

ソフトウェア開発技術者の単価推定に向けた分析

角田 雅照^{1,2} 松本 健一¹ 大岩 佐和子³ 押野 智樹³

概要: ソフトウェアを購入する企業にとって、ソフトウェアの価格は非常に重要である。受託開発ソフトウェアにおいて、価格の妥当性を判断するために、ソフトウェア開発工数に技術者単価を乗じて価格を推定することが行われる。単価を求めるためには、プロジェクトマネージャなどの各役割の技術者の参画割合を知る必要がある。ただし一般的にソフトウェア購入企業は、技術者の参画割合を知っているとは限らない。そこで本研究では、技術者の参画割合を用いずに、単価の推定を支援することを目的とする。そのために、単価に影響する要因を明らかにするとともに、要因が変化した場合、単価がどの程度変化するかについても示す。分析では企業横断的に収集されたソフトウェア開発プロジェクトデータから、99 件を対象に重回帰分析するとともに、200 から 400 件程度のデータを用いて二変量解析した。分析の結果、開発言語や基本設計工程の契約形態などが単価に影響していることがわかった。

Analysis toward Unit Price Estimation of Software Engineers

MASATERU TSUNODA^{1,2} KENICHI MATSUMOTO¹
SAWAKO OHIWA³ TOMOKI OSHINO³

Keywords: benchmarking, unit cost, software developers

1. はじめに

ソフトウェアを購入する企業にとって、ソフトウェアの価格は非常に重要である。企業が購入するソフトウェアとして、パッケージソフトウェアと受託開発ソフトウェアが存在する。前者はワードプロセッサや表計算ソフトなどの汎用性の高いソフトウェアであり、市場に流通しているソフトウェアを購入して利用する。後者は企業ごとの異なる業務フローに対応した個性性の高いソフトウェアであり、ソフトウェア開発企業（受託者）にユーザ（委託者）が発注して開発が行われる。経済産業省が 2018 年に行った調査によると、ソフトウェア開発企業の売上高ベースで、受注ソフトウェア開発の割合は 81.8%であるのに対し、ソフトウェアプロダクト（パッケージソフトウェア）の割合は 18.2%であり、前者が大部分を占める[4]。

パッケージソフトウェアの場合、市場に流通しているため、同等の機能を持つソフトウェア間の価格を比較し、価格の妥当性を判断することができる。これに対し、受託開発ソフトウェアの場合、受託者と委託者が個別に契約して作成されるため、他社で使われているソフトウェアの価格や機能を知ることができず、同等の機能を持つソフトウェアを比較して価格の妥当性を判断することが困難である。

そのため受託開発ソフトウェアの価格を、ソフトウェア開発工数に技術者単価を乗じることにより推定する場合がある（官公庁の調達では原価計算方式と呼ばれる）。例えば

政府が作成した「デジタルガバメント推進標準ガイドライン」[9]では、価格妥当性を確認できるように、工数、単価などの積算内訳を明確にすることが求められている。また、著者らの研究[12]でも、工数と技術者単価を推定できれば、高精度で価格を推定できることを示している。さらに、情報サービス産業協会（JISA）の価格モデル[2]でも、工数（モデル中では「生産量」×「生産性」で定義）と技術者単価に基づき、価格を決定している。

上記で述べたような、従来の価格推定方法では、プロジェクトマネージャ、プログラマーなどの各役割別の技術者単価を参照し、それぞれの役割の技術者が何%参画しているかにより単価を決定することを前提としている。例えばプロジェクトマネージャとプログラマーがそれぞれ 50%ずつ参画している場合、参画割合を重みとして単価を加重平均して算出する。ただし一般的に委託者は、各役割の技術者参画割合を知ることができるとは限らず、仮に知ることができたとしても、その参画割合が妥当かどうかを検討する必要がある。

そこで本研究では、従来の価格推定方法を利用することを前提としつつ、技術者の参画割合を用いずに、技術者単価の推定を支援することを目的とする。そのために、単価に影響する要因を明らかにするとともに、要因が変化した場合、単価がどの程度変化するかについても示す。

以降、2 章では本研究と関連する研究について述べる。3 章では分析で用いたデータの詳細について説明し、4 章ではデータを分析した結果得られた、単価の影響要因と、要因ごとに見た単価の分布について明らかにする。5 章では妥当性の脅威について検討するとともに、研究結果の活用方法について説明する。6 章ではまとめと今後の展望につ

1 奈良先端科学技術大学院大学
Nara Institute of Science and Technology
2 近畿大学
Kindai University
3 経済調査会
Economic Research Association

いて述べる。

2. 関連研究

これまで受託ソフトウェア開発の価格に関する研究は少なく、特に技術者の単価に影響する要因についてはほとんど分析されていない。著者らによる研究[12]では、受託ソフトウェア開発において、技術者単価に影響する可能性のある要因（価格推定モデルの説明変数の候補）をいくつか示している。ただし、相関係数（二変量解析）でのみ分析しており、重回帰分析を用いていないため、要因間の相互関係は考慮していない。また、単価の基準となる値（中央値や箱ひげ図）を一切示していないため、受託者が文献[12]を用いて単価を推定することはできない。さらに、文献[12]で用いているデータは 31 件であるのに対し、本研究では重回帰分析では 99 件、二変量解析では 200 から 400 件程度の大規模データを用いており、分析結果の外的妥当性がより高いと考えられる。

パッケージソフトウェアの価格を決定するための方法はいくつか提案されている[5][8]。ただし、これらは開発者が最大の利益を得るために用いるモデルであり、受託開発ソフトウェアの価格の妥当性判断に用いることはできない。また、（組込み系の）受託開発ソフトウェアの原価率（売上に対する原価の割合）に影響する要因を分析した研究[13]があるが、原価率と価格は直接の関係がなく、価格の妥当性判断に用いることはできない。委託側がコストを見積もるためには、工数見積もりモデル[1][6][11]が多用される。ただし、利益や技術者単価をどのように設定するかについてはモデルに含まれていないため、それらのモデルをそのまま受託側が用いても価格を推定することはできない。

