

自主研究

ソフトウェア開発における工程別生産性に関する分析 ～生産性変動要因に基づくリスク管理・予測に向けて～

ソフトウェア開発における工程別生産性に関する分析 ～生産性変動要因に基づくリスク管理・予測に向けて～

松本 健一 奈良先端科学技術大学院大学 情報科学研究科
大岩佐和子 押野 智樹 一般財団法人 経済調査会 調査研究部 第二調査研究室

1 はじめに

本稿は、著者らが平成23年に公表した論文(参考文献[1]、以下、文献[1]という)の続編である。同論文では、その前年に公表された2つの調査結果(参考文献[2]、[3])などに基づき、「開発期間」、「開発言語」、及び、「開発方法」に着目して、ソフトウェア開発の各工程における生産性の分析を行った。その結果として、開発期間等が生産性に与える影響や工程間での生産性の差異などに関するいくつかの知見を得ることができた。

本稿では、文献[1]の公表後に収集されたソフトウェア開発プロジェクトデータを分析対象に加えると共に、生産性に関する分析結果を、ソフトウェア開発におけるリスク管理・予測へと活用することを試みる。具体的には、経済調査会が定義した10個の生産性変動要因に基づいて、ソフトウェア開発の各工程における生産性を詳細に分析すると共に、生産性低下による工数超過といった開発リスクの回避に向け、開発後半となるテスト工程における工数密度の予測を試みる。なお、生産性の指標には、文献[1]と同様に、「工数密度」(最終成果物の規模を表す「実績ファンクションポイント(FP)規模」100FPあたりの工数)を用いることとする。

2 利用データ(分析対象データ)

分析に用いるのは、経済調査会が平成19年度から

26年度に実施した「ソフトウェア開発に関する調査」で収集したプロジェクトデータ(ソフトウェア開発データリポジトリ)である。同調査は、ソフトウェア開発における生産性、工数、費用に及ぼす要因の特定などを通じて、ソフトウェア開発の実態を明らかにし、その成果を公表することを目的として、平成10年度からほぼ毎年実施している。同調査では、分析用データとして平成13年度から26年度までの14年度分延べ2,009プロジェクトのデータを蓄積しているが、本稿で用いるのは、平成19年度から26年度の同調査で収集されたデータのうち、次の条件を満足する141プロジェクトのデータである。

- ・経済調査会が共通フレームと対応付けし、定義した開発工程区分のうち、基本5工程(基本設計、詳細設計、プログラム(以下PGと略す)設計製造、結合テスト、総合テスト(ベンダ確認))*¹がすべて実施され、各工数が記されている。
- ・ソフトウェアの規模を表すデータ「実績FP規模」が記されている。
- ・10個の生産性変動要因すべてについて生産性区分(同要因に対する主観評価値)がそれぞれ記されている。生産性変動要因の定義と生産性区分については付録を参照。
- ・開発全体の工数値と基本5工程の工程別工数合計値の乖離が30%以内である。(外れ値の除去が目的。誤記入の可能性が高いため。)
- ・基本5工程全体における工数密度の対数値が、平

*¹ 一般にソフトウェア開発では、設計を行い(設計工程)、設計に従ってプログラムを作成し(製造工程)、プログラムの動作をテストする(テスト工程)ことが実施される。経済調査会では、ソフトウェア開発の工程を共通フレームと対応付けて、基本設計(A)、基本設計(B)、詳細設計、PG設計・製造、結合テスト、総合テスト(ベンダ確認)、総合テスト(ユーザ確認)に区分した。各工程と共通フレームの対応関係については「月刊積算資料」(2011年3月号以降)を参照されたい。なお、本稿では基本設計(A)、基本設計(B)をまとめて基本設計とし、基本設計から総合テスト(ベンダ確認)までの基本5工程を分析対象とした。

均値±標準偏差×3の範囲である。(外れ値の除去が目的。工数密度の対数値は正規分布に従うと見なすことができる。この基準により除外されたプロジェクトは2件であった。)

