

氏 名 田上 悠太

学位(専攻分野) 博士(統計科学)

学位記番号 総研大甲第 1919 号

学位授与の日付 平成29年3月24日

学位授与の要件 複合科学研究科 統計科学専攻  
学位規則第6条第1項該当

学位論文題目 Statistical modeling and analysis of defaulted and normal  
states loan's expected loss with Japanese regional banks'  
integrated database

論文審査委員 主 査 教授 川崎 能典  
教授 山下 智志  
教授 黒木 学  
教授 大野 忠士 筑波大学  
教授 批々木 規雄 慶應義塾大学

論文の要旨

Summary (Abstract) of doctoral thesis contents

This thesis composed of five Chapters aims contributing to improvements of risk modeling of bank loans with my integrated database.

Chapter 1 gives backgrounds and outline of this thesis. As required by Basel III, banks are allowed to estimate their portfolio's credit risk. To estimate accurate credit risk, detailed analysis of credit risk (Probability of default (PD), Loss given default (LGD), Exposure at default (EAD)) and development of high accurate estimation models are required. However, because of the shortage of publicly accessible database, especially database containing loss information, credit risk has not been much researched. Further, although there are studies using not publicly accessible database, most of these database contains only limited information and composed of only single bank. Further, there are no publicly accessible bank's loan database that can be used for credit risk research in Japan. To acquire more general results and knowledge, I developed an integrated database composed of five Japanese regional banks' loan data, which contains a great deal of information about factors that might influence the credit risk. With my integrated credit risk database, I tackled three issues in Chapter 2-4.

Chapter 2 developed a model that estimates the probability of return-to-normal-state from default loan to performing loan with logistic regression model to identify the significant factors on return-to-normal-state. There are many researches on credit risk of the non-default state loans. On the other hand, there are only a few researches which study on the credit risk of the default state loans. Studying on credit risk of default state loans is important because default state loans also can cause credit risk. I used corporate loan dataset, which contains information on borrower characteristics, financial characteristics, and time period categorical variables as explanatory variables of the proposed model, from a regional bank recorded every half year from September 2007 to March 2012. I propose a YJ transformation for the explanatory variables, which is an expansion of Box-Cox transformation to negative values. As a result, our result confirmed that although YJ-transformation do not statistically significantly improve model predictive performance, this transformation can improve model predictive performance. Furthermore, I found that debtors with high neglog-transformed number of officers and employees at end of term, model score, YJ-transformed Deposit collateral coverage/EAD, YJ-transformed Land/Total asset and YJ-transformed Notes payable/Total asset are likely to return-to-normal-state. On the other hand, debtors

(別紙様式 2)  
(Separate Form 2)

with high YJ-transformed Long-term working capital/EAD are unlikely to return-to-normal-state.

In Chapter 3, we analyze the effect lending area on regional banks portfolio's credit risk. To develop accurate bank loan's credit risk estimation model, we have to analyze the influencing factors. Existing studies consider the effect of macroeconomic, global factors, country factors, industry factors and so on. The depopulation in provincial areas causes the decline of these area's economies in recent years in Japan. Accordingly, Japanese regional banks have been advancing into areas outside their original operational bases, especially into major cities and urban regions. Since amount of lending to areas outside the original operational base is increasing and such lendings' uncertainty is being high, the lendings are expected to significantly affect the credit risk of the regional banks' lending portfolios. Therefore, we analyze the effects of the lending areas on the credit risk of the regional banks' lending portfolios. Using five regional bank's credit data, we describe the fundamental statistics of each area, we develop credit risk estimation models to analyze the effect of lending area variables. These analysis indicate that lending area variables affect the credit risk, default is less likely to occur outside the local region and in Tokyo, and LGD is small in the local region as compared as outside the local region and in Tokyo.