3. 分析に用いたデータ

分析に用いたデータは、財団法人経済調査会により 2001 年から 2018 年にかけて収集されたものであり、2225 件のソフトウェア開発プロジェクトのデータが含まれている。システムの適用分野は、プロジェクトの 82%が事務系、10%が制御系、8%がその他である。表 1 にデータに含まれる業種の割合を示す。製造業が 25%であり、最も割合が高いが、業種の偏りは比較的小さいといえる。

データの抽出条件: プロジェクトの条件を揃えるため、開発 5 工程（要求分析から総合テストまでの 5 つの工程）が実施されているプロジェクトのみを分析対象とした。また、技術者単価を分析対象とするため、単価の算出に必要なソフトウェア開発見積金額と見積開発工数（時間）が記録されているデータを分析対象とした。その結果、477 件のプロジェクトが抽出された。

分析用変数の定義: 本研究では、原価計算方式（1 章参照）を前提とすると、適正な技術者単価 u は見積もり時点の金額 p と工数 e に基づくと考えられるため、以下のよう

表 1 データに含まれる業種の割合

業種	割合
建設業	3%
製造業	25%
電気・ガス・熱供給・水道業	6%
情報通信業	7%
流通業	20%
金融業・保険業	16%
サービス業	12%
公務	11%

に技術者単価を定義した。すなわち、技術者単価の単位は技術者 1 名の 1 ヶ月あたりの金額（金額の単位は 100 万円）となる。

$$u = p / e \quad (1)$$

分析に用いたデータでは、プロジェクトマネージャ、プログラマーなどの各役割の技術者が、開発の各工程に何%参画しているか（各工程において、何%が各役割の工数か）が記録されている。例えば、設計ではシステムエンジニアが 60%、プログラマーが 40%参画（合計 100%）参画しているなどが記録されている。よって、例えばプログラマーの総参画割合 p_a （開発の全工数において、何%がプログラマーの工数か）は、以下のようにして、これらの割合に対し、各工程の比率を乗じた値を合計することにより求められる。

$$p_a = r_r p_r + r_d p_d + r_m p_m + r_t p_t \quad (2)$$

ここで r_r , r_d , r_m , r_t はそれぞれ要求分析、設計、製造、テストの工程比率を示し、 p_r , p_d , p_m , p_t はそれぞれの工程におけるプログラマーの参画比率を示す。

分析では、要求分析比率などの各工程の比率については中央値を用いた。具体的には、文献[3]に新規案件と改造案件それぞれの各工程の比率の中央値が示されており、それらを各工程の比率として用いた。なお、データに含まれるプロジェクトでは新規案件が多かったため、新規案件かどうか不明なデータ 18 件（全体の 0.8%のデータ）については、新規案件として各工程の比率を用いた。これらにより、各役割の技術者の総参画割合（プロジェクトマネージャ総参画割合、プログラマー総参画割合など）を算出した。

重回帰分析でのデータの取り扱い: 分析では OS 種別と 4 章で説明する変数がすべて記録されている 99 件のデータを用いて重回帰分析をした（リストワイズ除去）。なお 4 章ではリストワイズ除去していない 477 件のデータについても同様の傾向が見られるか（99 件のデータが偏った傾向を持っていないか）を確かめた。

カテゴリ変数の扱い: ソフトウェア開発言語については、類似の言語についてはまとめて一つの区分として扱った（C, C++, VC++, C#など）。また、ソフトウェアの対象業種と開発言語において、プロジェクト件数の少ない区分

表 2 技術者単価との相関係数

	従業員数	IT 技術者数
相関係数	0.22	0.12
p 値	0.00	0.01
ケース数	456	457

表 3 各変数と技術者単価との関連

説明変数	標準化 偏回帰係数	p 値
契約形態/基本設計	-4.00	0.00
PG 合計	3.51	0.00
開発言語(VB)	-2.81	0.01
アナリストの経験と能力	2.49	0.02
発注要件の明確度と安定度	-2.42	0.02
先行モデルの流用と標準モデルの採用	2.19	0.03
開発言語(C 言語)	-2.13	0.04
業種(金融保険業)	2.10	0.04
プロジェクト管理の経験と能力	-1.88	0.06
実績 FP	-1.82	0.07

表 4 予測精度の比較

予測方法	BRE 平均	BRE 中央値	BRE 標準偏差
重回帰分析	0.17	0.26	0.26
平均値	0.22	0.31	0.32

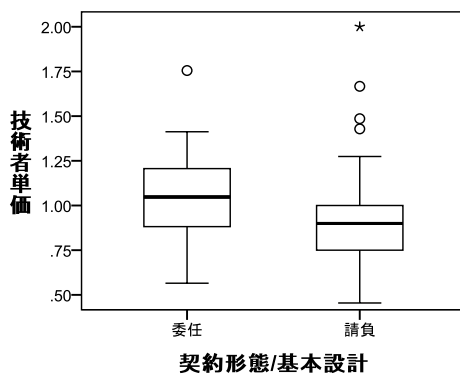


図 1 契約形態と技術者単価との関係
(重回帰分析データ)

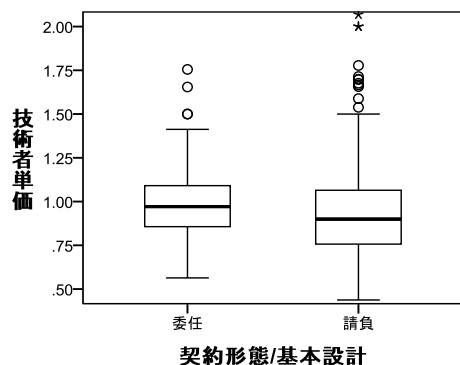


図 2 契約形態と技術者単価との関係

については分析から除外した（前者では電気・ガス・熱供給・水道業など、後者では ASP・ASP.NET など）。

4. 技術者単価の要因と基準値

本研究では、プロジェクトマネージャなど役割別の参画

割合を、受託者が知らないことを前提としているため、役割別の技術者単価を求めることは目的としていない。分析では、主に以下の 2 つを明らかにすることが目的である。

- 技術者単価は、(参画割合以外の) どの要因に着目すると推定可能なのか?
- 要因別に見た場合、技術者単価の基準値(中央値と分布)はどのようなものか?