分析に用いた主なプロジェクト特性値は、規模(実績FP規模)、工数、生産性変動要因、そして、生産性指標として(100FPあたりの)工数密度、の計4個である。これら特性値の定義を図表1に、規模、工数、工数密度の基本統計量を図表2に、生産性変動要因それぞれにおける生産性区別のプロジェクト件数を図表3に、それぞれ示す。なお、図表3の各区分において、該当するプロジェクト数が、分析対象である141プロジェクトの約5%にあたる7プロジェクト未満である場合は、濃い網掛けで図示し、以降では、原則、それ単独では分析対象とはしないものとする。

また、本稿では、先述の通り、生産性指標として、一般的な「規模/工数」ではなく、その逆数となる工数密度を用いている。ソフトウェア開発における生産性は、開発工程で生成・利用されたプロダクトの規模を開発工数で除した値を用いることが多い。しかし、本稿では工程別の生産性について論じようとしており、収集したプロジェクトデータには各工程で生成されたプロダクト規模データが含まれておらず、また把握も困難である。そのため、最終成果物の規模である「実績FP規模」を用いることとした。当然のことながら、実績FP規模は、工程に関係なくプロジェクト全体で一定である。この実績FP規模を分母にすることで、規模の異なるプロジェクト間で工数を比較できるだけでなく、工程間の工数が比較できることとなる。とすれば、「規模/工数」ではなく、「工数/規模」とする方が自然である。

図表1 分析に用いたプロジェクト特性値

プロジェクト特性値	単位	定義
規模	FP	未調整ファンクションポイントの実績値(実績FP規模)
工数	人月	ソフトウェア開発の実績工数
生産性変動要因	-	本稿「付録」を参照
(100FPあたりの)工数密度 【生産性指標】	人月/100FP	100FPあたりの工数を表す。生産性指標として工程毎に算出 (100FPあたりの)工数密度 = (当該工程における)工数 / (FP規模 / 100)

図表2 プロジェクト特性値の基本統計量

	平均値	標準偏差	最小値	第1四分位数	中央値	第3四分位数	最大値
規模(実績FP規模)	1,751	3,053	23	411	792	1,600	26,572
工数(基本5工程)	133	257	2	22	48	132	1,954
工数密度(基本5工程)	7.96	5.89	1.13	3.88	6.31	9.94	29.82

図表3 生産性変動要因の生産性区別のプロジェクト件数

生産性変動要因		生産性区分						
		1	1 & 2	2	3	4	4 & 5	5
委託者側	機能性	9	48	39	79	14	14	0
	信頼性	0	18	18	81	30	42	12
	プラットフォーム	1	13	12	91	25	37	12
	開発スケジュール要求	4	21	17	86	26	34	8
	発注要件の明確度と安定度	13	86	73	41	7	14	7
	ユーザの参画割合	3	21	18	72	29	48	19
受託者側	先行モデルの流用と標準モデルの採用	17	35	18	71	32	35	3
	プロジェクト管理者の経験と能力	2	29	27	61	40	51	11
	アナリストの経験と能力	15	45	30	64	31	32	1
	SE・プログラマの経験と能力	6	22	16	61	36	58	22

3 工程別工数密度

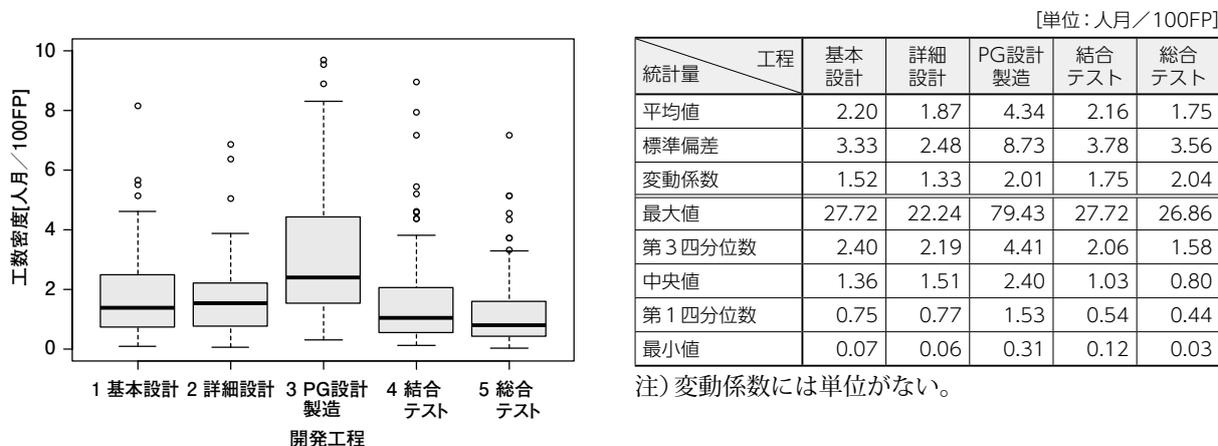
ここでは、まず、基本5工程における工数密度を見ていく。図表4 (a) は、文献 [1] からの再掲で、平成19年度から21年度に収集された92プロジェクトにおける工数密度の基本統計量と箱ひげ図を示す。同じく図表4 (b) は、本稿で分析対象とする141プロジェクトにおける工数密度であり、収集期間は平成19年度から24年度となる。収集期間からも分かるように、後者は、前者のプロジェクトに平成22年度から24年度のプロジェクトを追加したものであるが、「生産性区分(同要因に対する主観評価値)が記されていない生産性変動要因がある」プロジェクトや工数密度が外れ値となるプロジェクトは除外されている。

