防ぐことを目的として、株主が提出しようとする議案の「数」は10件までに制限された。一方で、株主提案の「内容」に制限を設けることは見送られた。ただし、2019年に公表された要綱案の段階では、内容に関する制限も含まれていた。具体的には、提案が「専ら人の名誉を侵害し、人を侮辱し、若しくは困惑させ、又は自己若しくは第三者の不正な利益を図る」ことを目的としていた場合や、「株主総会の適切な運営が著しく妨げられ、株主の共同の利益が害されるおそれがあると認められる場合」には、株主提案権の行使を制限することが予定されていた(法務省, 2019)^{*1}。

しかしながら、当該内容は後の国会における法案審議において撤回される。その理由として、「権利の濫用に該当する株主提案権の類型についてさらに精緻に分析を深めながら、引き続き

^{*1} 株主提案権には議題提案権(会社法第303条)、議案提出権(同法第304条)、議案要領通知請求権(同法第305条第1項)が含まれる。当該制限は、議案提出権および議案要領通知請求権を対象としていた。

検討していくべき」(竹林他, 2020, p.5)と述べられている。つまり、2021年の改正会社法において、株主提案の内容に関する制限が含まれなかったことは、関連する知見が少ないことに起因する。実際に、株主提案の要領について定量的観点から分析を行なった研究は存在しない。そこで本論文は、株主提案権に関する今後の検討材料に資するという意義の下、株主提案の内容に焦点を当てた基礎的分析を行う。

つづく第2節では、米国および日本の株主提案権について、提案内容および記述に関する制度上の規定について参照し、関連する先行研究を紹介する。第3節では、分析手法とサンプルについて述べる。第4節では分類したトピックの概要を確認し、第5節で主たる発見事項を提示する。最後に、第6節で本論文の要約を述べる。

2 株主提案の要領に関する制度上の規定とトピックモデルを利用した先行研究のレビュー

2.1 株主提案の要領に関する制度上の規定

本項では、米国および日本の株主提案権を対象に、記述量および提案内容に関する制度上の規定について述べる。米国では、記述量と提案内容のいずれにおいても、日本に比して制限が厳しい。米国の上場企業において、株主が企業の株主総会招集通知(proxy statement)に提案内容等の記載を求める場合、米国証券取引委員会(U. S. Securities and Exchange Commission; SEC)の規則に則る必要がある。まず、記述量は補足説明を含めて500単語以内に限られる。さらに、提案内容が取締役の選任に関するものであったり、具体的な配当金額に関するものであったりした場合、企業側は提案を排除できる^{*2}。

一方で、日本の上場企業において株主が提案できる内容は多様であり、記述量に関する法的な制限もない。ただし、提案の理由が明らかに虚偽である場合または専ら人の名誉を侵害し、若しくは侮辱する目的によるものと認められる場合には、提案理由を削除されることもある。また、企業側が株主提案議案を拒否できる事由として明文化されているものは、法令もしくは定款に違反する場合や、実質的に同一の議案につき総株主の議決権の10の1以上の賛成を得られなかった日から3年を経過していない場合が該当する(会社法第304条ただし書き、第305条第4項)。一般的には、株主が記述した提案内容や提案理由の原文に加えて、議案に対する取締役(取締役会設置会社の場合は取締役会)の意見(以下、取締役会の意見)が、株主総会招集通知内の株主総会参考書類に記載される。

2.2 トピックモデルを用いた金融、経済分野の先行研究

本論文に関連して、金融・経済分野のテキストを対象にトピックモデルを用いた分析を行っている以下3つの研究を挙げる。Feuerriegel and Pröllochs (2021)は、2004年から2013年の期

^{*2} ただし、企業が株主提案議案を排除する場合には、株主総会招集通知をSECに提出する80日以上前に、排除する理由をSECに提出する必要がある(17 C.F.R. §240.14a-8(j))。企業が当該提案を排除することに妥当性があると認められた場合には、SEC企業財務局スタッフ(division of corporation finance staff)がSECに法執行を勧告しないことを保証する、非公式かつ拘束力のない見解(ノー・アクション・レター)を述べる。

間に提出された Form 8-K を EDGAR から取得し、所定の要件を満たす 73,986 件の Form 8-k を 20 のトピックに分類した^{*3}。分類したトピックによって、投資家の反応は異なることが確認されている。例えば、「事業戦略」「医療セクター」「格付け」トピックは異常リターンがプラスであったのに対して、「決算」「M&A」トピックは異常リターンがマイナスであった。

竹澤他 (2021) は、過去の情報を伝達する新聞記事は株価の予測に役立つか、というリサーチ・クエスションのもと、2003 年から 2012 年までの新聞記事 77,814 件を 30 のトピックに分類し、日ごとのトピックの割合と市場リターン (TOPIX) との関係を分析している。分析を通じて、トピックの割合と市場リターンの間に相関は見られなかった一方で、一部のトピックの割合と市場ボラティリティの間に相関があることを発見した。

中野他 (2022) では 2004 年から 2021 年までの所定の要件を満たす有価証券報告書 43,289 件をもとに、MD&A (経営成績等の分析) の記述について分析を行い、近年のディスクロージャー制度改革が MD&A の開示に影響を及ぼしていることを報告している。具体的には、経営戦略および会計上の見積等に関する記述が充実していることが確認された。いずれの研究も、文書の分類にあたって機械学習の 1 つであるトピックモデルが採用されている。

3 分析手法とサンプル

3.1 分析の概要

分析の概要は以下の通りである。はじめに、株主提案の要領である 3 つのテキスト (提案内容、提案理由、取締役会の意見) を独立した文書とみなし、各文書の特徴量を取得する。つづいて、各文書をトピックに分類し、トピックごとの賛成率や、文書の特徴量と賛成率の関係を確認する。本論文では、文書の特徴量として以下 2 つを採用した。第 1 に、文字数である。定款で文字数に制限を設けている企業に対して、当該定款の削除を求める株主提案が存在したことから明らかなように、記述量は賛成率に影響を与えられられる。第 2 に、文字数に占める数字の割合である。特に剰余金の分配に関する提案など、特定の議案においては具体的な数字を用いた提案の方が説得力は高まると考えられる。以下は実際の提案理由に基づく特徴量であるが、例 A よりも例 B の方が賛成率は高いと予想する。