4.1 会社規模と技術者単価との関係

技術者の所属する会社規模(従業員数)と技術者単価には関連がある可能性がある。そこで、スピーアマンの順位相関係数を用いて従業員数と技術者単価との関連を確かめた。従業員数、IT 技術者数(4 段階の数値で、数値が大きいほど IT 技術者が多いことを示す)と単価との相関係数を表 2 に示す。それぞれ正の相関であり p 値は 0.05 を下回っていたが、ともに単価に対する相関係数は小さかった。

ソフトウェア開発プロジェクトには従業員数以外にも単価に影響すると考えられる特性が含まれている。例えばプログラミング言語はプロジェクトによって異なり、また、ソフトウェアの対象業種もプロジェクトごとに異なる。次節で詳細に説明するが、それらの複数の特性を同時に考慮するために、重回帰分析を用いると、会社規模は説明変数として採用されなかった。これは、技術者単価を推定するためには、会社規模は必須ではなく、その他の特性があれば推定可能であることを示している。

なお、従業員数と IT 技術者数を説明変数として、重回帰分析を行うと、 R^2 は 0.03 となった。このことから、従業員数だけを基準に技術者単価を推定することは困難であるといえる。

4.2 重回帰分析による技術者単価の要因の特定

本節では、ソフトウェア開発プロジェクトにおける複数の要因(ソフトウェアの対象業種や開発言語など)を同時に考慮するために、技術者単価を説明変数として重回帰分析を行った。以下を直接的、または間接的に示すものを説明変数の候補とした。

1. 開発対象ソフトウェアの特徴
2. 技術者に求められるスキル
3. プロジェクト体制
4. その他、単価を変動させる要因

項目 1 は例えば「業種」や「開発言語」であり、これは項目 2、項目 3 に影響し、そこから単価に影響すると考えられる。項目 2 は例えば「アナリストの経験と能力」であり、スキルの高低が単価に影響する。項目 3 は「各役割の技術者の参画割合(3 章で定義)」であり、例えばプログラマーの参画割合が高いと、単価が低くなる可能性などがある。項目 4 は例えば「契約形態」であり、コスト増のリスクを考慮して単価を高めに設定するなど、単価への影響がありうる変数である。

これらの候補すべてを用いて変数選択法を適用すると、

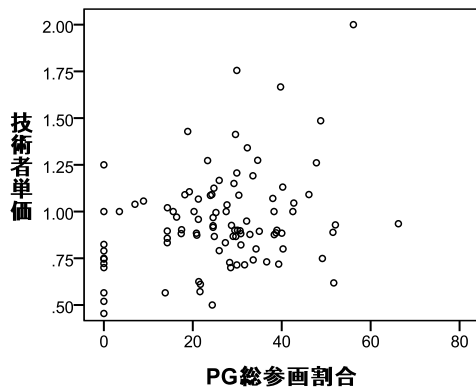


図 3 プログラマー総参画割合と技術者単価との関係（重回帰分析データ）

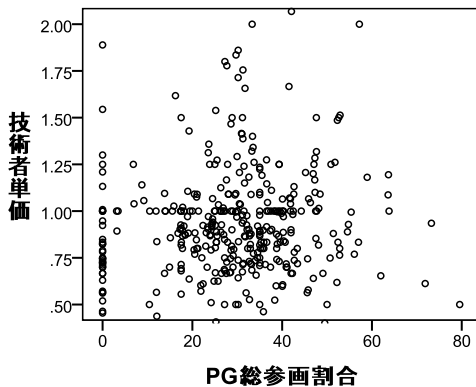


図 4 プログラマー総参画割合と技術者単価との関係（381 件のデータ）

リストワイズ除去（欠損値が 1 つでも含まれるケースは除外される）により、用いることができるケースが非常に少なくなる。そのため、予備分析により変数を絞り込み、その後に変数選択法により説明変数を絞り込んだ。その結果、表 3 に示す変数が採用された。以降では、標準化偏回帰係数（技術者単価への影響）が大きかった変数順に詳細に説明する。

PG 合計（技術者の参画割合）など、一部の説明変数は委託者が知ることができるとは限らないが、一方でそれらの変数を考慮せずに、技術者単価の影響要因を正しく分析することはできないため、説明変数に含めた。ただし 5.2 節で後述するように、委託者が表 3 に示す変数の一部しか知らない場合でも、分析結果を活用することが可能である。

モデルの説明力を表す R^2 は 0.44 となり、若干 0.5 を下回っていた。モデルの説明力を別の方法で評価するため、技術者単価の予測を行い、技術者単価の平均値を用いた場合と比べて、どの予測精度が高まるのかどうか（予測誤差が小さくなるのかどうか）を確かめた。予測精度の評価指標として、*BRE* (Balanced Relative Error)を採用した。*BRE*[7] は評価指標として広く用いられており ([6][10]など)、直感的には相対的な誤差を表している。*BRE* の算出にはリープ