分布状況を表す「最大値」から「最小値」までに着目すると、両者間で大きく異なるのは「最大値」のみで、「第3四分位数」から「最小値」までに大きな変化は見

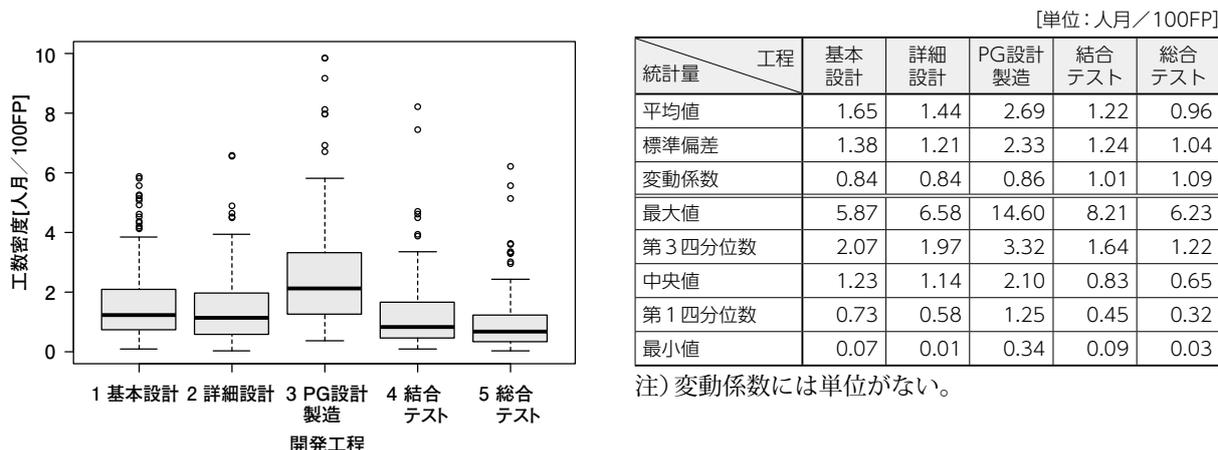
られない。一方、図表4 (a) と比較して図表4 (b) は平均値や標準偏差が小さくなっている。これは、開発全体の工数値と基本5工程の工程別工数合計値の乖離が30%を超えるプロジェクトは工数密度のばらつきが大きく、本稿における分析対象には、そのデータが含まれていないことが理由と考えられる。

本稿で分析対象とする141プロジェクトについても少し詳しく見てみる。図表4 (b) に示された平均値を見ると、PG設計製造工程の工数密度(2.69)が、他の工程の2倍程度と突出している。一方、変動係数で見ると、基本設計工程からPG設計製造工程まではほとんど変化がない(0.84~0.86)が、結合テスト工程に入ると約2割大きくなり(1.01)、総合テスト工程でさらに約1割大きくなっている(1.09)。テスト工程(結合テスト工程+総合テスト工程)では、工数密度そのものは決して高くないが、そのばらつきは相対的に大きく、生産性低下による工数超過のリスクが高いと見