例 A: 増配できる十分な利益がある。

- 文字数: 14 字
- 文字数に占める数字の割合: 0%

例 B: 貴社は、平成 30 年 9 月末の連結貸借対照表にて現預金及び投資有価証券の合計額 350 億円を有し、(省略) 妥当なものであると考えます。

- 文字数: 278 字
- 文字数に占める数字の割合: 6.47%

議案からテキストを抽出する際の留意事項は下記の通りである。提案内容が取締役等の役員

^{*3} 米国上場企業は、投資家の意思決定に影響を及ぼす可能性のある重要な事象 (決算発表、重要な契約の締結、倒産など) が生じた際、4 営業日以内に Form 8-k (current report) を SEC に提出する必要がある (U. S. Securities and Exchange Commission, 2021)。

選任議案の場合、当該候補者の経歴等は提案内容に含めない。また、会社注も含めない。さらに、提案内容と提案理由が混在している場合、可能であれば文章を分割したが、明確に区別されていなかった場合、提案理由の記述はないものとみなしている。加えて、取締役会の意見が複数の議案に対する集約的なものであった場合、各議案に同じ記述がなされたものとして扱った。ただし、個別の議案に対する意見に加えて、別立てで議案全体に対する取締役会の意見が付されていた場合は、個別の議案に対する意見のみを取締役会の意見として収集している。

3.2 テキストの処理方法および LDA モデルの構築

トピックモデルの構築にあたり、以下の通りに前処理および形態素解析を行う。表記の統一を目的として、全角コロン「,」を読点「,」に、全角ドット「.」を句点「。」に変換した上で、全角の英数字と記号(括弧や%など)を半角に変換する。数値に関しては桁数区切りのコンマを除いた上で、0に置き換える。そのため、例えば「100万円」は「0万円」、「10,000人」は「0人」に置き換わっている。形態素解析にあたっては解析器に MeCab を、辞書に mecab-ipadic-neologd (佐藤他, 2017) を利用した。品詞は名詞、形容詞、動詞を採用し、ストップワードとして SlothLib に出現頻度の高い単語を独自で追加した^{*4}。具体的には、昭和、平成、令和、当社、貴社、会社、株主、提案、議案をストップワードとして採用している。また、ひらがな 1、2 文字の単語も除外した。トピックの分類においては、潜在的ディリクレ配分法 (Blei et al., 2003) によるトピックモデルを採用した。LDA モデルの構築において、出現文書数が 13 件 (サンプルの 1%) 以下の単語および 20% 以上の文書に出現する単語を辞書から除いた^{*5}。

3.3 トピック数に関する検討

本論文で設定するトピック数に関して、各文書についてトピック数を 2~30 の範囲でシミュレーションを行い、Perplexity および Coherence の値をもとに決定した。佐藤・奥村 (2015) によると、Perplexity は分岐数または選択肢の数を表しており、より少ない値であれば、そのモデルまたは学習アルゴリズムの絞り込み性能が良いことを示す。例えば、「明日の天気は ** だ。」という文章の ** に入る単語の候補数に相当する。一方で、Coherence は単語間類似度の平均を示す。トピック内で出現確率の高い順に c 個の単語を選び、集合 $W_k = \{w_1, w_2, \dots, w_c\}$ とした際、 W_k 内の単語間の類似度が高いほど Coherence は高い。したがって、トピック全体の Coherence が高ければ良い学習アルゴリズムないしモデルであるとされる。

図 1~6 は各文書についてトピック数を 2~30 の範囲でシミュレーションを行った場合の Coherence (Mimno et al., 2011) および Perplexity の値を示している^{*6}。なお、図 1・3・5 はコーパスに対して TF-IDF による重み付けを実施しなかった場合の結果を示しており、図 2・4・6 はコーパスに対して TF-IDF による重み付けを実施した場合の結果を示している^{*7}。全体的な

^{*4} 多くの自然言語処理手法を用いた研究において、一般的に除外されるストップワードを集約したリストが SlothLib である。

^{*5} 例えば「取締役会の意見」のテキストをもとに作成した辞書からは、「反対」「取締役会」などの定型的な単語が除かれる。

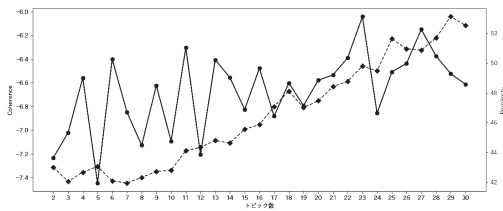
^{*6} Coherence の算定方法は複数存在するが、本論文では Python ライブラリの gensim で u.mass を指定して Coherence を算定した。

傾向として、トピック数が増えるほど、Coherence は低下し、Perplexity は上昇する。特にコーパスを TF-IDF で重み付けした場合は、Perplexity が著しく高くなる。例えば、図 1 において Perplexity は最大で 50 程度であるのに対し、図 2 では最大 900 程度である。しかしながら、図 2 では Coherence が他の図と同様に下降傾向であるのに対して、図 1 では Coherence が上昇傾向であり、トピック数が 1 つ増える度に値が大きく上下しているため、適切なトピック数を判断できない。したがって、本論文では、上記のシミュレーション結果に基づき、コーパスの作成に際して TF-IDF による重み付けを実施した上で、提案内容は 5、提案理由は 5 (ただし、テキストが取得できなかった議案のみの Topic 6 も設定)、取締役会の意見は 6 に設定した*⁸。

