

# ソフトウェア開発データの分析パターンの提案と評価

服部 昇<sup>†</sup> 副島 千鶴<sup>†</sup> 山本 修一郎<sup>†</sup>

ソフトウェア開発を工業的なものとするためには、ソフトウェア開発データを分析することで事実の推測を科学的に行い、生産性や品質等に影響を与える要因を明らかにしていく取り組みを、開発現場でも実践していく必要がある。そのためには、ソフトウェア開発データの分析の進め方に対応したデータ分析パターンを用意、展開するのが有効と考えられる。本論文では、データ分析方法を仮説、モデル式、検証方法、結果の解釈からなるデータ分析パターンとして整理することを提案するとともに、ソフトウェア開発データの分析に共通的に利用できると思われる10個のデータ分析パターンを提案する。評価の結果、提案するデータ分析パターンはソフトウェア開発に関する様々な仮説検証に適用可能であること、また、仮説検証の効率化に効果があることを確認できた。

## Effect of Using Data Analysis Patterns for Software Development

NOBORU HATTORI,<sup>†</sup> CHIZURU SOEJIMA<sup>†</sup> and SHUICHIRO YAMAMOTO<sup>†</sup>

We propose using data analysis patterns consisting of hypotheses, model formulae, verification methods and validation for use in analyzing software development project data sets. This will enable development organizations to identify the factors affecting software productivity and quality scientifically, which will promote the industrialization of software development. We also propose ten data analysis patterns that should be commonly applicable to the analysis of such data sets. Evaluation of these ten patterns demonstrated that they can be applied to various types of hypothesis verification and that they make hypothesis verification more effective.

### 1. はじめに

情報システムが経済社会の様々な活動を支える基盤となっている今、ソフトウェア開発の生産性の向上や安定した品質の確保が大きな課題となっている。このため、ソフトウェアの開発を属人的なものから工業的なものとするための取り組みへの期待が高まっている。日本においても、たとえば独立行政法人情報処理推進機構 (IPA) の傘下にソフトウェアエンジニアリングセンター (SEC)<sup>1),2)</sup> が立ち上げられ、国の取り組みとしてソフトウェア開発の工業化の促進が開始されている。ソフトウェアエンジニアリングセンターの中には、ソフトウェア産業においても定量データに基づく科学的なマネジメントを浸透、発展させることを目指した定量データ分析部会が立ち上げられ、国内の複数企業から収集したソフトウェア開発データをもとに、工数、工期、規模等の基本的なメトリクスの分析を行ったデータ分析白書の発行等の取り組みが開始されている。

現在、ソフトウェア開発を行う企業・組織における定量的な取り組みは、収集されたメトリクスを生データのままグラフ等で確認する等の形態にとどまっていることが少なくない。もちろん、それらは定量的取り組みの基本であり、開発プロセスの改善に一定の効果を発揮すると考えられる。しかし、ソフトウェア開発を工業的に行う取り組みを広く普及させるためには、さらに開発現場において統計的なデータ分析手法を用いて生産性や品質等に影響を与える要因を科学的に明らかにする取り組みを広く普及させることが必要であると考えられる。統計的なデータ分析手法のソフトウェア開発現場への普及のためには、データ分析の手法を単独ではなく、対応する仮説や、分析の結果得られる結果の解釈等と組み合わせるパターン化することが有効であると考えられる。さらに、用意するデータ分析のパターンは、ソフトウェア開発データの分析に共通的に用いることができるものに絞り込む必要があると考えられる。

本論文では、ソフトウェア開発データ分析の進め方に対応した、仮説、モデル式、検証方法、結果の解釈の4つの項目からなるデータ分析パターンを提案す

<sup>†</sup> 株式会社 NTT データ  
NTT DATA Corporation

る。また、開発データ分析において共通して用いることができると考えられる 10 個のデータ分析パターンを提案する。さらに、実際のソフトウェア開発データを用いて開発コストおよび品質に関する分析を行った際の適用例を示すことで、10 個のデータ分析パターンの有効性を確認する。

本論文の構成を以下に示す。2 章において、本研究の課題設定の背景となる関連研究について触れる。3 章で、ソフトウェア開発データ分析の進め方に対応したデータ分析パターンを提案する。4 章で、ソフトウェア開発データの分析に有効と考えられる 10 個のデータ分析パターンを提案する。5 章で、4 章のデータ分析パターンを、実際のソフトウェア開発データを用いた仮説検証にあてはめた例を示す。6 章で評価を行う。評価として、これらのパターンを実際にソフトウェア開発データにあてはめることによる分析作業の生産性の向上について確認する。7 章では、データ分析パターンの適用による生産性向上についてより詳細な考察を行うとともに、今回提案したデータ分析パターンの十分性について議論する。最後に 8 章で、まとめおよび今後の課題を述べる。

## 2. 関連研究

ソフトウェア工学の分野における実証的研究が他の分野に比べて遅れており、これを改善すべきであると指摘する調査研究は、以前から存在する。Tichy ら<sup>3)</sup> は、400 件以上の研究論文を調査した結果、コンピュータサイエンスの分野では実証実験による妥当性の確認が行われていない論文が 40% 以上あるが、特にソフトウェア工学の分野ではこれが 50% 以上になる。これに対して、比較対象として光学分野およびニューラルネットワークを例にあげると、それぞれ 15%、12% 程度である。理由の 1 つはコンピュータサイエンスに関する実験が、ヒューマンファクタが関与する場合が多くあることだが、心理学者はすでに人間を対象とした実験を行う技術を持っている。コンピュータ科学者も、この分野の発展のために、実証的研究をもっと行う必要があるとしている。

Zelkowitz ら<sup>4)</sup> はソフトウェア工学の分類法を提案し、600 件以上の論文を分類しているが、その中でも同様に、ソフトウェア工学における実証的研究を増やす必要があると主張している。

2006 年には、Zannier ら<sup>5)</sup> により、ICSE (International Conference on Software Engineering) の過去 29 年あまりの予稿集を評価した報告が行われている。これによると、実証的研究の増加という量的な面

では有意に状況が改善されているものの、質的な面では仮説や目標の記述がないもの等も存在し、引き続き改善が必要なが指摘されている。

Perry ら<sup>6)</sup> は、ソフトウェア工学分野における実証的研究において、仮説だけでなく結論の導出が十分でないという問題点を指摘している。仮説の検証は、(1) 仮説の設定、(2) 状況の観察、(3) 観察のデータへの抽象化、(4) データ分析、(5) 仮説検定による結論の導出で進められるが、このうち、最後のステップである (5) 仮説検定による結論の導出が最も重要であり、かつ最も改善の余地がある。また、実証的実験は理論と実践の対比であり、つねに仮説が必要であるが、実際には仮説が明確でない研究が見られる。また、実証的実験は高価であり時間もかかるため、コストを下げるためにも実験計画の効率化や、より焦点を絞った問題にあたる必要があるとともに、信頼性の向上のために他者が結果を再現することを支援できる方法を見つける必要があるとしている。

実証的研究の裾野を広げる必要性については、Dyba ら<sup>7)</sup> も問題提起を行っている。ここでは、医療分野の EBM (Evidence-based Medicine) にない、EBSE (Evidence-based Software Engineering) に取り組むべきであると主張している。ソフトウェア技術者は、新しい技術の採用において、技術の効果についての科学的証拠を得て誤った判断をしないようにする必要があり、そのために実証的研究や研究の評価が必要である。またソフトウェア工学分野では、実験室での実験だけでなく、産業界の実プロジェクトでの実証も試みるべきである。しかし、すでに研究実績が豊富な医療分野に比べ、ソフトウェア工学の研究は蓄積が少ない現状に鑑みると、研究者だけでなく実務者もより積極的にこの分野の実証に関与すべきであるとしている。

実証的ソフトウェア工学の研究を広げるための取り組みとして、ガイドラインの必要性を指摘するものもある。Sjöberg ら<sup>8)</sup> は、ソフトウェア工学の制御実験に関する 5,453 の論文について傾向を調査した結果として、報告があいまいで系統だっていないこと等を指摘している。高品質で、事実に基づいたソフトウェア工学の実験を実施、報告する際の方法論および現実の複雑さの扱いを支援するためのガイドラインが必要である。実験の報告がより統一的になされるのであれば、論文のレビュー、実験の複製、メタアナリシス、理論の構築がより改善できるであろうとしている。

また、Kitchenham ら<sup>9)</sup> は、自らの経験および医療分野の研究ガイドラインを参考に、実験の背景、実験計画、実験の実施およびデータ収集、分析、結果の提

示、結果の解釈の6つのトピックからなるガイドラインを提示している。このうち、分析のガイドラインとしては、複数のテスト結果を調整する手順の具体化、分析を望ましい結果に誘導しないような blind analysis、外れ値や多重共線性等の影響を確認する sensitivity analysis、データが検定の前提条件を満たしているかの確認の実施があげられている。これは、あるべき姿を示す指針であり、分析の実践を容易化するものにはなっていないと考える。