図表4 基本5工程における工数密度



(a) 平成19年度から21年度に収集した92プロジェクト(文献[1]における分析対象)



(b) 平成19年度から24年度に収集した141プロジェクト(本稿における分析対象)

ることができる。

4 生産性変動要因が工数密度に与える影響

4-1 基本5工程全体

基本5工程全体における平均工数密度を、10個の生産性変動要因それぞれの生産性区別に算出した結果を図表5に示す。同要因は、ソフトウェア開発に関わる委託者側の要因6個と受託者側の要因4個からなる。また、生産性区分は、同要因に対する1から5までの主観評価値である。いずれの要因においても、値が小さいほど、生産性に与える影響が大きい。すなわち、同要因により生産性が大きく低下したと調査の回答者が評価したことを意味する。改めてまとめると次の通りである。

区分「1」:同要因の生産性に与える影響は極めて大きかった(生産性が大きく低下した。工数密度が高くなった)。

区分「2」:同、影響は大きかった。

区分「3」:同、影響は平均的、標準的であった。

区分「4」:同、影響は小さかった。

区分「5」:同、影響は極めて小さかった。

また、同図表においては、各要因の区分「3」、つまり、

生産性に与える影響は平均的、標準的であったプロジェクトにおける平均工数密度を1として、その相対値によって平均工数密度を示している。例えば、同図表において、要因「機能性」の区分「1」に対する値は1.08となっているが、これは、同区分に属するプロジェクトにおける平均工数密度が、同じく区分「3」に属するプロジェクトの1.08倍であったことを示す。

また、図表5は図表4と同様に、区分に属するプロジェクト数が7プロジェクト未満である場合は、背景色を濃い網掛けで表示し、当該区分単独では分析対象外であることを示している。同図表より、10個の要件のうち9個については、区分「1」もしくは「5」が分析対象外であることが分かる。特に、要因「機能性」の区分「5」と要因「信頼性」の区分「1」には、該当するプロジェクトが存在しない。ただし、数が少ないとはいえ、生産性に対する影響が極めて大きい、あるいは、小さいと評価されたプロジェクトである。そもそも、この区分は主観評価によるものであり、「1」と「2」、「4」と「5」の間に大きな差があるとは限らない。そこで、区分「1」と区分「2」に属するプロジェクトを統合した区分「1 & 2」、同じく、区分「4」と区分「5」に属するプロジェクトを統合した区分「4 & 5」を追加し、その平均工数密度の相対値をそれぞれ示している。

区分「1 & 2」と区分「3」の間及び区分「3」と区分「4

図表5 基本5工程全体における平均工数密度
(生産性変動要因別:生産性区分「3」での平均値を「1」とした相対値)

生産性変動要因		生産性区分						
		1	1 & 2	2	3	4	4 & 5	5
委託者側	機能性	1.08	1.26	1.30	1.00	0.98	0.98	—
	信頼性	—	1.65	1.65	1.00	0.88	0.87	0.84
	プラットフォーム	1.24	1.13	1.12	1.00	1.63	1.37	0.84
	開発スケジュール要求	1.02	1.39	1.47	1.00	0.94	0.91	0.80
	発注要件の明確度と安定度	0.88	1.14	1.19	1.00	0.85	0.74	0.62
	ユーザの参画割合	0.77	1.25	1.33	1.00	1.14	1.09	1.01
受託者側	先行モデルの流用と標準モデルの採用	1.19	1.16	1.13	1.00	1.01	0.97	0.56
	プロジェクト管理者の経験と能力	1.26	1.16	1.15	1.00	1.10	1.20	1.55
	アナリストの経験と能力	1.41	1.19	1.08	1.00	1.39	1.37	0.84
	SE・プログラマの経験と能力	1.14	1.33	1.40	1.00	1.22	1.20	1.18