分析対象は 2011 年～2019 年の定時株主総会に提出され、かつ、賛成率を入手できた株主提案議案 1,352 件である*⁹。データの取得源は以下のとおりである。株主提案の要領に関しては、eol (株式会社アイ・エヌ情報センター) から取得した株主総会招集通知の PDF から手作業でテキストを収集した。賛成率については日経 NEEDS-Cges 株主総会関連データ (日経メディアマーケティング株式会社) から取得している。

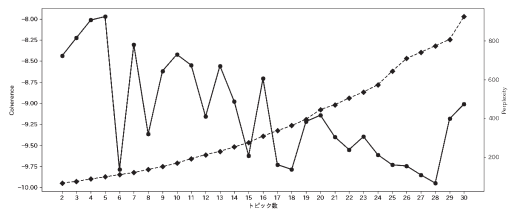
なお、株主提案の要領を分類する方法として、既存のデータベースを用いる方法もある。例えば、日経 NEEDS-Cges 株主総会関連データでは、提案内容に基づいて議案が 13 種類 (定款変更、剰余金処分、取締役選任、監査役選任、役員報酬、役員賞与、退職慰労金、会計監査人選任、ストックオプション、買収防衛策、自社株、企業再編、その他) に分類されており、その 1 つに含まれる定款変更については、さらに 14 種類に細分化されている。つまり、内容に応じて 26 種類に分類されたデータを利用することができる。しかしながら、そもそも提案内容に基づいて議案を分類する上で、何種類が適切であるのかを判断するのは難しい。また、提案理由、取締役会の意見についてはデータベース化されていないため、トピックモデル等によって分類する必要がある。上述のような背景から、本論文では機械学習の手法を用いて、全ての文書 (提案内容、提案理由、取締役会の意見) を分類することとした。

図 1 提案内容のトピック別 Coherence および Perplexity (重み付けなし)



(出所) 著者作成
 (注) 縦軸について、左軸および黒色実線は Coherence を表し、右軸および灰色点線は Perplexity を表す。横軸はトピック数 (2～30) を示す。

図 2 提案内容のトピック別 Coherence および Perplexity (重み付けあり)



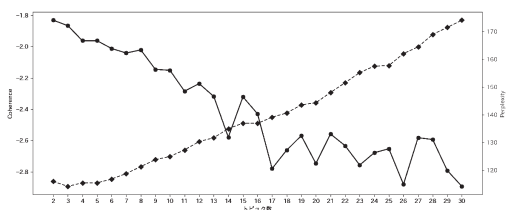
(出所) 著者作成
 (注) 詳細は図 1 の注を参照されたい。

*7 TF-IDF は文書中の単語の重要度を評価する指標である。

*8 追加分析として、TF-IDF による重み付けを行わず、かつ、トピック数を変更 (提案内容は 22、提案理由は 6 (テキストが取得できなかった議案のみの Topic 7 も設定)、取締役会の意見は 7) した場合についても検討したが、分析結果の解釈に基づく結論は変わらなかったため、記載を省略している。

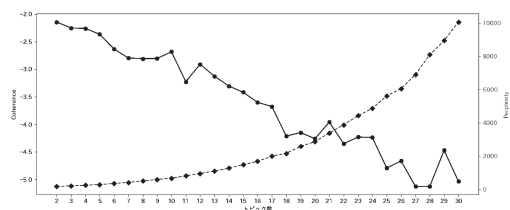
*9 付議の要件を満たさなかったなどの理由で決議されなかった議案については、賛成率が取得できない。

図3 提案理由のトピック別
Coherence および Perplexity (重み付けなし)



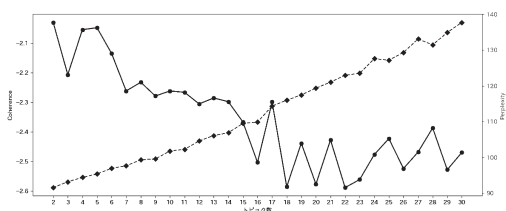
(出所) 著者作成
(注) 詳細は図1の注を参照されたい。

図4 提案理由のトピック別
Coherence および Perplexity (重み付けあり)



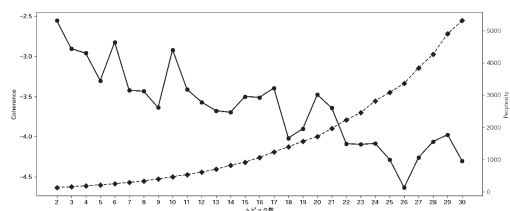
(出所) 著者作成
(注) 詳細は図1の注を参照されたい。

図5 取締役会の意見のトピック別
Coherence および Perplexity (重み付けなし)



(出所) 著者作成
(注) 詳細は図1の注を参照されたい。

図6 取締役会の意見のトピック別
Coherence および Perplexity (重み付けあり)



(出所) 著者作成
(注) 詳細は図1の注を参照されたい。

4 トピックの概要

4.1 基本統計量と相関係数

表1はサンプル全体の文書特徴量および賛成率に関する基本統計量を示している。変数名の括弧内について、「内容」は提案内容を、「理由」は提案理由を、「意見」は取締役会の意見をそれぞれ意味している。また、文字数に占める数字の割合については、「数字の割合」と表記を省略している。まず文字数に注目すると、平均値はそれぞれ、提案内容が131.23字、提案理由が417.43字、取締役会の意見が428.51字と、提案内容の文字数が比較的少ないことがわかる。ただし、本論文では取締役等の役員選任議案については、候補者の経歴等の記述を提案内容に含めていない点に注意が必要である。提案理由の文字数について、最小値は0字であるが、これは提案理由として明確な記述が得られなかった議案については文字数を0としていることに起因する。また、標準偏差は提案内容が246.60字、提案理由が426.89字、取締役会の意見が505.00字であり、特に取締役会の意見は議案によって記述量が異なる。