In Chapter 4, using a dataset of five Japanese regional banks, we propose LGD estimation models that combine a two-stage model with classification and regression tree (CART)-based boosting and SVR with a logistic transformation. We developed (I) Loss occurrence model of two-stage model with Platt scaling model in which we use the output of CART-based Boosting as the predictor and the second component is developed with SVR with a logistic transformation. As a result, we confirm that our model is superior to other models, indicating that a combination of the two-stage model with CART-based boosting and SVR with a logistic transformation works well for LGD estimation.

Chapter 5 describes the conclusion.

Summary of the results of the doctoral thesis screening

田上悠太氏の博士論文審査を、2017年1月23日午後1時30分から約2時間にわたって、本人と5名の委員全員の出席のもとに行い、論文発表会および審査のための会議を行った。

論文は全5章78ページからなり、英語で執筆されている。目的は、デフォルト状態に陥った債権の信用リスクを計測する統計的モデリング法の提案を、個々の貸付債権ごとの情報が利用できる独自のデータベースに基づく実証分析で裏付けながら行うことである。

第1章では、実務的動機としてのバーゼル規制と、銀行貸出に関して信用リスクが最重要であることを確認し、債権の正常状態とデフォルト状態に定義を与えている。一口にデフォルト状態と言っても深刻さの度合いに応じて複数の状態区分があること、従ってデフォルト状態は必ずしも吸収壁とは限らず、正常状態への復帰があり得ることが説明されている。そうした変動の統計的モデリングのために、複数の地方銀行から提供を受けた法人向け貸付債権データを統合し、独自のデータベースを構築したことが述べられている。第2章以降の分析は、全てこのデータベースに基づいている。

第2章では、デフォルトした企業が銀行融資を受け続けることによって正常状態に復帰する現象を分析している。正常復帰の成否を目的変数とし、財務データ、担保情報、保証情報を説明変数とした2項ロジットモデルで分析し、決定要因を論じている。さらにBox-Cox変換を負値変数にも適用可能に拡張したYeo-Johnson変換を導入して、説明力の改善度を報告している。

第3章では、銀行が持つ企業情報の量が信用リスクに与える影響を評価している。具体的には、地方銀行が営業基盤外に対して貸出した融資に着目し、そのデフォルト率と回収率が地元貸出とは異なることを実証的に示した。さらに共変量調整を行うことで、営業基盤外貸出の影響を抽出している。

第4章では、デフォルト時損失の推計精度向上に焦点を当て、複数のモデルの予測精度を比較、評価している。対象は、2項ロジットと一般化線型モデルを用いた2段階推定モデルと、決定木に基づくブースティングにロジット変換サポートベクトル回帰を組み合わせたモデルである。その結果、推計精度や安定性の観点から後者が優れていることが報告され、信用リスク分析においても弱学習器の結合やノンパラメトリックな接近法が精度改善につながる可能性を示唆している。第5章はまとめである。

デフォルト時損失の推計は、貸出先に関する秘匿情報が果たす役割が本質的であるため、比較的未開拓の研究分野である。出願者はデータの収集、データベースの構築から始め、正常復帰の決定要因分析や、企業に関する情報の多寡がデフォルト率や回収率に及ぼす影響について、データに基づく一定の結論を導いているところに最大の貢献がある。技術的には、信用リスク分析における非線形変換の効果、機械学習的接近法の有効性に関して新たな試みを行っている点が評価できる。なお、本論文の第2章の内容は、査読付和文誌「統計数理」(第63巻第1号、出願者が筆頭著者)に掲載済みであり、第3章、第4章は海外の英文誌に投稿、審査中である。また、本論文には含まれないものの、*International Journal of Forecasting*に掲載決定の関連論文(出願者が筆頭著者)がある。

(別紙様式 3)

(Separate Form 3)

総合研究大学院大学複合科学研究科における課程博士及び修士の学位の学位授与に係る論文審査等の手続き等に関する規程第10条に基づいて、口述による試験を実施した。その結果、出願者はその博士論文を中心としてそれに関連がある専門分野、その基礎となる分野および英語の能力について博士（統計科学）の学位の授与に十分な学識を有するものと判断し、合格と判定した。