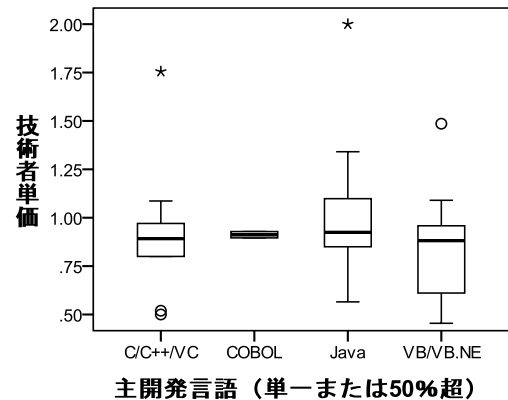


図 5 開発言語と技術者単価との関係（重回帰分析データ）

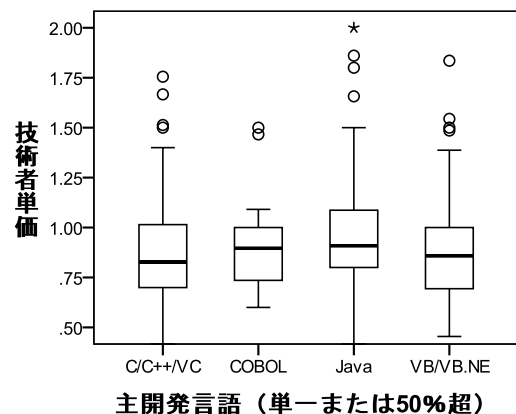


図 6 開発言語と技術者単価との関係（288 件のデータ）

ワンアウト法を用いた。

予測精度の評価結果を表 4 に示す。表に示すように、重回帰分析による予測のほうが、平均値による予測よりも誤差が小さくなっていた。このため、前者のほうが後者よりも説明力があり、より技術者単価の推定に適しているといえる。ただし、その差はあまり大きなものではないことに注意が必要である。

4.3 契約形態との関係

ソフトウェア開発では、開発工程ごとに契約形態が異なる場合がある。例えば要件定義工程では、どの程度開発工数が必要となるかが未確定の部分があるため、委任契約で実施する場合がある。逆にソフトウェアのプログラムを実際に作成する製造工程では、どんなソフトウェアを作るかの設計が確定しており、開発時間に関して不確定要因が小さいため、請負契約で行う場合がほとんどである。

表 3 の結果より、基本設計工程が請負か委任かが、技術者単価に影響しており、かつ単価への影響が最も大きいといえる。なお、73% (477 件の全プロジェクトにおいては 73%) のプロジェクトが、請負契約であった。偏回帰係数が負であることから、この工程が請負契約の場合、技術者単価が低くなるといえる。ソフトウェア開発に不確定要素

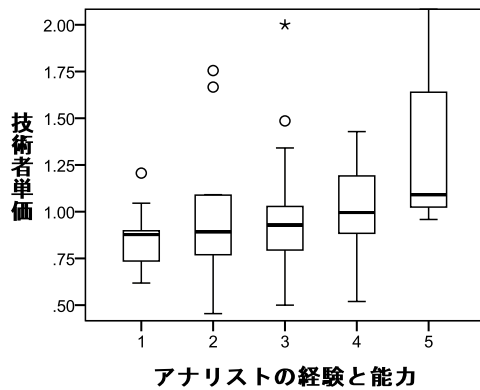


図 7 アナリストの経験・能力と技術者単価との関係（重回帰分析データ）

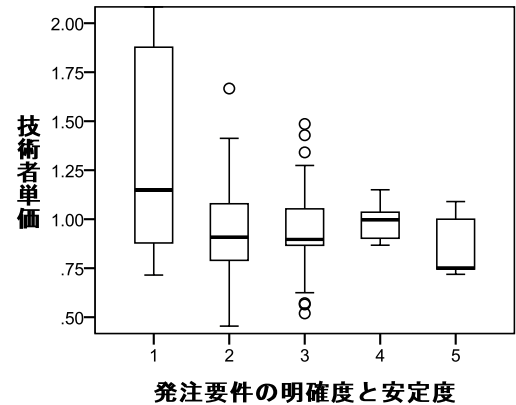


図 9 発注要件の明確度・安定度と技術者単価との関係（重回帰分析データ）

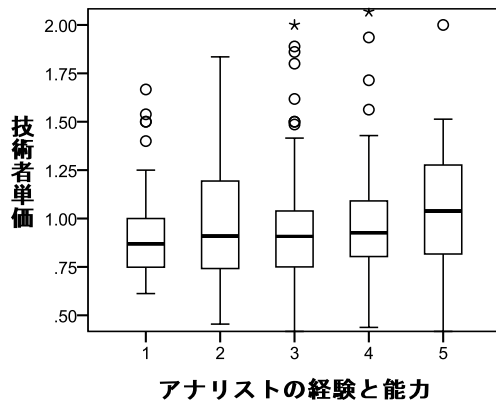


図 8 アナリストの経験・能力と技術者単価との関係（419 件のデータ）

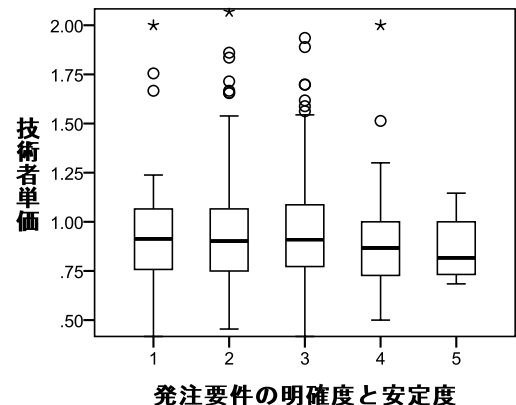


図 10 発注要件の明確度・安定度と技術者単価との関係（446 件のデータ）

が多い場合に、設計工程を委任契約にすることが考えられ、リスク発生時を考慮して単価が高くなっている可能性がある。

重回帰分析のデータのみを用いた場合と、すべてのデータを用いた場合の、技術者単価と契約形態との関係を図 1、図 2 に示す。どちらの場合においても、委任契約の場合、単価が高くなっていた。すなわち、重回帰分析のデータは特に偏っておらず、分析結果は妥当であると考えられる。