つづいて、文字数に占める数字の割合に着目すると、平均値はそれぞれ、提案内容が3.62%、提案理由が2.57%、取締役会の意見が1.07%である。提案内容の値が比較的高い理由は、分母の文字数が他の文書に比して小さいためである。その点を考慮すると、特に提案理由は記述量に比して数字を用いる割合が高いと言える。数字の数の平均値が最も高い文書は提案理由

表1 サンプル全体の基本統計量

変数	最小値	第1四分位	平均値	中央値	第3四分位	最大値	標準偏差	欠損値の個数
文字数(内容)	7.00	51.00	131.23	91.00	147.00	4,583.00	246.60	0
文字数(理由)	0.00	262.00	417.43	379.00	400.00	5,331.00	426.89	0
文字数(意見)	12.00	194.75	428.51	308.00	456.00	6,210.00	505.00	0
数字の数(内容)	0.00	0.00	5.16	2.00	5.00	164.00	11.59	0
数字の数(理由)	0.00	0.00	12.78	6.00	13.00	227.00	24.46	0
数字の数(意見)	0.00	0.00	7.13	1.00	5.00	411.00	24.51	0
数字の割合(内容)	0.00	0.00	3.62	2.08	5.36	43.18	4.85	0
数字の割合(理由)	0.00	0.24	2.57	1.70	3.68	28.57	3.22	15
数字の割合(意見)	0.00	0.00	1.07	0.43	1.44	23.41	1.93	0
賛成率	0.03	4.10	14.63	8.00	18.68	99.08	16.96	0

(出所) 著者作成

(注1) 表1は分析対象である1,352議案(168企業、347企業年度)に関する基本統計量を示している。

(注2) 変数名の括弧内について、提案内容は「内容」、提案理由は「理由」、取締役会の意見は「意見」と表記を省略している。

(注3) 単位について、文字数および数字の数は「字」、数字の割合および賛成率は「%」である。

(注4) 数字の割合(理由)における欠損値の個数は、提案理由として明確な記述がなかった15議案を意味する。

(注5) 賛成率について、特に役員選任に関する議案の場合は複数の賛成率を得られるが、議案ごとに平均値を取ったのちに全体的な基本統計量を算定した場合も上記と結果は大きく変わらない。

であり(12.78字)、文字数の平均値も提案内容より提案理由の方が高い。最後に、賛成率に注目すると、平均値が14.63%であり、第3四分位においても18.68%であることから、本論文のサンプルにおいてほとんどの株主提案議案が可決されていないことがわかる^{*10}。

表2はサンプル全体の文書特徴量および賛成率に関する相関係数を示している^{*11}。各文書間において、文字数と数字の数や、数字の数と数字の割合の間で高い相関が見られるが、これは当然の傾向と言える。例えば、表2において最も高い値(0.852)を示しているのは提案内容の文字数と提案内容の数字の数であるが、数字の数が増えると当然のこととして文字数も増える。注目すべきは、提案理由の文書特徴量と取締役会の意見の文書特徴量において、プラスの相関が確認されている点である。文字数(理由)と文字数(意見)の相関係数は0.271、数字の数(理由)と数字の数(意見)の相関係数は0.237、数字の割合(理由)と数字の割合(意見)の相関係数は0.180である。また、数字の割合(内容)と数字の割合(意見)の相関係数も0.247と比較的高い。つまり、株主提案議案において、提案理由の記述量が多いほど取締役会の意見も記述量が多くなる。また、数字が多く用いられたより具体的な提案理由あるいは提案内容であるほど、取締役会の意見においても、反対意見の説得力を高めるために、数字を多く用いたより具体的な記述がなされている可能性がある。

では、各文書の特徴量と賛成率の間にはどのような傾向が確認できるだろうか。全体的な傾向として、概ねプラスの相関が確認できる。特筆すべき点として、取締役会の意見に関する文

*10 議案の可決要件は決議事項によって異なるが、日経 NEEDS-Cges 株主総会関連データにおいて、2019年までの株主総会で可決された株主提案議案の最低賛成率は50.02%である(監査役の報酬金額に関する議案)。

*11 表1と同様に、変数名の括弧内について、「内容」は提案内容を、「理由」は提案理由を、「意見」は取締役会の意見をそれぞれ意味している。また、文字数に占める数字の割合については、「数字の割合」と表記を省略している。

書特徴量と賛成率の間でもプラスの相関が見られる(文字数(意見)は0.329、数字の数(意見)は0.184、数字の割合(意見)は0.065)点は、株主提案議案に対する賛成率が高いと予想される場合に、取締役会の意見における記述量や、数字を用いた記述が増えると解釈できるだろう。しかしながら、数字の割合(内容)との間でマイナスの相関が見られる点については、解釈が難しい。上述のとおり、提案内容において数字の割合が高い場合には取締役会の意見における数字の割合も高い傾向がある。また、取締役会の意見における数字の割合が高い場合には取締役会の意見における記述量も多い。加えて、賛成率との相関において文字数(内容)の係数は0.084、文字数(理由)の係数は0.082であるのに対して文字数(意見)の係数が0.329と高いことを考慮すると、数字が多く用いられたより具体的な提案内容であるほど、取締役会の意見の記述量が増える、あるいは数字が用いられることによる反対意見の具体性が高まり、結果として賛成率が低くなっている、と解釈できるかもしれない。

表2 サンプル全体の相関係数

	文字数 (内容)	文字数 (理由)	文字数 (意見)	数字の数 (内容)	数字の数 (理由)	数字の数 (意見)	数字の割合 (内容)	数字の割合 (理由)	数字の割合 (意見)	賛成率
文字数 (内容)	1									
文字数 (理由)	0.123	1								
文字数 (意見)	0.081	0.271	1							
数字の数 (内容)	0.852	0.114	0.102	1						
数字の数 (理由)	0.122	0.786	0.148	0.199	1					
数字の数 (意見)	0.067	0.250	0.624	0.107	0.237	1				
数字の割合 (内容)	-0.008	0.021	0.036	0.333	0.135	0.147	1			
数字の割合 (理由)	0.064	0.213	0.058	0.199	0.591	0.124	0.173	1		
数字の割合 (意見)	0.043	0.181	0.234	0.109	0.232	0.753	0.247	0.180	1	
賛成率	0.084	0.082	0.329	0.086	0.022	0.184	-0.052	0.002	0.065	1

