



Title	Empirical Characterization of Risky Software Projects Using Logistic Regression Analysis
Author(s)	高木, 徳生
Citation	大阪大学, 2005, 博士論文
Version Type	
URL	<a href="https://hdl.handle.net/11094/45738">https://hdl.handle.net/11094/45738</a>
rights	
Note	著者からインターネット公開の許諾が得られていないため、論文の要旨のみを公開しています。全文のご利用をご希望の場合は、<a href="https://www.library.osaka-u.ac.jp/thesis/#closed">大阪大学の博士論文について</a>をご参照ください。

*The University of Osaka Institutional Knowledge Archive : OUKA*

<https://ir.library.osaka-u.ac.jp/>

The University of Osaka

氏 名	高 木 徳 生
博士の専攻分野の名称	博 士 (情報科学)
学 位 記 番 号	第 19613 号
学 位 授 与 年 月 日	平成 17 年 3 月 25 日
学 位 授 与 の 要 件	学位規則第 4 条第 1 項該当 情報科学研究科情報システム工学専攻
学 位 論 文 名	Empirical Characterization of Risky Software Projects Using Logistic Regression Analysis (ロジスティック回帰分析による混乱ソフトウェアプロジェクトの実証的判別法の開発)
論 文 審 査 委 員	(主査) 教 授 菊 野 亨  (副査) 教 授 井 上 克 郎    教 授 尾 上 孝 雄    助 教 授 楠 本 真 二

### 論 文 内 容 の 要 旨

ソフトウェア開発プロジェクトでは、しばしばプロジェクトの実行が危うくなる状態に陥ることがある。もし、このような危険な事態を引き起こす要因（リスク）を見逃したり、その対処が遅れたりすると、プロジェクトは制御不能となり混乱した状態に陥ってしまう可能性がある。ソフトウェアが社会システムにおいて重要な位置を占める中で、そうしたプロジェクトの混乱を防止するために、ソフトウェアプロセスの改善に取り組む組織が増えてきている。

本論文では、ソフトウェアプロセス改善に係わる統計的分析手法を用いた 2 つの実証的研究について述べる。まずはじめに、「プロジェクト計画の妥当性」が「プロジェクトのコスト見積りと実績に関する逸脱の度合い」にどの程度影響を与えるか、さらに「プロジェクトの生産性や品質」に及ぼす影響についての分析を試みた。この分析の結果、プロジェクト計画の妥当性が高いほど、プロジェクトコスト見積りと実績の逸脱の度合いが低く、また、同時に生産性や品質が高くなることを統計的に示すことができた。

次に、実際のプロジェクトから得られるデータと統計的分析手法を用いて混乱プロジェクトを実証的に判別する新しい手法を提案する。提案する手法には、次の特徴がある。

実践的な質問表を使用していること：混乱プロジェクトの特性を抽出するために開発現場の経験や教訓に基づく簡単な質問表を設計した。質問表の一般性を確認するため、他の研究で提案されているソフトウェアリスク項目との関連性を評価した。

実証的アプローチであること：統計的モデルの構築および統計的モデルの検証には、実際のプロジェクトから収集したデータを用い、実証的な検証を行った。

はじめに、混乱プロジェクトの特性を抽出するため、5 つの観点と 22 の質問項目（混乱抽出要因）からなる質問表を設計した。5 つの観点とは「要件」「見積り」「計画策定」「チーム組織」および「プロジェクト管理活動」である。この質問表を用い、ある組織で実行された 32 のプロジェクトから評価データを収集した。次に、ソフトウェアメトリクスやインタビューなどで得たプロジェクトに関する情報を用い、32 のプロジェクトを「混乱あり」「混乱なし」の 2 グループに分類した。そして、得られた質問表の回答（即ち、プロジェクトの特性）とプロジェクトの混乱の程度と関係について最尤推定法に基づく標準的な分析手法であるロジスティック回帰分析を行い、混乱プロジェ

クトを判別する統計的モデルを構築した。

構築した統計的モデルの妥当性を検証するために、モデル構築に使用した 32 プロジェクトに対し、統計的モデルを使った混乱予測結果と実際のプロジェクトデータとの比較評価を行った。この結果、32 プロジェクト中で 28 プロジェクトの判別が正しく行われることを確認した。さらに、この統計的モデルによる混乱の危険性があるプロジェクトの判別能力に関する実証的評価を行うため、同じ組織の別の 8 プロジェクトに対してモデルの混乱予測を実施し、その結果と実プロジェクトデータを比較評価した。その結果、8 プロジェクト中の 7 プロジェクトの判別が正しく行われることが確認できた。

以上のように、本研究で提案する手法によりソフトウェア開発における混乱プロジェクトを判別する統計的モデルが構築できることを実証した。現在、本研究で手法は実際の開発組織でのプロジェクトリスク管理へ応用することを行っており、さらにリスクの動的な変化を捉える手法への展開を検討中である。

### 論文審査の結果の要旨

本論文は、ソフトウェア開発におけるプロジェクト管理に対し、情報科学的な手法の導入を試みたものである。基本的には情報科学的な手法の基礎となるモデル作成に関する理論的考察と実際の開発データを利用したモデル構築とその評価から構成されている。

モデル作成に当たっては、医療分野等への応用の成功で最近特に注目を集めているロジスティック回帰分析を採用している。分析の目標を混乱ソフトウェアプロジェクトの判別に置いている点で極めて挑戦的である。分析に利用するデータはプロジェクトの現場からアンケート方式で収集する。それは混乱プロジェクトの開発現場では直感的にその兆候が開発者と管理者によって観測されているという調査結果に基づいた決定である。

モデル作成のキーとなるアンケートで使用する質問表には特別の工夫をしていて、開発現場での経験や教訓がバランス良く反映されている。得られた質問表は「要件」「見積」「計画策定」「チーム組織」「プロジェクト管理活動」の 5 つの視点から構成されている。その下に、全部で 22 個の混乱要因が具体的な質問項目として用意されている。このような具体的な質問表の設計は従来研究では困難であるとして行われていない。その意味でこの質問表自体にも高い実用的な価値がある。

モデル作成には 32 個の実際の開発プロジェクトを用いている。まず、32 個を「混乱あり」と「混乱なし」の 2 つのグループに分けておいて、質問表の回答（これがプロジェクトの特性を表す）とプロジェクトの混乱のあり、なしとの間の関係をロジスティック回帰分析で明らかにした。その関係を表現する回帰式が求めるべき混乱ソフトウェアプロジェクトの判別モデルである。

引き続いて、モデルの妥当性評価を、上記とは別の 8 個の実際の開発プロジェクトを用いて行った。評価の手順としては、これら 8 個のプロジェクトの質問表の回答を判別モデルに適用して得られる混乱予測とプロジェクトの実際の（混乱のあり、なしに関する）結果との一致の度合いを調べている。その結果、8 個の内の 7 個について一致することが確認され、有効性の高さが示された。

以上のように、本論文で提案する実証的判別法によって混乱ソフトウェアプロジェクトを十分に高い精度で予測できる可能性が示されており、ソフトウェアプロジェクト管理手法の開発に貢献するところが非常に大きいだけでなく、学術的にも非常に意味のある幾つかの知見をもたらしている。

よって、博士（情報科学）の学位論文として価値のあるものと認める。