4.4 プログラマー総参画割合との関係

3 章の定義に基づくと、例えばプログラマー総参画割合は、開発に必要な作業時間（工数）のうち、何%をプログラマーが担当しているかを表している。一般に役割によって技術者単価が異なり、例えばプログラマーの総参画割合が増えると、プロジェクト全体の技術者単価の平均値（本研究における技術者単価：重回帰分析における目的変数）はプログラマーの技術者単価に近づく。

このため、各役割の総参画割合が（各役割の単価を平均した）技術者単価に影響することが考えられる。重回帰分析の結果においても、プログラマーの総参画割合が単価に影響する要因とみなされた（変数選択法の結果、説明変数とした採用された）。ただし、標準化偏回帰係数が正の値で

あることから、プログラマーの総参画割合が増加すると技術者単価が上昇することになり、直感に反する結果となった。

多重共線性が発生していないかを確かめるために、プログラマー総参画割合とその他の変数との相関係数を算出したが、特に相関係数は大きくなかったことから、多重共線性は影響していないと考えられる。

重回帰分析で用いたデータにおける、技術者単価とプログラマー総参画割合の散布図を図 3 に、全てのプロジェクトを用いた場合のものを図 4 に示す。前者の相関係数は 0.20 であり、後者の 0.09 よりも大きくなっていた。なお、後者についても負の相関は見られなかった。本節の分析結果は、重回帰分析に用いたプロジェクトのみの傾向である可能性があり、今後さらなる分析が必要である。

4.5 開発言語との関係

開発では複数の言語が用いられる場合がしばしばある。具体的には、開発全体の機能のうち 70%は COBOL で行い、残りの 30%は Java で行うことなどがある。本研究では 50%以上用いられている言語を主開発言語とし、これに着目して分析した。また、分析データにおいて使用している割合が高い言語（Java など）に着目して分析した。

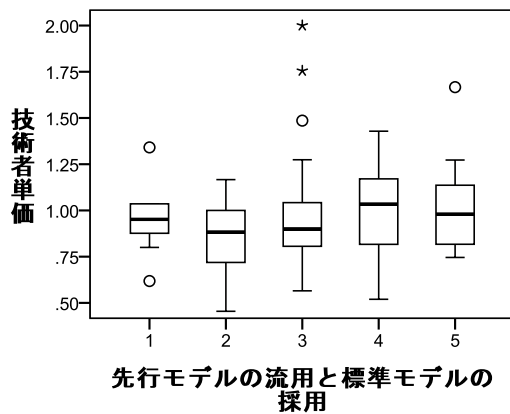


図 11 先行モデルの流用・標準モデルの採用と技術者単価との関係（重回帰分析データ）

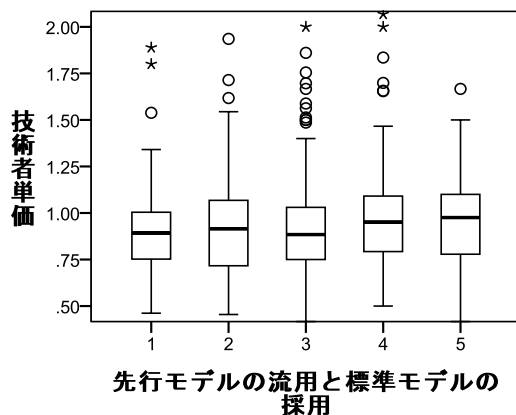


図 12 先行モデルの流用・標準モデルの採用と技術者単価との関係（446 件のデータ）

ソフトウェア技術者が、業務に必要とされるレベルで全ての開発言語を扱えることは少なく、技術者により用いることができる開発言語は異なる。そのため、ある言語が扱える技術者の需要と供給のバランスにより、技術者単価が異なることが考えられる。表 3 の重回帰分析の結果でも、C 言語を用いている場合と Visual Basic を用いている場合に技術者単価が変化する傾向が見られた。それぞれの標準化偏回帰係数は正と負であることから、前者を用いていると単価が高くなり、後者の場合は低くなることになる。

重回帰分析のデータのみを用いた場合と、すべてのデータを用いた場合の、技術者単価と開発言語との関係を図 5、図 6 に示す。前者の場合では、重回帰分析の結果と同様に C 言語を用いていると単価が高くなり、Visual Basic を用いている場合は低い傾向があった。ただし、後者の図では両者の間で技術者単価の違いは見られない。両グラフとも、Java の場合は若干技術者単価が高い傾向があり、開発言語は技術者単価に影響があると考えてよい。ただし、具体的にどの言語を用いると、どの程度技術者単価が変化するかについては、さらなる分析が必要である。