(出所) 著者作成

(注1) 分析対象である1,352議案(168企業、347企業年度)に関するピアソン相関係数を示している。ただし、数字の割合(理由)の相関係数を算定するにあたり、提案理由として明確な記述がなかった15議案は除いている。

(注2) 各変数の括弧内について、提案内容は「内容」、提案理由は「理由」、取締役会の意見は「意見」と表記を省略している。

4.2 トピックの概要

表3・4・5は、各文書の各トピックにおける頻出上位5単語を示している。本論文では、前章のシミュレーション結果を踏まえて最も適したトピック数を選択したが、表3・4・5のいずれにおいても各トピックに重複した単語が散見される。そのため、表3・4・5の頻出上位5単

表3 提案内容のトピック別頻出上位5単語

トピック	単語1	単語2	単語3	単語4	単語5
Topic 1	解任	監査役	加える	取得	配当
Topic 2	氏	候補者	選任	社外取締役	変更
Topic 3	氏	解任	条項	事業	候補者
Topic 4	選任	配当	氏	報酬	設置
Topic 5	選任	候補者	配当	氏	事業

(出所) 著者作成

表4 提案理由のトピック別頻出上位5単語

トピック	単語1	単語2	単語3	単語4	単語5
Topic 1	報酬	開示	原発	個別	事業
Topic 2	事業	配当	株	原発	自己株式
Topic 3	社外取締役	氏	取締役会	役員	独立
Topic 4	自己株式	保有	株式	配当	0円
Topic 5	0円	株式	保有	配当	原発

(出所) 著者作成

(注) 明確な記述がなかったものは Topic 6 として分類している。

表5 取締役会の意見のトピック別頻出上位5単語

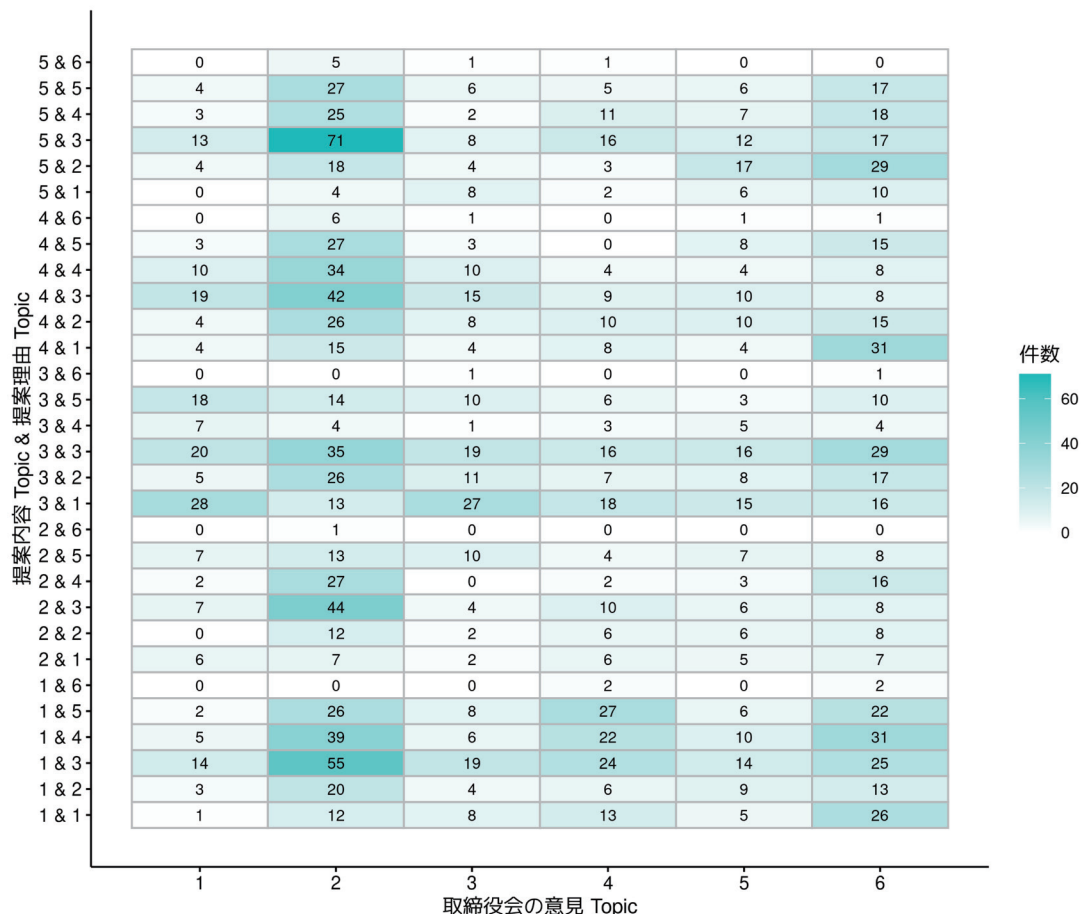
トピック	単語1	単語2	単語3	単語4	単語5
Topic 1	事項	候補者	自己株式	消却	業務
Topic 2	配当	報酬	監査役	選任	氏
Topic 3	報酬	社外取締役	執行	配当	設置
Topic 4	監査役	忠実	解任	事由	法令
Topic 5	記載	共通	報酬	設置	監査役
Topic 6	報酬	配当	開示	還元	事項

(出所) 著者作成

語に基づいて本論文の結果を正確に解釈することは困難ではあるが、例えば、提案内容トピックの Topic 2 には役員選任に関する議案、提案理由トピックの Topic 1 には報酬の個別開示を理由とした議案、取締役会の意見トピックの Topic 4 には「監査役は忠実に業務にあたっているため、解任する事由はない」との意見が表明された議案が、比較的多く分類されていると考えられる。