生産性区分 1:当該要因が生産性に与えた影響は極めて大きかった。

2:同、影響は大きかった。

3:同、影響は平均的、標準的であった。

4:同、影響は小さかった。

5:同、影響は極めて小さかった。

区分「3」における平均値との統計的有意差 太文字・斜体・下線:有意水準5%で有意差あり
太文字:有意水準10%で有意差あり

& 5」の間で、工数密度の平均値に統計的に有意な差があるかどうかをダネット検定(多重比較)で確かめた。有意水準5%で有意差が認められた平均工数密度(の相対値)を「太文字・斜体・下線」で、有意水準10%で有意差が認められた平均工数密度(の相対値)を「太文字」で、それぞれ示す。なお、有意差が認められたいずれの要因、区分も、「区分「3」よりも平均工数密度が有意に高い」と判定されている。すなわち、5工程全体で見ると、生産性変動要因や区分に関わらず、区分「3」よりも平均工数密度が低くなることはなかった、ということになる。

ここで、区分「3」における工数密度を「標準」と見なすと共に、少し緩やかな基準ではあるが、有意水準10%で平均工数密度が区分「3」、すなわち、標準よりも有意に高いと認められた区分をH(High)、有意差が認められなかった区分をM(Moderate)、標準よりも有意に低いと認められた区分をL(Low)、とそれぞれ表すこととする。先述の通り、区分「1」から区分「5」に進むにつれて、生産性変動要因が生産性に与える影響(工数密度が高くなる可能性)は小さくなる(と、回答者は評価している)。従って、同要因が工数密度に与える影響が十分に大きければ、区分「1 & 2」、区分「3」、区分「4 & 5」の順の三つ組み(パターン)で考えると、「H-M-L」となることが期待される。図表5の結果に基づき、このパターンで10個の生産性変動要因を分類した結果を図表6に示す。

平均工数密度が標準よりも有意に低いと認められた区分、すなわち、Lと表される区分は存在しなかったことから、生産性変動要因と工数密度の間で想定される関係に比較的近いパターンは、「H-M-M」となる。このパターンに該当する生産性変動要因は、「信頼性」と「開発スケジュール要求」の2つである。これら2つの要因が区分「1」や「2」に属するソフトウェア開発プロジェクトでは、工数密度が標準よりも高くなるため、生産性低下による工数超過等のリスクに備えるべきということになる。一方、それら要因が区分「4」や「5」に属するからといって、工数密度が標準よりも有意に低くなるわけではないことも、リスク管理を行う上で有益な知見の一つと言える。

パターン「M-M-M」は、いずれの区分においても、

工数密度は標準と有意な差はない、すなわち、生産性変動要因と工数密度の間に強い関係は見られないことを表す。このパターンに該当する生産性変動要因は、「機能性」、「発注要件の明確度と安定度」、「ユーザの参画割合」、「先行モデルの流用と標準モデルの採用」、「プロジェクト管理者の経験と能力」、及び、「SE・プログラマの経験と能力」の6つである。これら6つの生産性変動要因については、基本5工程全体としては工数密度との間に強い関係が認められなかっただけで、特定の工程においては、関係性が認められる可能性がある。詳しくは、4-3で述べる。

パターン「M-M-H」は、工数密度が高くなる可能性が小さい区分「4」や「5」において、工数密度がむしろ標準よりも高くなっており、生産性変動要因と工数密度の間で想定される関係とは合致していない。このパターンに該当する生産性変動要因は、「プラットフォーム」と「アナリストの経験と能力」の2つである。これら2つの要因については、工数密度との関係が、想定するような単純なものではない可能性がある。つまり、該当するプロジェクトが他の生産性変動要因によって大きく影響を受けている可能性があるため、さらに詳しく分析する必要がある。

4-2 工程別

ここでは、パターン「H-M-M」に該当する2つの生産性変動要因について、工程別に平均工数密度を見てみる。それら2つの要因について、工程別に平均工数密度を算出した結果を図表7に示す。ただし、図表5と同様に、各要因の区分「3」、つまり、生産性に与える影響は平均的、標準的であったプロジェクトにおける平均工数密度を1として、その相対値によって平均工数密度を示している。また、同じく、ダネット検定(多重比較)の結果、有意水準5%で標準との有意差が認められた平均工数密度(の相対値)は「太文字・斜体・下線」で、有意水準10%で標準との有意差が認められた平均工数密度(の相対値)は「太文字」で、それぞれ示す。