なお、技術者の役割によって、開発言語が技術者単価に

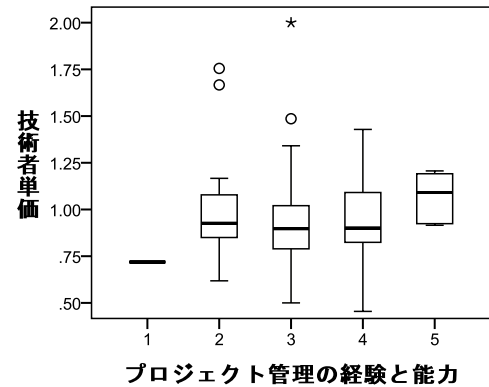


図 13 プロジェクト管理の経験・能力と技術者単価との関係（重回帰分析データ）

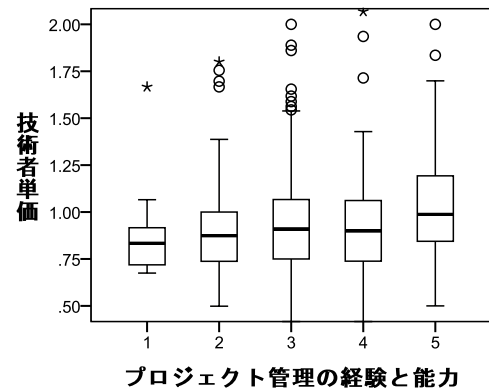


図 14 プロジェクト管理の経験・能力と技術者単価との関係（446 件のデータ）

与える影響は異なると考えられる。例えばプログラマーの場合は開発言語を実際に用いて開発する必要があるが、その言語を用いる能力が求められるが、プロジェクトマネージャが実際に言語を用いて開発することは少ない。分析に用いたデータでは、製造工程へのプロジェクトマネージャの参画割合は 7%となっている。そのため、プロジェクトマネージャの技術者単価に関して、開発言語の影響は小さいと考えられる。ただし、開発言語が開発対象システムの特徴を間接的に示すことはありえる（COBOL が用いられている場合、事務処理系のシステムであり、かつ長期間用いられているシステムであるなど）ため、全く影響がないともいえない。

4.6 アナリストの経験と能力との関係

経験豊富なアナリストが参画している場合、技術者単価が高くなることが想定される。データには、プロジェクトにおけるアナリストの経験と能力の 5 段階評価（数値が大きいほど経験と能力が豊富）が記録されており、これを用いた。重回帰分析より、アナリストの経験と能力に応じて、技術者単価が変化するといえ、表 3 の偏回帰係数が正であることより、経験と能力が高いほど技術者単価が高くなる傾向があるといえる。

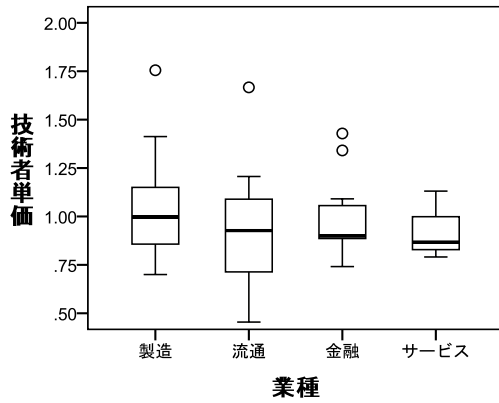


図 15 業種と技術者単価との関係
(重回帰分析データ)

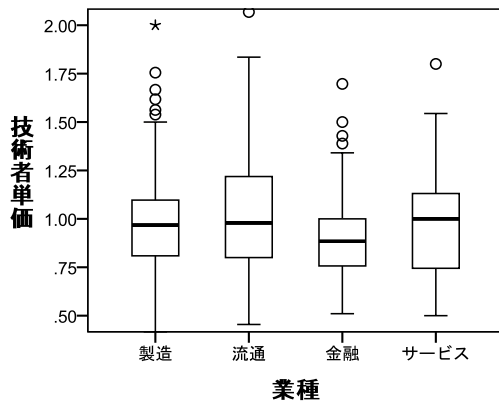


図 16 業種と技術者単価との関係
(333 件のデータ)

重回帰分析のデータのみを用いた場合と、すべてのデータを用いた場合の、技術者単価とアナリストの経験と能力との関係を図 7、図 8 に示す。どちらの場合においても、重回帰分析の結果と同様に、経験や能力が高いほど技術者単価が高い傾向が見られた。

4.7 発注要件の明確度と安定度との関係

ソフトウェアの発注要件が不明確、不安定である場合、ベンダーが開発時のリスク発生時を考慮し、単価を高く設定する可能性がある。データには、発注要件の明確度と安定度を 5 段階で評価したもの（数値が大きいほど安定、明確であることを示す）が記録されており、これを分析に用いた。重回帰分析の結果、発注要件の明確度と安定度に応じて、技術者単価が変化するといえ、表 3 の偏回帰係数が負であることより、発注要件の明確度と安定度が高いほど、技術者単価が低くなる傾向があった。

重回帰分析のデータのみを用いた場合と、すべてのデータを用いた場合の、技術者単価と発注要件の明確度と安定度との関係を図 9、図 10 に示す。どちらの場合においても、重回帰分析と同様の傾向が見られた。

4.8 先行モデルの流用と標準モデルの採用との関係

ソフトウェア開発では、既存のモデルを利用することが

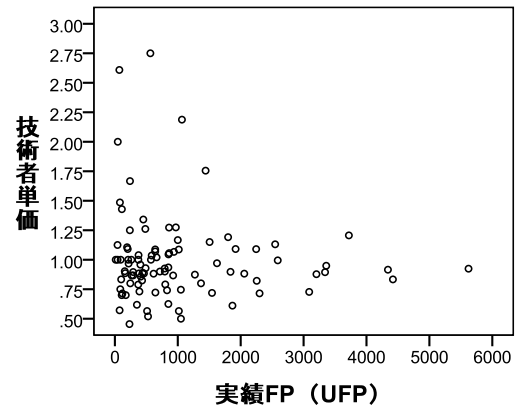


図 17 開発規模と技術者単価との関係
(重回帰分析データ)

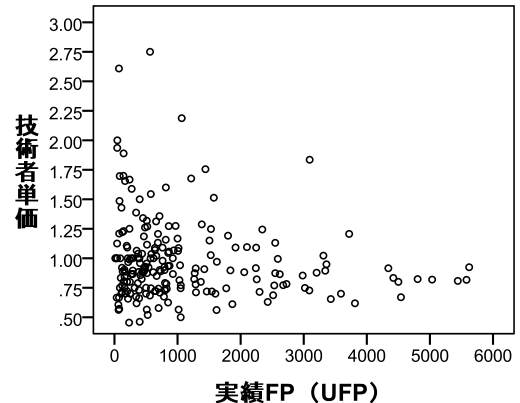


図 18 開発規模と技術者単価との関係
(220 件のデータ)