図7は文書別かつトピック別に分類された件数を示している。縦軸は提案内容トピックと提案理由トピックの組み合わせを示し、横軸は取締役会の意見トピックを示している。最も分類件数が多かったものは、提案内容トピックが Topic 5 かつ提案理由トピックが Topic 3 であり、取締役会の意見トピックが Topic 2 のケースである。表3・4・5の単語をもとに推測するならば、取締役会の独立性を理由に社外取締役等の役員選任議案が提出され、その提案に対して現任の社外取締役や監査役による監督・監査機能は十分である、との反対意見が表明されているケースが該当していると考えられる。

図7 文書別かつトピック別の件数



(出所) 著者作成

(注 1) 縦軸は提案内容トピックと提案理由トピックの組み合わせを示し、横軸は取締役会の意見トピックを示している。

(注 2) マス目の数値は分類された件数を示しており、マス目の色が濃いほど件数が多いことを意味する。

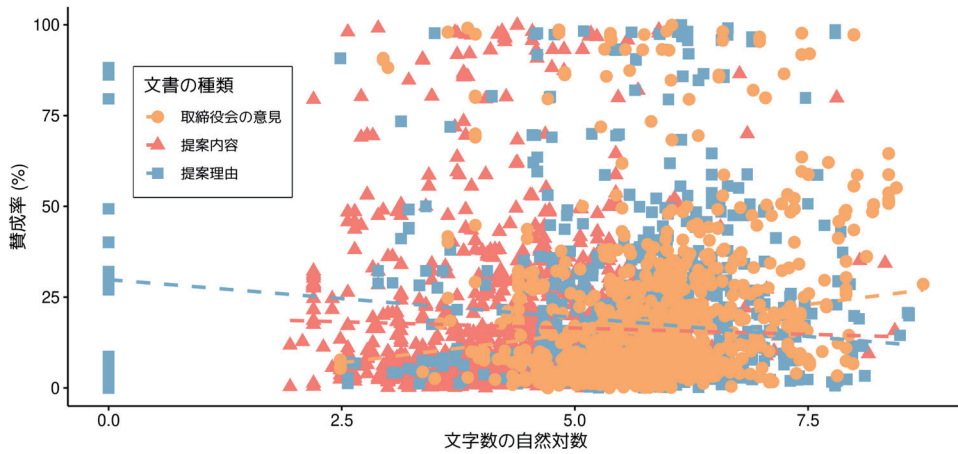
(注 3) 1つの議案が複数のトピックに分類されることを許容している。

5 分析結果

5.1 サンプル全体の傾向

はじめに、株主提案議案をトピックに分類する前の、文字数ないし文字数に占める数字の割合と賛成率の関係について確認する。図8はサンプルの1,352議案について、文字数の自然対数と賛成率の散布図、および被説明変数を賛成率、説明変数を文字数の自然対数とした単回帰の回帰直線を、文書別に示している。また、表6は単回帰の結果を要約している。図8と表6から明らかとなり、取締役会の意見については、記述量と賛成率の間にはプラスの関係があ

図8 文書別の文字数と賛成率



(出所) 著者作成

表6 賛成率に対する文字数の単回帰結果

文書の種類	回帰係数	p 値	決定係数
提案内容	-0.686	0.195	0.001
提案理由	-2.088	0.000	0.010
取締役会の意見	3.217	0.000	0.031

(出所) 著者作成

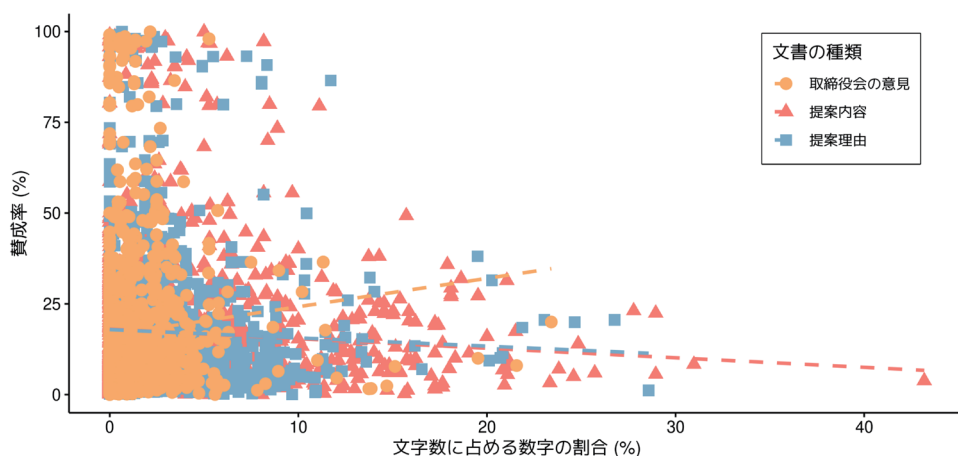
(注) 説明変数は文字数の自然対数をとった値である。

る (回帰係数は 3.217)。つまり、賛成率が高いと予想される場合に、取締役会の意見の記述量が増える可能性があると考えられる。これは、表 2 の相関において確認された傾向と同じである。

一方で、提案内容および提案理由については、記述量が増えると賛成率は低くなる傾向が見られる (回帰係数はそれぞれ -0.686 と -2.088)。提案内容については統計的な有意性を確認できていないが、少なくとも提案理由については、記述量が増えるほど他の株主にとっては吟味するためのコストが高くなるため、賛成票を投じる可能性が低くなるのかもしれない。