ソフトウェア開発データに関するデータ分析のフレームワークについては、Liu ら<sup>10)</sup>の取り組みがある。コスト予測には、COCOMO モデル等が広く用いられているが、プロジェクト実績データの多くは企業により異なるため、実際にこれを適用する前の準備として、各企業等が実施すべき手順をフレームワークとして提案している。フレームワークは、データ可視化、対数変換等のデータ変換、相関分析、回帰分析、ステップワイズ分散分析、説明変数間の独立性の確認、モデル式の解釈・引用、残差テストのステップからなるが、これはコスト予測モデルを現場に適用する前処理としてのフレームワークであり、仮説検証により事実の推測を科学的に行うためのものではない。

統計学の分野では、Unwin<sup>11)</sup>がデータ分析のパターン化について述べている。統計学が理論であるのに対して、データ分析は方法論である。このため、有効なデータ分析のパターンをボトムアップ的に用意し、繰返し使用を可能とすることを提起している。ただし、数個のデータ分析手法をパターンにあてはめたとときの記述例について示しているのみであり、特定の分野において実際に必要になるパターンの抽出、適用性や効果の検証等については特に議論されていない。

Dyba<sup>12)</sup>はソフトウェア開発に関する120組織を調査し、開発データの計測がプロセス改善に対して有効であることを統計的に実証した。そしてソフトウェア開発の進歩性が、開発プロセスと生産物の計測、ならびにその結果に基づく判断能力に依存することを示唆していると指摘した。

Gopal ら<sup>13)</sup>は、計測データの単なる表示・確認を上回る統計的手法等の洗練されたデータ分析手法が、ソフトウェア企業でのメトリクスプログラムの導入に好影響を与えることを実証的研究により示した。メトリクスプログラムがソフトウェア企業に受け入れられるためには、このほかにメトリクスの定期的、効率的な収集、結果のステークホルダへの報告等も影響をあたえるが、それらを実施するためには、組織の力や、マネジメント層のコミットメントが必要であることを

実証的に示している。

以上をもとに、本研究の位置づけを以下に整理する。

ソフトウェア工学分野における実証的研究を質、量ともに改善する必要性については、以前から指摘がなされているが、今なお十分な改善がなされているとはいえない。特に、ソフトウェア工学は実践的な学問・技術体系であるため、実験室だけでなく、産業界の実プロジェクトからも実証結果を得て、理論と突き合わせることが必須である。このためには、事実の推測を科学的に行うためにデータ分析の手法を広く展開し、実プロジェクトを対象とした分析を増やすことが有効であると考えられる。ソフトウェア工学分野においてデータ分析手法の普及・展開を行うためのこれまでの取り組みとしては、実証的研究の指針を示すガイドラインや、コスト予測の前処理のためのフレームワークが存在する。しかし、本論文では新たに、開発現場での仮説検証を目的としたデータ分析手法の普及・展開のために、データ分析手法を、仮説、モデル式、結果の解釈を含むパターンとして用意することを提案している。さらにソフトウェア工学の分野で実証の取り組みを広げるために有効と考えられる具体的なデータ分析パターンの提案と、それらを実プロジェクトでの仮説検証に適用した際の評価を行っている。データ分析のパターンに仮説やモデル式、結果の解釈を含むことで、データ分析と仮説や結論の関係の明確化や正しいモデル式、分析手法の適用等を通して、実証的研究の品質の向上にもつながることが期待できる。また、データ分析手法の開発現場への普及・展開は、開発現場におけるメトリクスプログラムの充実等を通して科学的マネジメントの浸透、発展にも寄与できると考えられる。

### 3. ソフトウェア開発データ分析の進め方に対応したデータ分析パターンの提案

今回提案するデータ分析パターンは、ソフトウェア開発データ分析の進め方に対応した、以下の4つの項目からなる。

【仮説】仮説の記述

【モデル式】仮説を検証するため、開発データに関して成立すると想定される不変式をモデル式として定義する。

【検証方法】モデル式を科学的にデータに基づいて検証するために必要となる統計的な手法を説明する。

【結果の解釈】モデル式が成立する場合と成立しない場合のそれぞれの意味と、推定される原因を示す。

このようなデータ分析の進め方に沿ったデータ分析パターンをあらかじめ用意し、それをソフトウェア開

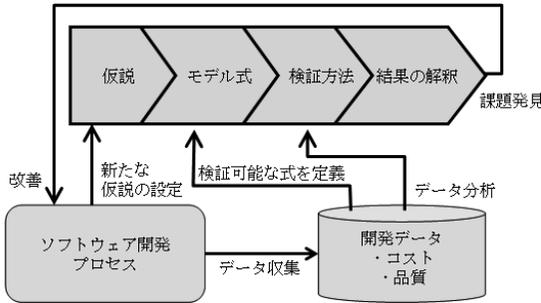


図 1 ソフトウェア開発データ分析の進め方とソフトウェア開発プロセス、開発データの関係

Fig. 1 Relationships among hypothesis verification, software development process and its data.

開発データの分析にあてはめることで、仮説設定や分析作業、得られた結果の解釈が容易になり、その結果データ分析作業が効率化されることが期待できる。これにより、ソフトウェア開発の現場へもデータ分析手法が普及、展開しやすくなることが期待できる。ソフトウェア開発データ分析の進め方とソフトウェア開発プロセス、開発データの関係について図 1 に示す。

#### 4. 具体的なソフトウェア開発データ分析パターンの提案

本章では、ソフトウェア開発分析で共通的に使用できると思われるデータ分析パターンを示す。データ分析パターンは、開発現場で利用されることを目的としているので、パターンの利用者にとってできるだけ容易なものとする必要がある。そのためには、1 つの仮説を検証するためには、1 つのデータ分析パターンのみを適用すればよいようにすることが有効と考える。

このため、他のデータ分析パターンと内容が重複する部分も、データ分析パターンごとに、別々に用意するものとする。表 1 に提案する 10 個のデータ分析パターンを示す。

##### 4.1 変数間の関係の確認に関するデータ分析パターン

###### パターン 1

パターンの目的：2 変数  $X, Y$  の相関関係を確認する。

パターン利用時に具体化する情報： $X, Y$

【仮説】 $X$  が大きくなると、 $Y$  は大きく（小さく）なる傾向にある。

【モデル式】 $Y = kX$

【検証方法】

適用手法 相関分析、具体的には、 $X, Y$  の分布により、以下のいずれかを用いる。

(ア)  $X, Y$  がともに正規分布と大きく異なる

場合：ピアソンの相関係数

(イ)  $X, Y$  のいずれかが正規分布から大きく異なる分布の場合：スピアマンの順位相関係数（ノンパラメトリック検定）

着目する指標

- $X$  と  $Y$  の相関係数 ( $\rho$ )

有効な可視化ツール 散布図

注意すべき点 事前に  $X$  と  $Y$  の分布を確認し、用いる相関係数の種別を決定すること。

【結果の解釈】

- ①  $\rho$  が 0 より有意に大きい場合： $X$  が大きくなると  $Y$  も大きくなる傾向にある。
- ②  $\rho$  が 0 と有意に異なる場合： $X$  と  $Y$  の関係には特に傾向は見出せない。
- ③  $\rho$  が 0 より有意に小さい場合： $X$  が大きくなると  $Y$  は小さくなる傾向にある。

###### パターン 2

パターンの目的：2 変数  $X, Y$  の比例関係を確認する。

パターン利用時に具体化する情報： $X, Y$

【仮説】 $Y$  は  $X$  に比例する

【モデル式】 $Y = \beta_0 + \beta_1 X$

【検証方法】

適用手法 単回帰分析

着目する指標

- 決定係数 ( $R^2$ ):  $Y$  を  $X$  でどれくらい説明できるかを示す指標である。たとえば、 $R^2 = 0.9$  なら、 $X$  のみで  $Y$  の 90% は説明できることになる。決定係数は相関係数の 2 乗の値となる。
- $X$  の係数 ( $\beta_1$ ): 単位  $X$  あたりの  $Y$  の推定値となる。

有効な可視化ツール 散布図

注意すべき点  $X$  や  $Y$  のデータの一部に、極端に大きな値が含まれる場合、決定係数は過大評価されやすい。その場合、 $X$  や  $Y$  の大きさにより分析範囲を層別する等の作業が必要になることがある。回帰分析に先立ち、 $X$  や  $Y$  の分布、ばらつき等を散布図で概観することが必要である。

【結果の解釈】

- ①  $K < R^2 \leq 1$  の場合： $Y$  は  $X$  により十分に説明できる。多くの観測対象において、単位  $X$  あたりの  $Y$  の値はあまり変わらない値である。また、回帰直線から大きく離れた観測点を発見し、その原因を詳細調査する等の活用ができる。
- ②  $0 \leq R^2 \leq K$  の場合： $Y$  は  $X$  ではあまり説明できない。観測対象により、単位  $X$  あたりの  $Y$  の値にばらつきがある。

表 1 提案するソフトウェア開発データ分析のための 10 個のデータ分析パターン  
Table 1 Ten proposed data analysis patterns for analyzing software development project data sets.

