

ソフトウェアメトリクス ーソフトウェア開発データの分析技法

平成26年度シラバス

2014年1月10日

国立情報学研究所

トップエスイープロジェクト

代表者 本位田 真一

1. 科目名

ソフトウェアメトリクスーソフトウェア開発データの分析技法

2. 担当者

野中 誠、阿萬 裕久、水野 修

3. 本科目の目的

本科目の目標は、ソフトウェア開発、プロジェクト管理、品質マネジメントに関する各種データについて、統計解析ソフトウェアを用いて自ら分析できる知識とスキルを修得することである。ただし、ソフトウェア開発データの分析そのものは本来の目的ではない。ソフトウェア開発のマネジメントにおいて「事実にもとづく管理」を実現する手段である。ソフトウェア開発データを分析することにより、プロダクトやプロジェクトの状況を定量的に把握し、何らかの対策をとる必要性をデータで示せるようになる。これは、品質向上のために組織的なプロセス改善を進める上での重要な原動力の一つになる。

ソフトウェア開発データを正しく分析するには作法があり、必要な知識とスキルがある。統計手法（検定、相関分析、回帰分析など）はその一つである。ただし、ソフトウェア開発データの分析において高度な統計手法を必要とする場面は多くなく、データの可視化方法を工夫すれば十分な場合もある。統計手法にとらわれすぎず、分析目的を見誤らないことが何よりも大事である。このような点も含めて、ソフトウェア開発データを測定し分析する際の留意点なども議論する。

一方で、ソフトウェア開発データの分析における視野を広げることも重要である。最近の研究成果を紹介し、受講生とともに議論する機会も設ける。

4. 本科目のオリジナリティ

ソフトウェア開発に関するデータを用いて統計解析を行う講義科目は、講師が知る限りほとんどない。類似のセミナーもあるが、15回の講義という時間をかけて実施されているものはない。また、気鋭の研究者を招いて研究内容を議論する場を提供している講義科目もほとんどないだろう。こうした場を提供しながら、ソフトウェア開発データの分析技法を演習主体で学べることが本科目のオリジナリティである。

5. 本科目で扱う難しさ

ソフトウェア開発マネジメントの実務において、「事実に基づく管理」「データに基づく意思決定」は容易ではない。そもそも信頼できるデータが収集できていなかったり、せっかくデータを収集しても分析されずに放置されていたりすることが一因である。本科目では、これらの課題に対処すべく、ソフトウェア開発データに対して適切な分析技法を用いて、意思決定に役立つ有益な情報を得るプロセスを扱う。

6. 本科目で習得する技術

本科目で習得目標とする技術は次の通りである。詳細は「8. 講義計画」を参照。

- ・ ソフトウェアプロダクトおよびプロセスに関するメトリクス
- ・ ソフトウェア開発データの可視化・分析手法
- ・ ソフトウェア開発データに基づく予測モデルの構築手法

7. 前提知識

次の項目について、概要を理解した上で受講することが望ましい。

- ・ 基本的なプロダクトメトリクス（ソースコード行数、サイクロマチック複雑度）
- ・ オブジェクト指向設計の基本概念（クラス、インスタンス、メンバ変数、メソッド等）
- ・ 統計学の基礎（基本統計量、相関と回帰、正規分布、検定）

8. 講義計画

節番号は教科書の該当箇所を指している。都合により順番が変更になる場合がある。

第1週：イントロダクション – 品質データ分析の基本（1.1節）

第2週：基本的なメトリクス – 欠陥・工数・規模の測定（6.1～6.3節）、その他

第3週：品質状況の可視化(1) – 工数推移（2.1節）、開発終盤の不具合の可視化（2.2節）

第4週：品質状況の可視化(2) – レビュープロセスの制御と評価（2.3節）

第5週：影響要因の把握(1) – レビュー欠陥指摘数に関わるメカニズムの把握（3.1節）

第6週：影響要因の把握(2) – リリース後品質を予見するメトリクス（3.2節）

第7週：工数見積りモデルの構築(1) – 工数見積りの妥当性確認（4.1節）

第8週：工数見積りモデルの構築(2) – 多変量データに基づく工数見積りモデル

第9週：不具合検出数の予測(1) – ソフトウェア信頼性モデルの適用

第10週：不具合検出数の予測(2) – 多変量データに基づく欠陥数予測モデル（4.2節）

第11週：fault-proneモジュール予測(1) – 判別分析による方法

第12週：fault-proneモジュール予測(2) – 設計メトリクスによる方法（4.3節）

第13週：研究テーマに関する紹介と議論（1）

第14週：研究テーマに関する紹介と議論（2）

第15週：まとめ

9. 教育効果

データを実際に分析しながら各種技法を学ぶことで、理解を深めることができる。また、演習で用いるデータはいずれもソフトウェア開発に関わるデータであるため、実務での応用に結びつけることができる。

10. 使用ツール

- R：統計解析ツール（フリーソフトウェア）。
- Microsoft Excel：表計算ソフトウェア。

11. 実験及び演習

ノートパソコンに使用ツールをインストールし、実際にデータを解析する。

12. 評価

出席日数とレポート課題への取り組み状況を総合して評価する。

13. 教科書/参考書

教科書は次を指定する。事前に予習をした上で講義に臨んでいただきたい。

- 野中誠・小池利和・小室睦, データ指向のソフトウェア品質マネジメント, 日科技連出版社, 2012.

参考書は次の通り。

- Linda M. Laird and M. Carol Brennan, 野中誠・鷲崎弘宜訳, 演習で学ぶソフトウェアメトリクスの基礎, 日経 BP 社, 2009.
- Norman E. Fenton and Shari Lawrence Pfleeger, Software Metrics: A Rigorous and Practical Approach, 2nd edition, Course Technology, 1998.
- Stephen H. Kan, 古山恒夫・富野壽訳, ソフトウェア品質工学の尺度とモデル, 構造計画研究所, 2004.
- Katrina D. Maxwell, Applied Statistics for Software Managers, Prentice Hall PTR, 2002.