

配当金支払金額の探索的データ解析

関西学院大学 商学部 柳 麻衣
関西学院大学 商学部 阪 智香
関西学院大学 商学部 地道 正行

研究概要

本研究では、日経メディアマーケティング社から提供される「日経 NEEDS 財務データ (一般事業会社)」の MT 版をもとに構築されたデータベース KGUSBADES¹⁾ から抽出された東京証券取引所第一部上場企業の財務データ (配当金支払金額, 資本金, 当期利益, 利益剰余金など) を利用した。このデータを、探索的データ解析 (Exploratory Data Analysis) (Tukey (1977)) にもとづいて、データ可視化 (data visualization) による知見から、配当金支払金額を予測するための統計モデリング (statistical modeling) を行った。さらに赤池情報量規準 (cf. Akaike (1973), Konishi and Kitagawa(2008)) によるモデル選択 (model selection) を行った。また、今回の研究結果を文書化する際には、Sweave (Leisch(2002)) を利用することによって動的に生成し、再現可能研究 (reproducible research) を実現することを試みた。

具体的なデータ分析には、配当金支払金額が、企業規模 (資本金) や利益 (当期利益や利益剰余金) によって説明できるのかを考察した。配当金支払金額, 資本金, 当期利益, 利益剰余金の分布を可視化によって確認したところ、歪みが大きいことがわかったため、対数変換を行った。図 1 は、対数スケールでの配当金支払金額のヒストグラムである。この図の結果と歪度の値から、対数をとったものもさらに右に歪んでいることが確認できた。そのため誤差項が非対称分布 (非対称正規分布, 非対称ティー分布) に従う両対数モデルを考察した。ベンチマークとして、正規誤差をもつ両対数モデルも当てはめ比較を行った。説明変数としては資本金, 当期利益, 利益剰余金の 3 つの変数を採用した場合を考えたところ、回帰診断と赤池情報量規準によるモデル選択の結果から、誤差項に非対称ティー分布をもつ両対数モデルが最も良いということが分かった。

参考文献

- [1] Akaike, H. (1973) Information theory and an extension of the maximum likelihood principle, *Proceedings of the 2nd International Symposium on*

¹⁾ Kwansei Gakuin University, School of Administration, Data Extraction System の略。いわゆる、LAMP 環境によって構築されている。詳細については、地道 (2010) を参照されたい。

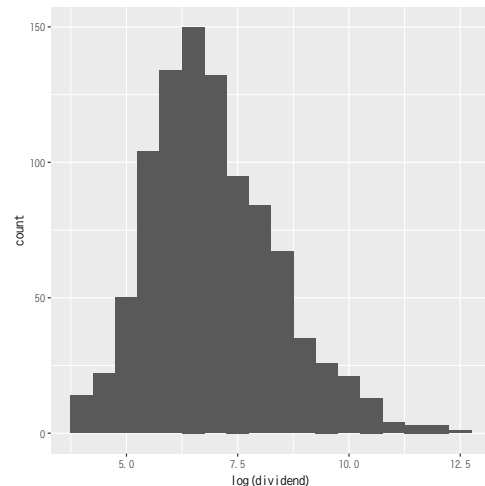


図 1 配当金支払金額のヒストグラム (対数スケール):
歪度: 0.649

- Information Theory*, Petrov, B. N., and Caski, F. (eds.), Akademiai Kiado, Budapest: pp. 267–281.
- [2] Azzalini, A. (1985) A class of distributions which includes the normal Ones, *Scandinavian Journal of Statistics*, Vol. 12, No. 2, pp. 171–178.
- [3] Azzalini, A. with the collaboration of A. Capitanio (2014) *The Skew-Normal and Related Families*, Cambridge University Press, Institute of Mathematical Statistics Monographs.
- [4] 地道 正行 (2010) 『財務データベースサーバの構築』, 関西学院大学リポジトリ, <http://kgur.kwansei.ac.jp/dspace/handle/10236/6013>, ISBN:9784990553005.
- [5] Jimichi, M., Miyamoto, D., Saka, C. and Nagata, S. (2018) *Visualization and Statistical Modeling of Financial Big Data: Log-Linear Modeling with Skew Error*, SSRN: <https://ssrn.com/abstract=3166440>
- [6] Konishi, S. and G. Kitagawa (2008) *Information Criteria and Statistical Modeling*, Springer.
- [7] Leisch, F. (2002) *Sweave: Dynamic generation of statistical reports using literate data analysis*, In Wolfgang Härdle and Bernd Rönz, editors, *Compstat 2002 - Proceedings in Computational Statistics*, pp. 575–580. Physica Verlag, Heidelberg. ISBN 3-7908-1517-9.
- [8] Tukey, J. W. (1977) *Exploratory Data Analysis*, Addison-Wesley Publishing Co.