パターン ID	パターン名	分類	仮説	モデル式
パターン 1	相関	変数間の関係の分析	$X$ が大きくなると、 $Y$ は大きく（小さく）なる傾向にある	$Y = kX$
パターン 2	比例		$Y$ は $X$ に比例する	$Y = \beta_0 + \beta_1 X$
パターン 3	べき乗比例		$Y$ は $X$ のべき乗に比例する	$Y = e^{\beta_0} \cdot X^{\beta_1}$
パターン 4	2 群の比例関係の違いの確認		$X$ が $Y$ に与える影響は、グループ A の方がグループ B よりも大きい（小さい）	$Y = \beta_{00} + \beta_{01}D + \beta_{10}X + \beta_{11}DX$
パターン 5	2 群のべき乗比例関係の違いの確認		$X$ が $Y$ に与える影響は、グループ A の方がグループ B よりも大きい（小さい）	$Y = e^{(\beta_{00} + \beta_{01}D)} \cdot X^{(\beta_{10} + \beta_{11}D)}$
パターン 6	重回帰		$Y$ は $X_1, X_2, \dots, X_n$ に比例する	$Y = \beta_0 + \sum \beta_i X_i, \quad i = 1, 2, \dots, n$
パターン 7	べき乗重回帰		$Y$ は $X_1, X_2, \dots, X_n$ のべき乗に比例する	$Y = e^{\beta_0} \cdot \prod X_i^{\beta_i}, \quad i = 1, 2, \dots, n$
パターン 8	2 群の値の比較	属性が与える影響の分析	値 $Q$ について、グループ A とグループ B を比較するとグループ A の方が大きい	$Q_A > Q_B$
パターン 9	2 群のばらつきの比較		ある値 $Q$ について、グループ A とグループ B のばらつきを比較するとグループ A の方が大きい	$V_A > V_B$
パターン 10	2 群の危険度の比較		事象 P が発生する危険度は、グループ A の方がグループ B よりも大きい	$\frac{R_A}{R_B} > 1$

①のときは、決定係数  $R^2$  の値が有意であること、すなわち該当の回帰モデル式が有意であることを確認すること。有意でない場合は、標本の追加を検討すること。

なお、 $K$  は結果の解釈を分ける閾値である。 $K$  の具体値は分析対象の特性や、有識者の意見等を総合的に考慮しながら決定する必要がある。

**パターン 3**

パターンの目的：2 変数  $X, Y$  のべき乗の関係を確認する。

パターン利用時に具体化する情報： $X, Y$

【仮説】 $Y$  は  $X$  のべき乗に比例する

【モデル式】 $Y = e^{\beta_0} \cdot X^{\beta_1}$

【検証方法】

適用手法  $X$  および  $Y$  を対数変換した後に単回帰分析を実施する。

着目する指標

- 決定係数 ( $R^2$ )
- $X$  のべき乗の係数 ( $\beta_1$ ): これは  $X$  が 1% 増加したとき  $Y$  が何%増加するかを示す値となる。

有効な可視化ツール 散布図

注意すべき点  $X$  や  $Y$  の値が 0 の場合、対数変換が実施できない。この場合、各変数に無条件に 1 を足してから対数変換を実施する。

【結果の解釈】

(ア) モデルの説明力に着目する場合

- ①  $K < R^2 \leq 1$  の場合：多くの観測対象におい

て、 $X$  が 1% 増加したとき  $Y$  が何%増加するかはあまり変わらない。

- ②  $0 \leq R^2 \leq K$  の場合：観測対象により、 $X$  が 1% 増加したとき  $Y$  が何%増加するかにばらつきがある。

上記①のときは、決定係数  $R^2$  の値が有意であることを、すなわち該当の回帰モデル式が有意であることを確認すること。有意でない場合は、標本の追加を検討すること。

なお、 $K$  は結果の解釈を分ける閾値である。 $K$  の具体値は分析対象の特性や、有識者の意見等を総合的に考慮しながら決定する必要がある。

(イ)  $X$  と  $Y$  の関係に着目する場合

- ①  $\beta_1 \cong 0$  の場合： $X$  の値にかかわらず  $Y$  は不変となり、 $X$  と  $Y$  の値の間には特に関係がない。
- ②  $0 < \beta_1 < 1$  の場合： $X$  が大きくなるほど、 $X$  の増加による  $Y$  の増加の割合は小さくなる。
- ③  $\beta_1 \cong 1$  の場合： $X$  の増加による  $Y$  の増加の割合は一定である。
- ④  $\beta_1 > 1$  の場合： $X$  が大きくなるほど、 $X$  の増加による  $Y$  の増加の割合は大きくなる。

図 2 に、 $\beta_1$  の値の違いによる、 $X$  と  $Y$  の関係の違いを示す。

**パターン 4**

パターンの目的： $X$  と  $Y$  の間に比例関係があてはまるとき、すなわちパターン 2 を適用した分析で  $Y$  は  $X$  により十分に説明できるという結果が得られたと

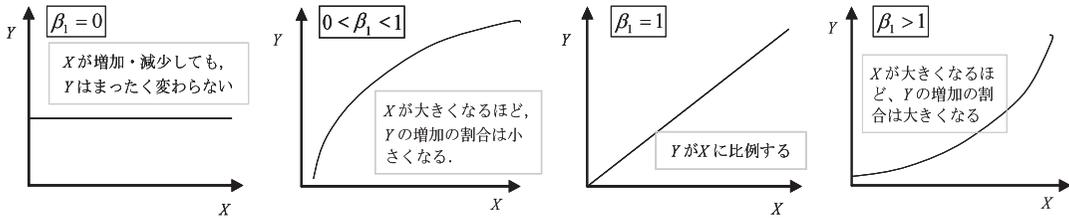


図 2  $Y = e^{\beta_0} \cdot X^{\beta_1}$  のときの、 $\beta$  の値による  $X$  と  $Y$  の関係の違いのイメージ  
 Fig.2 Graphs relating  $X$  and  $Y$  in given value of  $\beta$  in  $Y = e^{\beta_0} \cdot X^{\beta_1}$ .

き、グループ A とグループ B で、 $X$  が  $Y$  に与える影響が異なるかどうかを確認する。また、異なる場合に、どのように異なるのかもあわせて確認する。パターン利用時に具体化する情報： $X, Y, \text{グループ A}, \text{グループ B}$

【仮説】 $X$  が  $Y$  に与える影響は、グループ A の方がグループ B よりも大きい (小さい)。

【モデル式】 $Y = \beta_{00} + \beta_{01}D + \beta_{10}X + \beta_{11}DX$   
 $D$  : ダミー変数  
 (グループ A のとき 1, グループ B のとき 0)

【検証方法】

適用手法 重回帰分析

着目する指標

- $D \cdot X$  の係数 ( $\beta_{11}$ ): 0 と有意に異なるかどうかを確認する。

有効な可視化ツール 散布図

注意すべき点 残差 ( $Y$  の推定値と観測値の差) がばらついていると、係数  $\beta$  の有意性の計算が正しく行われなくなる。残差グラフ、散布図等により、残差がだいたい同じ値であることを確認すること。残差にばらつきがある場合、しばしば  $X$  と  $Y$  をべき乗の関係で見直すことが解決になる。

【結果の解釈】

$\beta_{10} > 0$  のとき、

- ①  $\beta_{11} > 0$  の場合： $X$  が  $Y$  に与える影響は、グループ A の方がグループ B よりも大きい。
- ②  $\beta_{11} \cong 0$  の場合： $X$  が  $Y$  に与える影響は、グループ A とグループ B で差はない。
- ③  $\beta_{11} < 0$  の場合： $X$  が  $Y$  に与える影響は、グループ A の方がグループ B よりも小さい。

パターン 5

パターンの目的： $X$  と  $Y$  の間にべき乗の関係があてはまるとき、すなわちパターン 3 を適用した分析で  $Y$  は  $X$  により十分に説明できるという結果が得られたとき、グループ A とグループ B での、 $X$  が  $Y$  に与える影響の関係の違いを確認する。また、異なる場合に、どのように異なるのかもあわせて確認する。

パターン利用時に具体化する情報： $X, Y, \text{グループ A}, \text{グループ B}$

【仮説】 $X$  が  $Y$  に与える影響は、グループ A の方がグループ B よりも大きい (小さい)。

【モデル式】 $Y = \ell^{(\beta_{00} + \beta_{01}D)} \cdot X^{(\beta_{10} + \beta_{11}D)}$   
 $D$  : ダミー変数  
 (グループ A のとき 1, グループ B のとき 0)

【検証方法】

適用手法 重回帰分析

着目する指標

- 係数  $\beta_{11}$ : 0 と有意に異なるかどうかを確認する。
- 有効な可視化ツール 散布図

【結果の解釈】

- ①  $\beta_{11} > 0$  の場合： $X$  が  $Y$  に与える影響は、グループ A の方がグループ B よりも大きい。
- ②  $\beta_{11} \cong 0$  の場合： $X$  が  $Y$  に与える影響は、グループ A とグループ B で差はない。
- ③  $\beta_{11} < 0$  の場合： $X$  が  $Y$  に与える影響は、グループ A の方がグループ B よりも小さい。