ある。これにより、開発の効率や品質を高めること期待できる一方、そのようなモデルを適用できるようなプロジェクトでは、高いスキルの技術者が要求され、結果的に技術者単価が高まる可能性もある。先行モデルの流用と標準モデルの採用の程度を 5 段階で評価したもの（数値が大きいほど、既存モデルを利用できたことを示す）が記録されており、これを用いた。

重回帰分析の結果より、先行モデルの流用と標準モデルの採用の程度に応じて、技術者単価が変化するといえ、表 3 の偏回帰係数が負であることより、流用・採用の程度が高いほど、技術者単価が高くなる傾向があるといえる。重回帰分析のデータのみを用いた場合と、すべてのデータを用いた場合の、技術者単価と先行モデルの流用と標準モデルの採用の程度との関係を図 11、図 12 に示す。どちらの場合においても、重回帰分析の結果と同様に、流用・採用の程度が高いほど技術者単価が高い傾向が見られた。

4.9 プロジェクト管理の経験と能力との関係

技術者（プロジェクトマネージャ）のプロジェクト管理の経験と能力が高い場合、技術者単価も高くなると思われる。データには、プロジェクト管理の経験と能力を 5 段階で評価したものが記録されており、これを用いた。重回帰

分析では、プロジェクト管理の経験と能力に応じて、技術者単価が変化するという結果となった。ただし表 3 の偏回帰係数は負であり、経験と能力が高いと技術者単価が低くなるという、直感に反する結果となった。

重回帰分析のデータのみを用いた場合と、すべてのデータを用いた場合の、技術者単価とプロジェクト管理の経験と能力との関係を図 13、図 14 に示す。どちらの場合においても、中央値（箱中の太線位置）については、経験と能力が高いほど技術者単価が高い傾向があり、重回帰分析と逆の傾向が見られる。詳細に見ると、重回帰分析の対象データ（図 13）では、経験と能力が 1 または 5 のプロジェクト数が少なく、2 から 4 のプロジェクトでは、経験と能力が高くなるとわずかに中央値が低くなっていた。これが重回帰分析の結果とグラフでの傾向が不一致となった原因であると考えられる。

4.10 業種との関係

開発するソフトウェアの対象業種により求められる機能や品質が異なる。例えば金融業を対象とするソフトウェアの場合、比較的高い品質が求められることが多い。また、開発規模も業種によって異なる。業種によりソフトウェアの特徴が異なり、それにより必要とされる技術者も異なるため、業種が技術者単価に影響していると考えられる。

用いたデータにおいて、割合が高かった業種を分析対象とした。重回帰分析の結果より、金融保険業を対象とするシステムの場合、技術者単価が変化する傾向が見られた。表 3 に示すように、偏回帰係数が正であることから、金融保険業の場合、技術者単価が他の業種と比較して高くなる傾向であることになる。

重回帰分析のデータのみを用いた場合と、すべてのデータを用いた場合の、技術者単価と業種との関係を図 15、図 16 に示す。どちらの場合においても、中央値では、金融保険業の技術者単価が低くなっていた。重回帰分析の対象データ（図 15）では、金融保険業の技術者単価の中央値以下のプロジェクトが少なく、技術者単価の（中央値ではなく）平均値で見ると、他の業種よりも高くなっていた。これが重回帰分析の結果とグラフでの傾向が不一致となった原因であると考えられる。

なお、金融保険業のシステムは比較的大規模であり、契約金額が大きくなりやすく、結果的に技術者単価が低くなりやすい可能性がある。例えば技術者 1 人、技術者費用 100 万円、その他費用 10 万円のプロジェクトの場合、技術者単価は 110 万円になるが、技術者 10 人、技術者費用 1000 万円、その他費用 10 万円のプロジェクトの場合、技術者単価は 101 万円と低くなる。

4.11 開発規模との関係

ソフトウェアの開発規模（開発する機能量）はプロジェクトにより大きく異なる。開発規模が大きい場合と小さい場合とでは、プロジェクトの体制（プロジェクトマネー

ジャの必要性や、必要な技術者数など）や、ソフトウェアに求められる信頼性なども傾向が異なる可能性がある。

重回帰分析においても、開発規模（実績 FP）が技術者単価に影響しているという結果になった。表 3 に示すように、偏回帰係数が負の値となったことから、開発規模が大きくなると、技術者単価がやや低下する傾向であるといえる。これは前節で述べた、金融保険業と技術者単価との関係と同様の理由であると考えられる。なお、表では開発規模の偏回帰係数が最も小さく、分析に用いた変数の中では最も単価への影響が小さいことに留意する必要がある。

重回帰分析のデータのみを用いた場合と、すべてのデータを用いた場合の、技術者単価と開発規模との関係を図 17、図 18 に示す。どちらの場合においても、実績 FP（開発規模）がおおよそ 2000 以上の場合、技術者単価が大きい（1 を大きく超える）プロジェクトが少ない傾向が見られた。これは重回帰分析の結果と合致しているといえる。

5. 考察

5.1 妥当性への脅威

内的妥当性: 本研究で用いた説明変数以外に、価格に影響する要因が存在する可能性がある。ただし、モデルの説明力を表す R^2 は 0.44 であり、目的変数（技術者単価）が、ある説明変数と明確な関連がないことを考慮すると、用いた説明変数で、目的変数をかなり説明できている（その他に価格に影響する要因が存在するとしても、その影響はあまり大きくない）と考えられる。ここで、目的変数と説明変数に明確な関連があるとは、例えば工数見積もりにおける工数（目的変数）と開発規模（説明変数）との関係を指し、このような場合はモデルの説明力が高くなりやすい。一方で、生産性を目的変数とすると、一般には、目的変数に明確に関連する説明変数は存在しないため、説明力は低下しやすい。