同図より、要因「信頼性」が区分「1」や「2」に属するソフトウェア開発プロジェクト、すなわち、「ソフト

図表6 基本5工程全体における、生産性変動要因と工数密度の関係パターン

パターン	生産性変動要因と工数密度との関係	該当する生産性変動要因
H-M-M	区分「1」や「2」では、工数密度は標準より高いが、区分「4」や「5」では標準との有意差はない。生産性変動要因と工数密度の間で想定される関係とある程度合致する。	信頼性 開発スケジュール要求
M-M-M	いずれの区分においても、工数密度に標準と有意差はない。生産性変動要因と工数密度の間には強い関係は見られない。	機能性 発注要件の明確度と安定度 ユーザの参画割合 先行モデルの流用と標準モデルの採用 プロジェクト管理者の経験と能力 SE・プログラマの経験と能力
M-M-H	区分「1」や「2」では、工数密度に標準との有意差はないが、区分「4」や「5」では標準より高い。生産性変動要因と工数密度の間で想定される関係と著しく合致しない。	プラットフォーム アナリストの経験と能力

図表7 生産性区分別・工程別の平均工数密度
(生産性区分「3」での平均値を「1」とした相対値)

生産性変動要因：信頼性

生産性区分	工程	基本設計	詳細設計	PG設計製造	結合テスト	総合テスト	5工程全体
1 & 2		1.45	1.38	1.60	2.24	1.91	1.65
4 & 5		0.84	0.71 ↓	0.98	1.02	0.67	0.87

生産性変動要因：開発スケジュール要求

生産性区分	工程	基本設計	詳細設計	PG設計製造	結合テスト	総合テスト	5工程全体
1 & 2		1.35	1.25	1.47	1.71	1.03	1.39
4 & 5		0.92	0.89	0.88	0.89	1.02	0.91

区分「3」における平均値との統計的有意差

太文字・斜体・下線：有意水準5%で有意差あり

太文字：有意水準10%で有意差あり

なお、「区分「3」の平均工数密度よりも有意に低い」とされたのは、要因「信頼性」の工程「詳細設計」のみである(“↓”を付記)。

ウェア障害が人命や社会基盤にかかわる損失を招く、もしくは、財政上の大規模な損失を招くと想定されたプロジェクトにおいては、詳細設計工程以外の工程において、工数密度が標準よりも高くなる事が分かる。特に、結合テスト工程と総合テスト工程では、工数密度が標準の2倍程度となり、生産性低下による工数超過のリスクが極めて高いと言える。一方、区分「4」や「5」に属するプロジェクト、すなわち、「ソフトウェア障害が簡単に復旧可能な小規模な損失しか招かない、もしくは、軽微な損失しか招かないと想定された」プロジェクトにおいては、基本5工程全体では標準との有意差は見られないに関わらず、詳細設計工程の工数密度だけが標準よりも低くなる事が分かる。別の見方をすれば、詳細設計工程で一旦低くなった工数密度は、その後の工程で標準的な値に戻るということであり、詳細設計工程における工数密度の低下を過大

評価し、その後も工数密度は高くならないと考えることは、工数超過のリスクをかえって高めることになる。

要因「開発スケジュール要求」が区分「1」や「2」に属するソフトウェア開発プロジェクト、すなわち、「開発スケジュールが当初の想定よりも10%以上短縮された」プロジェクトにおいては、PG設計製造工程と結合テスト工程において工数密度が高くなる事が分かる。一方、区分「4」や「5」に属するプロジェクト、すなわち、「開発スケジュールが当初の想定よりも10%以上許容された(10%以上のスケジュール超過が許容された)」プロジェクトにおける工数密度は、工程によって多少の高低はあるものの、標準との有意な差は見られない。工数密度、あるいは、生産性の観点からすると、「開発スケジュールが当初の想定通りだった」プロジェクトと違いはないことになる。

4-3 「設計」、「製造」、「テスト」

4-2では、生産性変動要因「信頼性」の生産性区分「4 & 5」のように、基本5工程全体では工数密度に標準との有意差はないが、特定の工程では有意差がある例を示した。このように特定の工程に限定すれば、パターン「H-M-M」、さらには、生産性変動要因と工数密度の間で想定される関係により合致するパターン「H-M-L」を示す要因が見つかるかもしれない。そこで、ここでは、工程別に生産性変動要因と工数密度の関係を見ていくこととする。ただし、基本5工程を、「設計：基本設計+詳細設計」、「製造：PG設計製造」、「テスト：結合テスト+総合テスト」の3つに縮退した上で、工数密度との関係を見ることとする。