パターン 6

パターンの目的：複数の変数 ( $X_1, \dots, X_n$ ) の比例関係で 1 つの変数  $Y$  をどれくらい説明できるかを確認する。

パターン利用時に具体化する情報： $X_1, \dots, X_n, Y$

【仮説】 $Y$  は  $X_1, \dots, X_n$  に比例する。

【モデル式】 $Y = \beta_0 + \sum \beta_i X_i, i = 1, \dots, n$

【検証方法】

適用手法 重回帰分析

着目する指標

- 決定係数 ( $R^2$ )
- $X_i$  の係数 ( $\beta_i$ ):  $\beta_i$  の推定値を比較することで、 $X_i$  が  $Y$  に与える影響の差を比較することができる。

有効な可視化ツール 3次元の散布図 (説明変数が 2 つの場合のみ)

注意すべき点 重回帰分析の場合、説明変数  $X_i$  の間の相関が大きいと、 $\beta_i$  の推定が不安定になる問題が

発生する．事前に説明変数間の相関があまり大きくないことを確認する必要がある．

【結果の解釈】

(ア) モデルの説明力に着目する場合

- ①  $K < R^2 \leq 1$  の場合： $Y$  は  $X_1, \dots, X_n$  により十分に説明できる．
- ②  $0 \leq R^2 \leq K$  の場合： $Y$  は  $X_1, \dots, X_n$  により十分に説明できない．

上記①のときは，決定係数  $R^2$  の値が有意であること，すなわち該当の回帰モデル式が有意であることを確認すること．有意でない場合は，標本の追加を検討すること．

なお， $K$  は結果の解釈を分ける閾値である． $K$  の具体値は分析対象の特性や，有識者の意見等を総合的に考慮しながら決定する必要がある．

(イ)  $\beta_i$  と  $\beta_j$  の比較に着目する場合（ただし， $\beta_i, \beta_j > 0, j = 1, \dots, n, j \neq i$ ）

- ①  $\beta_i > \beta_j$  の場合： $X_i$  は  $X_j$  よりも， $Y$  に大きな影響を及ぼす．
- ②  $\beta_i \cong \beta_j$  の場合： $X_i$  と  $X_j$  が  $Y$  に及ぼす影響はあまり差がない．

パターン 7

パターンの目的：複数の変数 ( $X_1, \dots, X_n$ ) のべき乗の関係で 1 つの変数  $Y$  をどれくらい説明できるかを確認する．

パターン利用時に具体化する情報： $X_1, \dots, X_n, Y$

【仮説】 $Y$  は  $X_1, \dots, X_n$  のべき乗に比例する．

【モデル式】 $Y = \ell^{\beta_0} \cdot \prod X_i^{\beta_i}, i = 1, \dots, n$

【検証方法】

適用手法 重回帰分析

着目する指標

- 決定係数 ( $R^2$ )
- $X_i$  のべき乗の係数 ( $\beta_i$ ):  $\beta_i$  の推定値を比較することで， $X_i$  が  $Y$  に与える影響の差を比較することができる．

有効な可視化ツール 3次元の散布図（説明変数が2つの場合のみ）

注意すべき点 重回帰分析の場合，説明変数（対数変換後の  $X_i$ ）の間の相関が大きいと， $\beta_i$  の推定が不安定になる問題が発生する．事前に説明変数間の相関があまり大きくないことを確認する必要がある．

【結果の解釈】

(ア) モデルの説明力に着目する場合

- ①  $K < R^2 \leq 1$  の場合： $Y$  は  $X_1, \dots, X_n$  のべき乗により十分に説明できる．
- ②  $0 \leq R^2 \leq K$  の場合： $Y$  は  $X_1, \dots, X_n$  のべ

き乗により十分に説明できない．

①のときは，決定係数  $R^2$  の値が有意であること，すなわち該当の回帰モデル式が有意であることを確認すること．有意でない場合は，標本の追加を検討すること．

なお， $K$  は結果の解釈を分ける閾値である． $K$  の具体値は分析対象の特性や，有識者の意見等を総合的に考慮しながら決定する必要がある．

(イ)  $X_i$  と  $Y$  の関係に着目する場合

- ①  $\beta_i \cong 0$  の場合： $X_i$  の値にかかわらず  $Y$  は不変となり， $X$  と  $Y$  の値の間には特に関係がない．
- ②  $0 < \beta_i < 1$  の場合： $X$  が大きくなるほど， $X_i$  の増加による  $Y$  の増加の割合は小さくなる．
- ③  $\beta_i = 1$  の場合： $X_i$  の増加による  $Y$  の増加の割合は一定である．
- ④  $\beta_i > 1$  の場合： $X_i$  が大きくなるほど， $X_i$  の増加による  $Y$  の増加の割合は大きくなる．

(ウ)  $\beta_i$  の比較に着目する場合（ただし， $\beta_i, \beta_j > 0, j = 1, \dots, n, j \neq i$ ）

- ①  $\beta_i > \beta_j$  の場合： $X_i$  は  $X_j$  よりも， $Y$  に大きな影響を及ぼす．
- ②  $\beta_i \cong \beta_j$  の場合： $X_i$  と  $X_j$  が  $Y$  に及ぼす影響は，あまり差がない．

4.2 属性が与える影響の確認に関するデータ分析パターン

パターン 8

パターンの目的：ある値  $Q$  について，グループ A とグループ B の値の違いを確認する．

パターン利用時に具体化する情報：値  $Q$ ，グループ A，グループ B

【仮説】値  $Q$  について，グループ A とグループ B を比較すると，グループ A の方が大きい．

【モデル式】 $Q_A > Q_B$

$Q_A$ ：値  $Q$  について，グループ A に属するプロジェクトの集合値

$Q_B$ ：値  $Q$  について，グループ B に属するプロジェクトの集合値

【検証方法】

適用手法

- (ア)  $Q_A, Q_B$  が共に正規分布の場合： $t$  検定
- (イ) 上記以外の場合：マン・ホイットニーの順位和検定等（ノンパラメトリック検定）

着目する指標

- (ア)  $t$  検定の場合：仮説  $Q_A = Q_B$  の有意確率，および  $Q_A, Q_B$  の平均と分散
- (イ) ノンパラメトリック検定の場合：仮説  $Q_A =$

表 2 2×2 分割表  
Table 2 Two-by-two contingency table.

	事象 P 発生	事象 P 発生せず	計	危険度 (オッズ比)
グループ A	a	b	a + b	ad / bc
グループ B	c	d	c + d	1
計	a + c	b + d	a + b + c + d	

$Q_B$  の有意確率

有効な可視化ツール 箱ひげ図

【結果の解釈】

- ①  $Q_A > Q_B$  の場合：グループ A の値の方が、グループ B の値よりも大きい。
- ②  $Q_A \cong Q_B$  の場合：グループ A の値とグループ B の値に有意な差はない。
- ③  $Q_A < Q_B$  の場合：グループ A の値の方が、グループ B の値よりも小さい。

**パターン 9**

パターンの目的：ある値  $Q$  について、グループ A とグループ B のばらつきの違いを確認する。

パターン利用時に具体化する情報：値  $Q$ ，グループ A，グループ B

【仮説】ある値  $Q$  について、グループ A とグループ B を比較すると、グループ A の方がばらつきが大きい。

【モデル式】  $V_A > V_B$

$V_A$ ：グループ A に属するプロジェクトの値  $Q$  のばらつき

$V_B$ ：グループ B に属するプロジェクトの値  $Q$  のばらつき

【検証方法】

適用手法

- (ア) 値  $Q$  が、グループ A，グループ B 共に正規分布の場合：F 検定
- (イ) 上記以外の場合：アンサリ・ブラッドレイ検定等 (ノンパラメトリック検定)

着目する指標

- (ア) F 検定を用いる場合：値  $Q$  について、グループ A，グループ B それぞれの分散の値および両者の分散の比。F 検定の対象となる F 値は、グループ A，グループ B の分散の値をそれぞれ  $V_A, V_B$  とすると、 $V_A/V_B$  により算出する。
- (イ) ノンパラメトリック検定の場合：仮説  $V_A = V_B$  の有意確率

有効な可視化ツール 箱ひげ図

【結果の解釈】

- ①  $V_A > V_B$  の場合：値  $Q$  のばらつきは、グループ A の方が、グループ B よりも有意に大きい。
- ②  $V_A \cong V_B$  の場合：値  $Q$  のばらつきは、グループ A とグループ B で有意な差はない。
- ③  $V_A < V_B$  の場合：値  $Q$  のばらつきは、グループ B の方が、グループ A よりも有意に大きい。

グループ A とグループ B で有意な差はない。

- ③  $V_A < V_B$  の場合：値  $Q$  のばらつきは、グループ A の方が、グループ B よりも有意に小さい。

**パターン 10**

パターンの目的：ある事象 P が発生する危険度について、グループ A とグループ B の値の違いを確認する。  
パターン利用時に具体化する情報：事象 P，グループ A，グループ B

【仮説】事象 P が発生する危険度は、グループ A の方がグループ B よりも大きい。

【モデル式】  $R_A/R_B > 1$

$R_A$ ：グループ A に属するプロジェクトで、ある事象 P が発生する危険度

$R_B$ ：グループ B に属するプロジェクトで、ある事象 P が発生する危険度

【検証方法】

適用手法 カイ二乗検定

着目する指標

相対危険度 (オッズ比)  $ad/bc$  の値およびその有意確率：表 2 の 2×2 分割表を参照。a, b, c, d にはそれぞれのセルに属するプロジェクトの件数が入る。たとえば、表 2 においては、事象 P が発生し、かつグループ A に属するものが a 件ということになる。  
有効な可視化ツール 2×2 分割表 (表 2 参照)  
注意すべき点 a, b, c, d のいずれかの件数が少ない (目安として 5 件以下) の場合、Fisher の直接確率計算の方法を用いる。

【結果の解釈】

- ①  $R_A/R_B > 1$  の場合：グループ A の方が、グループ B よりも事象 P が発生する危険度が高い。
- ②  $R_A/R_B \cong 1$  の場合：グループ A とグループ B の間で、事象 P が発生する危険度に有意な差はない。
- ③  $R_A/R_B < 1$  の場合：グループ A の方が、グループ B よりも事象 P が発生する危険度が低い。

**5. 分析パターンのソフトウェア開発に関する仮説への適用事例**

本章では、4 章で述べた分析パターンを実際にソフトウェア開発に関する仮説検証にあてはめて利用した

表 3 検証仮説の例  
Table 3 Example hypotheses.