類似した傾向は、図 9 および表 7 でも確認できる。図 9 は文字数に占める数字の割合と賛成率の散布図、および被説明変数を賛成率、説明変数を文字数に占める数字の割合とした単回帰の回帰直線を、文書別に示している。また、表 7 は単回帰の結果を要約している。回帰係数はそれぞれ、提案内容が -0.263、提案理由は -0.228、取締役会の意見が 0.780 である。やはり、賛成率が高いと予想される場合に、取締役会の意見の記述に占める数字の割合が高くなると思われる。株主提案議案の全体的な傾向として、提案内容および提案理由の記述量を増やしたり、記述内に数字を多く含んで具体性を高めたりしても、必ずしも多くの賛成票を獲得できているとは言えない。

図9 文書別の文字数に占める数字の割合と賛成率



(出所) 著者作成

表7 賛成率に対する文字数に占める数字の割合の単回帰結果

文書の種類	回帰係数	p 値	決定係数
提案内容	-0.263	0.008	0.004
提案理由	-0.228	0.180	0.001
取締役会の意見	0.780	0.005	0.004

(出所) 著者作成

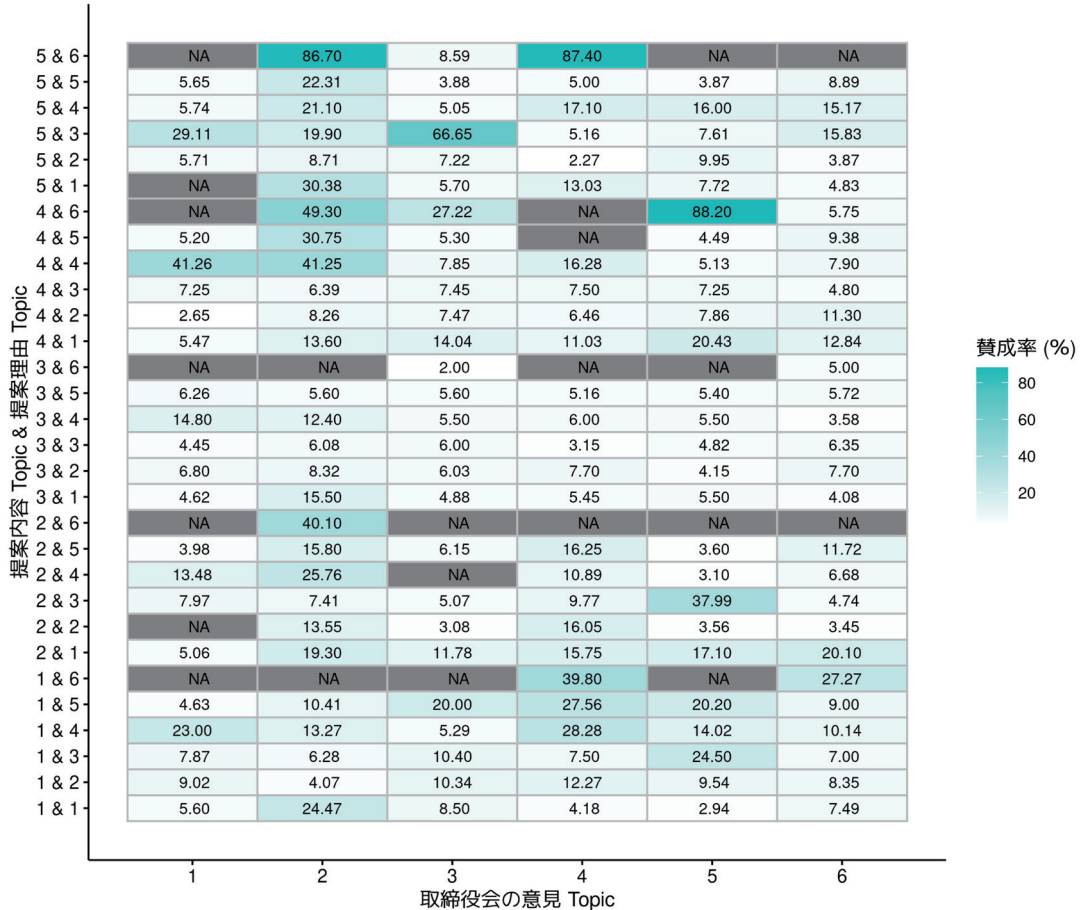
では、トピックに分類した場合にはどのような傾向が確認できるだろうか。図10は文書別かつトピック別に議案を分類した場合の、賛成率の中央値を示している。図7と同様に、縦軸は提案内容トピックおよび提案理由トピックの組み合わせを示しており、横軸は取締役会の意見トピックを示している。該当するケースがなかったものは、NAで表記している。賛成率が80%を超えているケースはいずれも、提案理由トピックのTopic6に分類されている。これは、提案理由として明確な記述がないにも関わらず高い賛成率を得ていることを意味するが、支配株主が可決を見越して十分な説明なしに提案を提出した場合が該当する^{*12}。

図10が示す重要な点は、提案内容トピック、提案理由トピック、取締役会の意見トピックの組み合わせによって、賛成率は異なるという点である。図10において、各マス目は概ね2%～40%台の範囲(上述のような支配株主による株主提案議案が多く分類されたマス目は除く)で賛成率が異なっていることがわかる。これは、他の株主が株主提案の内容や提案理由、それに対する取締役会の意見などを踏まえて、議決権行使を行っている可能性を示唆する。

最後に、トピックによって賛成率と文書特徴量との関係性も異なる可能性があることを提示

^{*12} 同様に、賛成率の中央値が66.65%であるマス目(提案内容トピック5、提案理由トピック3、取締役会の意見トピック3)についても、支配株主が取締役の選任に関する株主提案議案を提出し、かつ、取締役会の意見として賛成が表明された結果、賛成率が90%を超えたケースが分類されている。

図 10 文書別かつトピック別の賛成率



(出所) 著者作成

(注1) 縦軸は提案内容トピックと提案理由トピックの組み合わせを示し、横軸は取締役会の意見トピックを示している。

(注2) マス目の数値は分類された議案における賛成率の中央値を示しており、マス目の色が濃いほど賛成率が高いことを意味する。NAは分類された議案がないことを意味する。