外的妥当性: 本研究では、従業員数が多い企業から少ない企業まで、多様な企業から横断的に収集されたデータを用いて分析している。表 1 で示したように業種の偏りも比較的小さく、さらに、二変量解析で用いたデータは 200 から 400 件程度であり、本研究の分析結果は比較的外的妥当性が高いと考えられる。ただし、データの収集対象企業は国内企業のみであり、日本国外の企業に対する、分析結果の外的妥当性は不明である。日本国外の企業における外的妥当性を高めることは今後の課題である。

5.2 分析結果の活用方法

本研究の分析結果をソフトウェア価格の推定に活用するための、標準的な手順を以下に示す。

1. 技術者単価に関連のある要因（開発言語など）の各箱ひげ図から、自社が当てはまっているデータで最も箱の大きさ（データの散らばり）が小さいものを選ぶ。

2. 二変量解析による箱ひげ図を参考に、単価を決定する。例えば図 6 において、開発言語 (C 言語) の場合の技術者単価の中央値は約 0.9 であるため、開発言語が C 言語ならば 1 人月あたりの単価を 90 万円とする。
3. COCOMO[1]などの工数見積もり方法を用いて、見積工数を人月単位で算出する。
4. 手順 2 で求めた単価に手順 3 で求めた工数を乗じ、推定価格とする。

ただし、単価の分布には幅があり、開発言語が C 言語の場合の妥当な単価は、75 万円 (箱の下側) である可能性もあれば、100 万円 (箱の上側) の可能性もあることに注意が必要である。同様に、手順 3 で求めた工数にも誤差が含まれることに注意し、推定価格を絶対視しすぎるべきではない。手順 3 以降の代わりに、1 章で述べた価格推定方法を用いることも可能である。

6. おわりに

本研究では、ソフトウェア技術者の単価に影響する要因を明らかにするとともに、要因が変化した場合、(1 人月あたりの) 単価がどの程度変化するかについても示した。要求分析、基本設計、詳細設計、製造、試験の開発工程が実施されているプロジェクト 99 件に対して重回帰分析を行った結果、以下の傾向が見られた。

- 基本設計工程が請負の場合、技術者単価が低い傾向がある。
- アナリストの経験と能力が高い場合、技術者単価も高い傾向がある。
- 発注要件の明確度と安定度が高い場合、技術者単価は低い傾向がある。
- 先行モデルの流用と標準モデルの採用の程度が高い場合、技術者単価も高い傾向がある。
- ソフトウェアの規模が大きい場合、技術者単価は低い傾向がある。

上記については、箱ひげ図を用いて 200 から 400 件のデータを分析した場合でも、同様の傾向が見られた。

分析では上記の要因の違いによる技術者単価の分布についても示した。この分布を用いて、どの要因が変化すると、どの程度単価が変化するかを、大まかにではあるが推定することができる。これはソフトウェア開発の委託者(企業)が価格の妥当性を検討する際に有用な情報となりうる。ただし、箱ひげ図からもわかるように、技術者単価の分布は比較的広いため、技術者単価の中央値は絶対視すべきでなく、あくまで参考にとどめるべきである。

今後の課題は、さらに分析対象のデータを増加させ、委託者による技術者単価の推定精度を高めることである。

謝辞 本研究の一部は、日本学術振興会科学研究費補助

金 (基盤 C : 課題番号 21K11840, 基盤 S : 課題番号 20H05706) による助成を受けた。

参考文献

- [1] B. Boehm, Software Engineering Economics, Prentice Hall (1981).
- [2] 情報サービス産業協会: 平成 21 年度 新たな価格モデルに関する調査研究報告書 (2010).
- [3] 情報処理推進機構: ソフトウェア開発データ白書 2016-2017, 独立行政法人情報処理推進機構 (2016).
- [4] 経済産業省: 特定サービス産業実態調査報告書 -ソフトウェア業, 情報処理・提供サービス業及びインターネット附随サービス業編-, 経済産業省 (2019).
- [5] Y. Liu, H. Cheng, Q. Tang, and E. Eryarsoy: Optimal software pricing in the presence of piracy and word-of-mouth effect, Decision Support Systems, vol.51, no.1, pp.99-107 (2011).
- [6] T. Menzies, E. Kocaguneli and J. Keung: On the Value of Ensemble Effort Estimation, IEEE Transactions on Software Engineering, vol.38, no.6, pp.1403-1416 (2012).
- [7] Y. Miyazaki, M. Terakado, K. Ozaki, and H. Nozaki: Robust Regression for Developing Software Estimation Models, Journal of Systems and Software, vol.27, no.1, pp.3-16 (1994).
- [8] M. Murtojärvi, J. Järvinen, M. Johnsson, T. Leipälä, and O. Nevalainen: Determining the Proper Number and Price of Software Licenses, IEEE Transactions on Software Engineering, vol.33, no.5, pp.305-315 (2007).
- [9] 内閣官房 情報通信技術 総合戦略室: デジタルガバメント推進標準ガイドライン, 内閣官房 (2014).
- [10] Y. Seo, and D. Bae: On the value of outlier elimination on software effort estimation research, Empirical Software Engineering, vol.18, no.4, pp.659-698 (2013).
- [11] M. Shepperd, C. Schofield and B. Kitchenham, "Effort estimation using analogy," Proc. of International Conference on Software Engineering (ICSE), pp.170-178 (1996).
- [12] M. Tsunoda, A. Monden, K. Matsumoto, S. Ohiwa, and T. Oshino: Analysis of Attributes Relating to Custom Software Price, Proc. International Workshop on Empirical Software Engineering in Practice (IWESEP), pp.16-22 (2012).
- [13] 上野秀剛, 亀井靖高, 門田暁人, 松本健一: 原価率とプロジェクトメトリクスに着目したソフトウェア開発プロジェクトの特徴分析, プロジェクトマネジメント学会誌, vol.12, no.5, pp.25-30 (2010).