仮説検証の目的	仮説 ID	仮説検証内容	利用するパターン
開発コストと規模の関係を確認したい	仮説 1-1	製造規模が大きくなると、開発コストは大きくなる傾向にある	パターン 1
	仮説 1-2	開発コストは製造規模に比例する	パターン 2
	仮説 1-3	開発コストは製造規模のべき乗に比例する	パターン 3
	仮説 1-4	製造規模が開発コストに与える影響は、開発形態が機能追加のときの方が新規のときよりも大きい	パターン 5
	仮説 1-5	開発コストは製造規模と再利用規模のべき乗に比例する	パターン 7
総合運転テスト工程の品質に影響を与える要因を知りたい	仮説 2-1	総合運転テスト工程の品質に問題があったプロジェクト群は、問題がなかったプロジェクト群よりも工程 A のエラー・バグ密度が大きい	パターン 8
	仮説 2-2	総合運転テスト工程の品質に問題が発生する危険度は、ある工程で作業 T を実施しなかった場合の方が、作業 T を実施した場合よりも大きい	パターン 10

際の仮説の一覧と仮説検証の進め方について述べる。

表 3 に、検証仮説の一覧を示す。

### 5.1 ソフトウェア開発データの定義

ソフトウェア開発データを以下で定義する。なお、工程の定義は、IPA/SEC の「ソフトウェア開発データ白書 2005」<sup>1)</sup>を参考にした。

#### (1) コスト

- 開発コストは、直営労務費と外注委託費の合計である。

#### (2) 工数

- 全工数とは、要求定義工数、設計工数、製作工数、テスト工数の総和である。
- 要求定義工数とは、要求定義工程に要した工数の総和である。
- 設計工数とは、基本設計工程、詳細設計工程に要した工数の総和である。
- 製作工数とは、製作工程に要した工数の総和である。
- テスト工数とは、結合テスト工程、総合テスト（ユーザ確認、ベンダ確認）工程に要した工数の総和である。

#### (3) プロジェクト属性

プロジェクト属性の例を以下に示す。

- MF/CSS とは、①MF（メインフレーム）、②CSS（クライアントサーバシステム）のいずれかである。
- 開発形態は、①新規、②機能追加のいずれかである。

#### (4) 規模

- 製造規模は、新規製造部分と改造部分の総和である。
- 再利用規模は、手を加えずにそのまま流用する部

分を指す。

#### (5) 品質

- 基本設計エラー密度とは、基本設計工程で抽出したエラー数/基本設計書ページ数である。
- 詳細設計エラー密度とは、詳細設計工程で抽出したエラー数/詳細設計書ページ数である。
- 単体テスト工程バグ密度とは、単体テスト工程で検出したバグ数/製造規模である。
- 結合テスト工程バグ密度とは、結合テスト工程で検出したバグ数/製造規模である。
- 総合テスト工程バグ密度とは、総合テスト（ベンダ確認）工程で検出したバグ数/製造規模である。
- 総合運転テスト工程バグ密度とは、総合テスト（ユーザ確認）工程で検出したバグ数/製造規模である。

#### 5.2 回帰分析の決定係数の閾値 K の決定

パターン 2, 3, 6, 7 で、回帰分析の決定係数に基づく結果の解釈を分ける閾値  $K$  の値は、ある開発現場において試験的に実施したいいくつかの回帰分析の結果と、データ分析に関する有識者との意見をもとに、0.6 と決定した。すなわち、該当のモデルの説明力が 60%以上であるときに、十分な説明力を持つと判定することとした。なお、この値については、今後ソフトウェア開発データの分析結果が蓄積されていく過程で、適宜見直しを行っていく方向とした。

#### 5.3 開発コストと規模の関係に関する仮説検証

開発コストと規模の関係を明確にすることで、単位規模あたりの所要コストの推測や、プロジェクト属性が製造規模と開発コストの関係に与える影響の違いの確認等を実施することとした。

なお、この仮説検証は、実プロジェクトのデータ約

150 件を用いて実施した。

Step1: 製造規模と開発コストの相関関係に関する仮説検証

【仮説 1-1】製造規模が大きくなると、開発コストは大きくなる傾向にある。

利用したデータ分析パターン パターン 1  
データ分析パターン利用時に具体化した情報

$X$  は、製造規模とした。

$Y$  は、開発コストとした。

検証手法 検証手法は、ピアソンの相関係数を用いた。検証結果として、以下を確認した。

$\rho$  が 0 より有意に大きい: 製造規模が大きくなると、開発コストも大きくなる傾向にある。

Step2: 製造規模と開発コストの比例関係に関する仮説検証

【仮説 1-2】開発コストは製造規模に比例する。

利用したデータ分析パターン パターン 2  
データ分析パターン利用時に具体化した情報

$X$  は、製造規模とした。

$Y$  は、開発コストとした。

その他 散布図により製造規模と開発コストの関係を確認し、規模による層別を行った後の回帰分析をあわせて実施した。

仮説検証結果として以下を確認した。

$0 \leq R^2 \leq K$ : 開発コストは製造規模ではあまり説明できない。観測対象により、単位規模あたりの開発コストの値にばらつきがある。

すなわち、製造規模と開発コストの関係は比例関係で説明することは適切でない。

Step3: 製造規模と開発コストのべき乗の関係に関する仮説検証

【仮説 1-3】開発コストは製造規模のべき乗に比例する。利用したデータ分析パターン パターン 3

データ分析パターン利用時に具体化した情報

$X$  は、製造規模とした。

$Y$  は、開発コストとした。

その他 散布図により製造規模と開発コストの関係を確認し、両対数変換を行った方が、より直線の関係に近いことを確認した。

仮説検証結果として以下を確認した。

(ア) モデルの説明力に関する評価

$R^2$  が 0 より有意に大きく、かつ  $K < R^2 \leq 1$ :

開発コストと製造規模の関係は、べき乗の関係で説明できる。

Step4: プロジェクト属性が機能追加のときと新規のときで、製造規模が開発コストに与える影響の違いに

関する仮説検証

【仮説 1-4】製造規模が開発コストに与える影響は、開発形態が機能追加のときの方が、新規のときよりも大きい。

利用したデータ分析パターン パターン 5  
データ分析パターン利用時に具体化した情報

$X$  は、製造規模とした。

$Y$  は、開発コストとした。

グループ A は、開発形態が機能追加のプロジェクトとした。

グループ B は、開発形態が新規のプロジェクトとした。

仮説検証結果として以下を得た。

$\beta_{11} > 0$ : 製造規模が開発コストに与える影響は、開発形態が機能追加のときの方が、新規のときよりも大きい。

Step5: 製造規模、再利用規模と開発コストのべき乗の関係に関する仮説検証

【仮説 1-5】開発コストは製造規模と再利用規模のべき乗に比例する。

利用したデータ分析パターン パターン 7  
データ分析パターン利用時に具体化した情報

$X_1$  は、製造規模とした。

$X_2$  は、再利用規模とした。

$Y$  は、開発コストとした。

その他 3次元の散布図により、製造規模、再利用規模、開発コストのそれぞれを対数変換した後の値の関係を確認し、平面の係数に近いことを確認した。

また、対数変換後の製造規模と再利用規模の間には、ほとんど相関関係はないことを確認した。

仮説検証結果として以下を確認した。

(ア) モデルの説明力に関する評価

$R^2$  が 0 より有意に大きく、かつ  $K < R^2 \leq 1$ :

開発コストは、製造規模と再利用規模のべき乗の関係で説明できる。

(イ)  $\beta_1$  と  $\beta_2$  の比較に関する評価

$\beta_1 > \beta_2$ : 製造規模は再利用規模よりも、開発コストに大きな影響を及ぼす。

5.4 総合運転テスト工程の品質に影響を与える要因に関する仮説検証

総合運転テスト工程の品質は、開発中の最終品質である。このため、この品質指標に影響を与える要因を、総合運転テスト工程の品質に問題があったプロジェクトとなかったプロジェクトの違いを比較することで明らかにしようとした。