(注3) 役員選任に関する株主提案議案においては、候補者が複数名存在する場合など、複数の賛成率が得られる場合がある。それらは全て、1つのマスに集約されている。例えば、上記のような提案内容であるために10%と30%の賛成率が存在している1つの議案が1つのマス目に分類された場合、両者の中央値である20%がマス目に表記されることとなる。

する。表8および表9はそれぞれ、提案内容トピックでサンプルを分類した場合の賛成率に対する文字数(自然対数をとったもの)の単回帰結果と、賛成率に対する文字数に占める数字の割合の単回帰結果を示している。サンプル全体の結果を示す図8・9および表6・7では、提案内容と提案理由の文書特徴量と賛成率との間にはマイナスの関係がある可能性が示唆されていた。しかしながら、議案内容によっては異なる傾向が見られる。表8のTopic 1、Topic 4、Topic 5では、提案内容の記述量が増えるほど、賛成率が高くなることが示されている。表9においても、統計的な有意性は確認できないが、Topic 1、Topic 3、Topic 4では、提案理由の文

字数に占める数字の割合はプラスに推定されている。つまり、株主提案議案を提案内容トピックで分類した場合には、提案内容や提案理由の記述量が増える、もしくは記述に占める数字の割合が高まるほど、必ずしも賛成率が低くなるとは限らない可能性がある。

表8 提案内容トピック別かつ文書別の賛成率に対する文字数の単回帰結果

提案内容トピック	文書の種類	回帰係数	p 値	決定係数
Topic 1	提案内容	1.298	0.011	0.013
	提案理由	-2.134	0.001	0.024
	取締役会の意見	2.221	0.000	0.031
Topic 2	提案内容	1.701	0.104	0.009
	提案理由	-8.423	0.000	0.157
	取締役会の意見	0.257	0.819	0.000
Topic 3	提案内容	-0.395	0.528	0.001
	提案理由	-0.341	0.708	0.000
	取締役会の意見	2.089	0.000	0.035
Topic 4	提案内容	2.779	0.006	0.019
	提案理由	0.097	0.930	0.000
	取締役会の意見	5.904	0.000	0.078
Topic 5	提案内容	2.331	0.052	0.009
	提案理由	-1.262	0.309	0.003
	取締役会の意見	2.245	0.071	0.009

(出所) 著者作成

(注) 説明変数は文字数の自然対数をとった値である。

表9 提案内容トピック別かつ文書別の賛成率に対する文字数に占める数字の割合の単回帰結果

提案内容トピック	文書の種類	回帰係数	p 値	決定係数
Topic 1	提案内容	-0.352	0.000	0.026
	提案理由	0.094	0.614	0.001
	取締役会の意見	0.109	0.744	0.000
Topic 2	提案内容	-0.713	0.000	0.060
	提案理由	-0.935	0.045	0.018
	取締役会の意見	1.367	0.083	0.013
Topic 3	提案内容	-0.120	0.278	0.003
	提案理由	0.428	0.104	0.007
	取締役会の意見	0.642	0.259	0.003
Topic 4	提案内容	-0.604	0.000	0.031
	提案理由	0.550	0.321	0.003
	取締役会の意見	1.492	0.071	0.009
Topic 5	提案内容	-0.953	0.000	0.052
	提案理由	-0.213	0.682	0.000
	取締役会の意見	0.083	0.905	0.000

(出所) 著者作成

6 本論文の結論

本論文では、株主提案の要領に関する基礎研究を行った。分析の結果、提案内容や提案理由、取締役会の意見のトピックによって、得られる賛成率が異なることを確認した。また、各文書の特徴量として、文字数および文字数に占める数字の割合に着目し、賛成率との関係について分析を行ったところ、株主提案議案の全体的な傾向としては、提案内容および提案理由に関しては記述量および記述に占める数字の割合が賛成率とマイナスの関係にあることを確認した。対照的に、取締役会の意見については、記述量および記述に占める数字の割合が賛成率とプラスの関係にあることが確認された。この結果は、賛成率が高いと予想される場合には、取締役会の意見の記述量が増える、あるいは、数字を用いた具体的な記述内容が提示されることで、より説得力のある反対意見が表明されると解釈できる。

しかしながら、このような株主提案議案の全体的な傾向は、議案内容によっては異なる可能性がある。提案内容トピックで議案を分類した場合、提案内容に関しては記述量と賛成率の間にプラスの関係が確認された。さらに、提案理由についても、統計的な有意性は確認されなかったものの、提案理由の記述に占める数字の割合と賛成率の関係は必ずしもマイナスであるとは限らない可能性を示唆する分析結果を得た。本論文で得られた結果は、株主提案の要領が他の株主の議決権行使に影響を及ぼしている可能性があること、そして、株主提案議案の賛成率は要領に記載された3つの文書のトピック、および文書特徴量などの複合的な要因に影響される可能性があることを示している。

本論文の限界として、トピックの分類において時系列的な傾向は考慮していない点を挙げる。企業統治元年と言われた2015年頃から特にコーポレート・ガバナンスに関する提案が増えている可能性はあるが、そのような変化は捉えられていない。今後の研究課題として、株主提案議案の賛成率を予測するモデルの構築を行う際には、より精度の高い分類方法を模索しつつ、トピックの時系列的な変化も捉えられるよう検討したい。

謝辞

本論文はJSPS科研費JP20K13642の助成を受けた研究成果の一部である。また、本論文は日本経済会計学会第3回秋季大会(2022年12月17日)の自由論題報告「自然言語処理を用いた株主提案の要領に関する分析」および言語処理学会第29回年次大会(2023年3月15日)のテーマセッション「株主提案の要領におけるトピック別の文書特徴量と賛成率」の報告内容に一部修正を施したものである。

参考文献

Blei, D. M., A. Y. Ng, and M. I. Jordan (2003) “Latent dirichlet allocation,” *Journal of machine Learning research*, Vol. 3, No. Jan, pp. 993–1022.