なお、この検証は、実プロジェクトのデータ約 300

件を用いて実施した。

Step1: 総合運転テスト工程よりも前の工程のエラー・バグ密度との関係に関する仮説検証

【仮説 2-1】総合運転テスト工程の品質に問題があったプロジェクトは、問題がなかったプロジェクトよりも、工程 A のエラー・バグ密度が大きい。

利用したデータ分析パターン パターン 8

データ分析パターン利用時に具体化した情報

値  $Q$  は、工程 A のエラー・バグ密度とした。

ただし、工程 A は基本設計工程、詳細設計工程、単体テスト工程、結合テスト工程、総合テスト工程（ベンダ確認）のいずれかである。

グループ A は、総合運転テスト工程の品質に問題があったプロジェクトとした。

グループ B は、総合運転テスト工程の品質に問題がなかったプロジェクトとした。

検証手法 検証手法は、ノンパラメトリック検定を用いた。

仮説検証結果として、工程により、以下のいずれの場合も存在することを確認した。

・ $Q_A > Q_B$ : 総合運転テスト工程の品質に問題があったプロジェクトの方が、総合運転テスト工程の品質に問題がなかったプロジェクトよりも、工程 A のエラー・バグ密度が大きい。

・ $Q_A \cong Q_B$ : 総合運転テスト工程の品質に問題があったプロジェクトと、総合運転テスト工程の品質に問題がなかったプロジェクトとの間で、工程 A のエラー・バグ密度に有意な差がない。

・ $Q_A < Q_B$ : 総合運転テスト工程の品質に問題があったプロジェクトの方が、総合運転テスト工程の品質に問題がなかったプロジェクトよりも、工程 A のエラー・バグ密度が小さい。

Step2: ある工程での作業 T の実施有無との関係に関する仮説検証

【仮説 2-2】総合運転テスト工程の品質に問題が発生する危険度は、ある工程で作業 T を実施しなかった場合の方が、作業 T を実施した場合よりも大きい。

利用したデータ分析パターン パターン 10

データ分析パターン利用時に具体化した情報

事象 P は、総合運転テスト工程の品質に問題が発生することとした。

グループ A は、ある工程で作業 T を実施しなかったプロジェクトとした。

グループ B は、ある工程で作業 T を実施したプロジェクトとした。

検証手法 検証手法は、 $2 \times 2$  分割表のどのセルのプロ

ジェクト数も 5 件以上であったため、カイ二乗検定を用いた。

仮説検証結果として、以下の結果が得られ、ある工程で作業 T を実施しないことが、総合運転テスト工程の品質に問題が発生する要因であることを明らかにした。

$R_A/R_B > 1$ : ある工程で作業 T を実施しなかったプロジェクトの方が、ある工程で作業 T を実施したプロジェクトよりも、総合運転テスト工程の品質に問題が発生する危険度が高い。

## 6. 評価

本章では、今回検討したデータ分析パターンを用いたソフトウェア開発データ分析による効果について評価した結果を述べる。

### 6.1 仮説検証作業の効率化の評価

今回提案したデータ分析パターンにより、検証活動がどれくらい効率化されるかを評価実験により確認する。

#### 6.1.1 検証活動の進め方と作業項目

表 4 に、評価対象とした検証活動の作業項目の内訳を示す。検証活動は「仮説設定」「データ分析」「資料作成」「レビュー」の 4 つの作業項目から構成され、基本的にこの順序で検証活動が進められる。

なお、各分析活動において分析対象となる開発データは、これらの検証活動に入る前に、1 枚の Excel シート上に基本的にすべて整理されている。すなわち、データ収集・投入等の作業時間は、評価対象とした検証活動には含まれない。

#### 6.1.2 評価対象とする検証活動

表 5 に、評価対象とした検証活動の概要を示す。表の各項目について以下で説明する。

#### (1) 分析者

分析者は 3 人 (A, B, C) である。表 6 に各分析者の評価実験開始時点におけるデータ分析手法の経験の有無を示す。また表 7 に、各分析者の、評価実験開始時点における商用ソフトウェア開発の経験年数および関連保有資格を示す。

#### (2) 分析対象開発現場

分析対象開発現場は 2 拠点 (P, Q) である。

#### (3) 分析者から見た検証仮説群の新規性

該当の検証活動において、学習効果による作業効率の向上が発生している可能性があるかどうかを見るための目安である。検証仮説群が分析者から見て「一部経験あり」「経験あり」である検証活動 2-2、検証活動 2-3 は、分析者がパターン確立前に、類似の仮説検証

表 4 検証活動の作業項目

Table 4 Tasks for data analysis activities.

	作業項目	作業の概要
1	仮説設定	・ 経営者、管理者等との仮説検証の方向性についての意識あわせ ・ 方向性に対応した検証仮説の検討
2	データ分析	・ 統計解析ツール等を用いたデータ分析作業
3	資料作成	・ 経営者、管理者等を対象とした報告資料の作成
4	レビュー	・ 報告資料のレビュー・修正

表 5 評価対象とした検証活動の概要

Table 5 Profiles of evaluated data analysis activities.

パターン適用分類	検証活動 ID	分析者	対象開発現場	分析者から見た検証仮説の新規性	業務種別	検証仮説数
パターン確立前	検証活動 1-1	A	P	新規	実業務	9
	検証活動 1-2	B	P	新規	実業務	14
パターン確立後	検証活動 2-1	A	Q	新規	実業務	16
	検証活動 2-2	B	Q	一部経験あり	実業務	24
	検証活動 2-3	B	P	経験あり	訓練	14
	検証活動 2-4	C	P	新規	訓練	4

表 6 分析者の評価実験開始時点におけるデータ分析手法の経験有無

Table 6 Experience about statistical data analysis in subjects at start of experiments.

分析方法種別	分析手法	分析者		
		A	B	C
パラメトリック	相関分析 (ピアソンの相関係数)	○	○	×
	単回帰分析	○	○	×
	重回帰分析	○	×	×
	べき乗の回帰分析	○	×	×
	ダミー変数を用いた回帰分析	○	×	×
	2 群の差の検定 (t 検定)	○	×	×
	2 群のばらつき差の検定 (F 検定)	○	×	×
ノンパラメトリック	相関分析 (スピアマンの順位相関係数)	×	×	×
	2 群の差の検定 (マンホイットニー-U 検定など)	×	×	×
	2 群のばらつき差の検定 (アンサリ・ブラッドレイ検定など)	×	×	×
	独立性の検定	×	×	×

(凡例 ○ : 経験あり, × : 経験なし)

表 7 分析者の評価実験開始時点におけるソフトウェア開発の経験・スキルの概要

Table 7 Experience of software development in subjects at start of experiments.

	分析者		
	A	B	C
商用ソフトウェア開発プロジェクト経験年数	4	0	0
関連保有資格	テクニカルエンジニア (ネットワーク)	ソフトウェア開発技術者	基本情報処理技術者

を実施した経験があるものである。これらについては、分析者自身の学習効果による作業効率の向上の影響が無視できないものと考えられる。

(4) 業務種別

以下の 2 つに分類される。

① 業務種別が「実業務」の場合

仮説検証の方向性の確認および仮説検証結果の報告を、該当の開発現場を担当する経営者、管理者等を対

表 8 各検証活動の作業別所要日数と 1 日あたりの仮説検証数

Table 8 Days required for each work per one task and number of hypothesis verified per day.

検証活動 ID	各作業の所要日数 (単位: 日)					仮説検証の生産性 (1 日あたりの検証仮説数) (単位: 個/日)				
	合計	仮説設定	データ分析	資料作成	レビュー	全作業	仮説設定	データ分析	資料作成	レビュー
検証活動 1-1	42	8	18	12	4	0.21	1.13	0.50	0.75	2.25
検証活動 1-2	54	10	20	18	6	0.26	1.40	0.70	0.78	2.33
検証活動 2-1	32	11	10	7	4	0.50	1.45	1.60	2.29	4.00
検証活動 2-2	40	7	21	9	3	0.60	3.43	1.14	2.67	8.00
検証活動 2-3	13	0.5	8	4	0.5	-	-	1.75	-	-
検証活動 2-4	8	0.5	4	3	0.5	-	-	1.00	-	-

(注)検証活動 2-3, 検証活動 2-4 は業務種別が訓練であるため, 仮説検証の生産性の算出はデータ分析作業についてのみ行っている。

表 9 分析パターン確立前と確立後の生産性の変化の比較

Table 9 Productivity before and after establishment of data analysis patterns.

項番	比較の組み合わせ		パターン確立後に生産性が何倍になったか?				
	パターン確立前	パターン確立後	全作業	仮説設定	データ分析	資料作成	レビュー
比較 1	検証活動 1-1	検証活動 2-1	2.38	1.28	3.20	3.05	1.78
比較 2	検証活動 1-2	検証活動 2-2	2.31	2.45	1.63	3.42	3.43
比較 3	検証活動 1-2	検証活動 2-3	-	-	2.50	-	-
比較 4	検証活動 1-2	検証活動 2-4	-	-	1.43	-	-

象に実施する。これは実際の業務活動である。

## ②業務種別が「訓練」の場合

仮説検証の方向性の確認および仮説検証結果の報告を, 分析関係者内部で実施する。これは主に分析者のスキル向上を目的とした訓練である。

業務種別が「訓練」である, 検証活動 2-3, 検証活動 2-4 は「実業務」の検証活動と単純に比較することはできない。業務種別が「訓練」である検証活動の生産性の算出は「訓練」であることによる影響が比較的小さいと考えられるデータ分析作業のみを対象とする。

### 6.1.3 評価方法

#### (1) 各検証活動の生産性の算出

検証活動ごとの所要時間を「仮説設定」「データ分析」「資料作成」「レビュー」の作業項目ごとに確認する。そして, 1 日あたりに処理できた検証仮説数, すなわち仮説検証の生産性を, 検証活動全体および作業項目ごとに確認する。

#### (2) パターン確立前と確立後の生産性の比較

(1) で算出した各検証活動の生産性をもとに, 分析パターンの確立により生産性がどのように変化したかを確認する。変化の度合いは, パターン確立前の生産性を 1 としたときの, 分析パターン確立後の生産性の相対値で確認する。

### 6.1.4 評価結果

表 8 に, 各検証活動の全作業および作業項目別の生産性の算出結果を, 検証活動の所要日数とともに示す。また, 表 9 に, パターン確立前の生産性を 1 としたときの, パターン確立後の生産性の相対値を示す。この比較は, パターン確立前後で, 分析者が同一, あるいは関連スキルがほぼ同等の組合せについて実施している。いずれの比較においても, 分析パターン確立後の生産性が向上していることが確認できる。

### 6.2 その他の効果

表現方法を【仮説】【モデル式】【検証方法】【結果の解釈】の 4 項目に整理して記述することで, 報告の受け手に応じた報告資料の作成や, 説明の重点化が容易になるという効果が確認できた。たとえば, 経営者, 管理者等には, 限られた時間で状況を把握してもらう必要があるため, 主に【仮説】および【結果の解釈】に絞って説明を行う。これに対し, 分析者間での意見交換等の場合は【モデル式】や【検証方法】も重要である。表現方法を上記のように整理しておくことで, 報告対象や状況に応じた説明項目の使い分け等が容易になる。

## 7. 考 察

### 7.1 作業項目別の生産性の変化

表9の比較1と比較2で、パターン確立前後の各作業項目の生産性を比較した値を確認すると、すべての作業項目において、パターン確立後に生産性の向上が見られる。特に資料作成作業の生産性の向上が大きい。分析者A、分析者Bによると、資料作成作業およびレビュー作業で生産性が向上した要因の1つとして、資料の記述項目も、データ分析パターンにあわせて【仮説】、【モデル式】、【検証方法】、【結果の解釈】にパターン化できたことが考えられるとのことであった。なお、比較1と比較2では、仮説設定作業の生産性向上に大きな差が見られるが、これは開発現場Qでは、分析者Aが経営者、管理者等の問題意識に基づくトップダウン的な仮説検証を実施したのに対し、分析者Bは分析パターンと開発データをもとにボトムアップ的な仮説検証を実施したことによるところが大きいものと考えられる。

### 7.2 学習効果の影響

データ分析パターンの適用による効果は、できるだけ分析者自身の学習効果を排除した条件で評価を行う必要がある。分析者Cのみ、パターン確立後に初めて検証活動を実施しているが、表6、表7により、分析者Bと分析者Cは、評価実験開始前の関連スキルが比較的近いことが確認できる。このため、パターン確立後に初めて検証活動を実施した分析者Cと、パターン確立前の分析者Bと比較した表9の比較4の結果を確認する。これによると、データ分析パターン確立後のデータ分析作業の生産性は1.4倍となっている。この生産性の向上要因からは学習効果の影響は排除されており、データ分析パターンの確立の効果を、より正確に反映していると考えられる。

なお、表9の比較3は、パターン確立前の分析を経験した分析者が、パターン確立後にまったく同じ仮説検証を、同じ開発現場で実施した際の生産性の変化を見たものである。パターン確立後のデータ分析作業の生産性は2.5倍となっており、比較4よりも生産性の向上が大きい。

### 7.3 分析者のスキルとパターンの効果の関係

分析者のスキルとパターンの効果の関係を確認するため、分析者によるパターン確立前後での生産性の変化の違いを確認する。分析者Aと分析者Bのパターン確立前後の分析活動の生産性の変化、すなわち、表9の比較1と比較2の結果を確認すると、分析者Aの方が、データ分析作業の生産性の向上が大きいという

結果が得られている。すなわち、データ分析パターンの効果は、分析者のスキルが高い方が大きいと考えられる。分析者Aによると、データ分析作業が効率化できた要因として、対応する仮説や結果の解釈とデータ分析作業の関係を明確に意識することで必要なデータ分析作業を絞り込めたことがあるのではないかと意見が得られた。なお、それ以外の作業の生産性の上昇は比較2、すなわち分析者Bの方が大きい、これは検証活動2-2では一部経験のある仮説検証を行っていることが大きいと考えられる。

### 7.4 パターンによる分析品質の向上

データ分析パターンの効果は効率面だけではなく、正しい式や検証方法を用いることによる分析の質の向上もあると考えられる。たとえばパターン確立前の分析では、データの分布が正規分布と大きく異なり、本来はノンパラメトリックな手法が用いられるべき場合にも、一律に正規分布を前提としたパラメトリックな手法が用いられてしまっている事例がいくつか見られた。パターン確立後は、対象データの確認により、パラメトリックな手法とノンパラメトリックな手法を適切に使い分けるといった改善が見られた。ただし、品質向上についての定量的評価は今後の課題である。

### 7.5 パターンの十分性の検討

今回提案した10個のデータ分析パターンにより、どれくらい実際の開発現場での仮説検証を実施できるかが、データ分析パターンの有効性を評価するうえで重要である。ここでは、1拠点ではあるが、開発現場Qでの仮説検証の状況を、主にデータ分析パターンの適用可否の観点から確認する。

開発現場Qでは、経営者、管理者等の問題意識を確認してから設定したトップダウン的な仮説と、分析者が開発データとデータ分析パターンを確認しながら設定したボトムアップ的な仮説の双方の検証を実施している。このうち、データ分析パターンの十分性については、トップダウン的な仮説へのデータ分析パターンの適用状況により議論する必要がある。仮説検証の実施可否は、データ分析パターンの有無だけでなく、分析対象となる開発データの有無にも依存する。表10に、開発現場Qの経営者、管理者等の問題意識に基づく仮説で、データ分析パターンの適用、開発データの入手がとも可能であったため仮説検証が実施できた仮説の概要と、データ分析パターンの適用状況を示す。また、表11に、経営者、管理者等の問題意識に基づく仮説のうち、適切な開発データが入手できずに検証を断念した仮説群の概要を、新たに必要と考えられた開発データと、適用を想定したデータ分析パター

表 10 開発現場 Q でのデータ分析パターンの適用状況（経営者，管理者等の問題意識に基づく仮説で，検証が実施できたもの）

Table 10 How data analysis patterns applied to hypothesis verification at development organization Q: VERIFIABLE hypotheses based on executives' or managers' consciousness.

分類	検証仮説群の概要	適用したデータ分析パターン									
		1	2	3	4	5	6	7	8	9	10
1-1	あるプロジェクト特性の生産性に関するリスクの可視化	×	×	×	×	×	×	○	○	○	○
1-2	プロジェクトの要員数と生産性の関係	×	○	○	×	○	○	○	×	×	×
1-3	生産性と原価率の関係	○	×	×	×	×	×	×	×	×	○
1-4	工数と製造規模，試験項目数の関係	○	○	○	×	×	×	○	×	×	×

(凡例 ○：適用，×：不適用)

表 11 開発現場 Q でのデータ分析パターンの適用状況（経営者，管理者等の問題意識に基づく仮説で，検証に適切な開発データがなかったもの）

Table 11 How data analysis patterns applied to hypothesis verification at development organization Q: UNVERIFIABLE hypotheses due to lack of proper metrics based on executives' or managers' consciousness.

分類	検討した検証仮説群の概要	適用可能と判断したデータ分析パターン										新たに必要な開発データ
		1	2	3	4	5	6	7	8	9	10	
2-1	開発者が同時に所属するプロジェクト数と生産性の関係	○	○	△	△	△	-	-	○	△	○	開発者単位で見た生産量
2-2	全ての工程を同一の開発者が担当する場合の生産性の向上の有無	○	○	△	-	-	-	-	○	-	○	プロジェクトの工程分担に関する情報
2-3	共通業務 A の負荷がプロジェクト生産性に与える影響	○	○	○	-	-	-	-	○	-	-	プロジェクトごとの共通業務 A の工数
2-4	他からの流用が多い開発と新規開発の生産性の比較	○	○	△	△	△	-	-	○	-	-	流用するプログラムの規模
2-5	障害対応等によるプログラムの改修回数と生産性の関係	○	○	△	△	△	-	-	○	-	△	プログラムとプロジェクトの対応付けに関する情報

(凡例 ○：適用可能，△：適用可能性あり，-：適用可能性小)

ンとともに示す。今回，開発現場 Q の経営者，管理者等の問題意識に基づく仮説は，表 10 と表 11 にすべて含まれている。すなわち，今回確認した開発現場 Q の経営者，管理者等の問題意識に基づく仮説は，すべて分析パターンによる検証が可能である。

また，表 12 に，開発現場 Q での検証仮説のうち，分析者が開発データと分析パターンをもとにボトムアップ的に設定，検証し，経営者，管理者等に結果を報告した仮説の概要を示す。こちらは分析者がデータ分析パターンを適用しながら提案した仮説であるため，すべてデータ分析パターンを適用したものとなっている。

図 3 に，経営者，管理者等が検証したい仮説，データ分析パターンにより検証が可能な仮説，開発データ

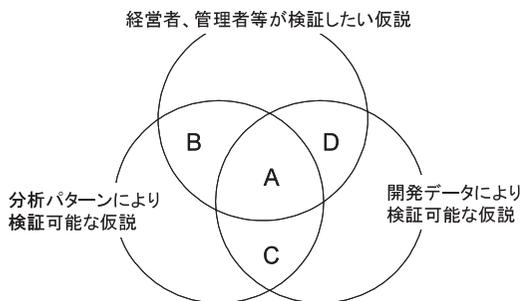
により検証が可能な仮説のイメージを示す。図 3 において，領域 A の仮説のみが，データ分析パターン，開発データの双方が存在し，かつ経営者，管理者等の問題意識に合致した仮説である。図 3 で，検証したい仮説が領域 A 以外に属する場合は，不足しているものを補い，該当の仮説を領域 A に移行できるかどうか検討する必要がある。たとえば領域 B では，不足している開発データの用意や，代用データの適用を検討する必要がある。また，領域 C では，分析者が仮説案や仮説検証結果を経営者，管理者等に提案，報告し，仮説が問題意識に合致しているか，あるいは新たな問題意識を喚起できるかを確認する必要がある。

ソフトウェア開発データ分析に適用可能なデータ分析パターンにより，基本的な分析手法を開発現場等に

表 12 開発現場 Q でのデータ分析パターンの適用状況（分析者の提案に基づく仮説）  
 Table 12 How data analysis patterns applied to hypothesis verification at development organization Q: VERIFIABLE hypotheses proposed by data analysts.

分類	検証仮説群の概要	適用したデータ分析パターン									
		1	2	3	4	5	6	7	8	9	10
3-1	開発コストと工数の関係	○	○	○	○	○	×	×	×	×	×
3-2	開発コストと製造規模の関係	○	○	○	×	○	○	○	×	×	×
3-3	最終工程の品質と上流工程の工数比率の関係	○	×	×	×	×	×	×	○	×	×
3-4	最終工程の品質と上流工程の品質の関係	○	×	×	×	×	×	×	○	×	○
3-5	プロジェクト特性によるある指標の値の差の確認	×	×	×	×	×	×	×	○	×	×

(凡例 ○：適用，×：不適用)



領域	実施する作業
A	経営者、管理者等の問題意識に基づく仮説検証の実施
B	新たな開発データの定義・収集の検討
C	分析者による経営者、管理者等への仮説検証の提案
D	新たな分析パターンの確立の検討

図 3 経営者、管理者等が検証したい仮説と分析パターン，開発データの関係のイメージ  
 Fig. 3 Diagram among hypotheses based on executives' or managers' consciousness, data analysis patterns and software development project data sets.

展開することができれば，図 3 のイメージのように，開発現場において収集する開発データの見直しや，分析者から経営者，管理者等への新たな仮説の提案等が進むことが期待できると考えられる。

8. おわりに

本論文では，ソフトウェア開発データの分析を開発現場に普及，展開するために，データ分析の適用知識を，仮説，モデル式，検証方法，結果の解釈の 4 項目の組み合わせたパターンとして用意することを提案した。

また，ソフトウェア開発データの分析に共通的に使用できると考えられる 10 個のデータ分析パターンの提案を行い，それらのデータ分析パターンを，実際にソフトウェア開発に関する仮説検証にあてはめた例を示した。評価として，データ分析パターンによる仮説検証の生産性向上の効果を示し，さらにデータ分析パターンの十分性について議論を行った。

今後は，データ分析パターンの十分性の検証を進めるとともに，必要なデータ分析パターンの拡充につい

て検討していく予定である。

謝辞 本研究を行うにあたり，貴重なご指導をいただいた株式会社 NTT データ山下徹代表取締役副社長執行役員に感謝いたします。また，様々なご助言をいただいた同技術開発本部ソフトウェア工学推進センタ木谷強センタ長，菅原昌平シニアスペシャリスト（現ニューソン株式会社品質保証部長）に感謝いたします。有益なコメントをいただいた本論文の査読者の方々に感謝いたします。ソフトウェア開発データの分析にご協力いただいた皆様に感謝いたします。

参 考 文 献

- 1) 独立行政法人情報処理推進機構，ソフトウェア・エンジニアリング・センター：ソフトウェア開発データ白書 2005，日経コンピュータ（編），日経 BP 社，東京（2005）。
- 2) 独立行政法人情報処理推進機構，ソフトウェア・エンジニアリング・センター：ソフトウェア開発データ白書 2006，日経コンピュータ（編），日経 BP 社，東京（2006）。
- 3) Tichy, W.F., Lukowicz, P., Prechelt, L. and

- Heinz, E.A.: Experimental Evaluation in Computer Science: A Quantitative Study, *Journal of Systems and Software*, Vol.28, No.1, pp.9-18, Elsevier (1995).
- 4) Zelkowitz, M.V. and Wallace, V.: Experimental Validation in Software Technology, *Information and Software Technology*, Vol.39, No.11, pp.735-744, Elsevier (1997).
  - 5) Zannier, C., Melnik, G. and Maurer, F.: On the Success of Empirical Studies in the International Conference on Software Engineering, *Proc. 28th ICSE 2006*, pp.341-350, ACM (2006).
  - 6) Perry, D., Porter, A. and Votta, L.: Empirical Studies of Software Engineering: A Roadmap, *Proc. Conference on The Future of Software Engineering*, pp.345-355, ACM (2000).
  - 7) Dybå, T., Kitchenham, B.A. and Jørgensen, M.: Evidence-Based Software Engineering for Practitioners, *IEEE Software*, Vol.22, No.1, pp.58-65 (2005).
  - 8) Sjøberg, D., et al.: A Survey of Controlled Experiments in Software Engineering, *IEEE Trans. Softw. Eng.*, Vol.31, No.9, pp.733-753 (2005).
  - 9) Kitchenham, B.A., et al.: Preliminary Guidelines for Empirical Research in Software Engineering, *IEEE Trans. Softw. Eng.*, Vol.28, No.8, pp.721-734 (2002).
  - 10) Liu, Q. and Mintram, R.: Preliminary Data Analysis Methods in Software Estimation, *Software Quality Control*, Vol.13, No.1, pp.91-115, Kluwer Academic Publishers (2005).
  - 11) Unwin, A.: Patterns of Data Analysis?, *Journal of the Korean Statistical Society*, Vol.30, No.2, pp.219-230 (2001).
  - 12) Dybå, T.: An Empirical Investigation of the Key Factors for Success in Software Process Improvement, *IEEE Trans. Softw. Eng.*, Vol.31, No.5, pp.410-424 (2005).
  - 13) Gopal, A., Mukhopadhyay, T. and Krishnan, M.F.: The Impact of Institutional Forces on Software Metrics Programs, *IEEE Trans. Softw. Eng.*, Vol.31, No.8, pp.679-694 (2005).

(平成 18 年 11 月 30 日受付)

(平成 19 年 5 月 9 日採録)



服部 昇 (正会員)

(株)NTT データ技術開発本部ソフトウェア工学推進センタ所属。平成元年慶應義塾大学経済学部卒業。同年(株)NTT データ入社。通信プロトコルの標準化, 実用化, 研究開発マネジメント業務等を経て, 平成 17 年より同所属。実証的ソフトウェア工学の研究に従事。プロジェクトマネジメント学会, ACM 各会員。



副島 千鶴

(株)NTT データ技術開発本部ソフトウェア工学推進センタ所属。平成 15 年名古屋大学大学院工学研究科分子化学工学専攻修了。同年(株)NTT データ入社。グリッドコンピューティングの研究開発を経て平成 17 年より同所属。実証的ソフトウェア工学の研究に従事。



山本修一郎 (正会員)

(株)NTT データフェロー。技術開発本部システム科学研究所長。博士(工学)。昭和 54 年名古屋大学大学院工学研究科情報工学専攻修了。同年日本電信電話公社入社。平成 2 年日本電信電話株式会社ソフトウェア研究所主幹研究員, 平成 11 年同社情報流通プラットフォーム研究所主幹研究員を経て, 平成 14 年(株)NTT データ技術開発本部副本部長となり, 平成 19 年より現職。ソフトウェア工学, コピキタスコンピューティング, 知識創造社会の実現に向けた知識創造デザイン技術の研究に従事。電子情報通信学会, 日本ソフトウェア科学会, 人工知能学会, 日本データベース学会各会員。平成 13 年度情報処理学会業績賞。平成 14 年度電子情報通信学会業績賞。平成 15 年度逓信協会前島賞。