10個の生産性変動要因それぞれについて、「設計」、「製造」、「テスト」の別で平均工数密度を算出した結果を図表8に示す。これまでと同様に、同図で示すのは各要因の区分「3」、つまり、生産性に与える影響は平均的、標準的であったプロジェクトにおける平均工数密度を1とした平均工数密度の相対値であり、ダネット検定(多重比較)の結果、有意水準5%で標準との有意差が認められた場合は「太文字・斜体・下線」で、有意水準10%で標準との有意差が認められた場合は「太文字」で、それぞれ示す。なお、全体を通して、工数密度が標準よりも有意に低いと認められた要因、工程はなかった。

まず、基本5工程全体ではパターン「M-M-M」、すなわち、いずれに区分においても工数密度に標準との有意差が見られなかった6つの要因(図表6参照)について見てみる。要因「機能性」では、製造工程の区分「1 & 2」において、工数密度が標準よりも有意に高いと認められる。「機能性が過度に要求された」プロジェクトにおいては、製造工程における工数超過(生産性低下)に特に注意し、備える必要があることになる。要因「SE・プログラマの経験と能力」では、設計工程の区分「1 & 2」において、工数密度が標準よりも有意に高いと認められる。「SE・プログラマの経験や能力が十分ではない」プロジェクトでは、設計工程における工数超過(生産性低下)に特に注意し、備える必要があることになる。残る4つの要因「発注要件の明確度と安定度」、「ユー

図表8 各工程における平均工数密度
(生産性区分「3」での平均値を「1」とした相対値)

工程 生産性 変動要因	設計 (基本設計 + 詳細設計)	製造 (PG設計製造)	テスト (結合テスト + 総合テスト)	5工程 全体
機能性				
1 & 2	1.05	1.47	1.33	1.26
4 & 5	0.89	1.14	0.92	0.98
信頼性				
1 & 2	1.42	1.60	2.09	1.65
4 & 5	0.78	0.98	0.86	0.87
プラットフォーム				
1 & 2	1.10	1.16	1.14	1.13
4 & 5	1.32	1.38	1.45	1.37
開発スケジュール要求				
1 & 2	1.30	1.47	1.40	1.39
4 & 5	0.90	0.88	0.95	0.91
発注要件の明確度と安定度				
1 & 2	1.02	1.21	1.25	1.14
4 & 5	0.68	0.78	0.77	0.74
ユーザの参画割合				
1 & 2	1.11	1.30	1.38	1.25
4 & 5	1.12	1.09	1.05	1.09
先行モデルの流用と標準モデルの採用				
1 & 2	1.08	1.19	1.22	1.16
4 & 5	1.01	0.91	0.99	0.97
プロジェクト管理者の経験と能力				
1 & 2	1.20	1.22	1.02	1.16
4 & 5	1.33	1.04	1.23	1.20
アナリストの経験と能力				
1 & 2	1.16	1.26	1.14	1.19
4 & 5	1.39	1.21	1.55	1.37
SE・プログラマの経験と能力				
1 & 2	1.49	1.39	1.03	1.33
4 & 5	1.13	1.28	1.22	1.20

太文字・斜体・下線：有意水準5%で有意(帰無仮説棄却)

太文字：有意水準10%で有意(帰無仮説棄却)

なお、「区分「3」の平均工数密度よりも有意に低い」とされた要因、工程はなし。

ザの参画割合」、「先行モデルの流用と標準モデルの採用」、および、「プロジェクト管理者の経験と能力」では、工程別でも工数密度に有意差は見られない。

次に、図表6において、パターン「M-M-H」を示し、工数密度との間で想定される関係と合致しない2つの要因について見てみる。要因「プラットフォーム」では、設計工程の区分「4 & 5」においては、工数密度に標準との有意差は見られないが、製造工程とテスト工程では、やはり、有意差が認められる。特に、テスト工程においては、平均工数密度が標準の1.45倍と非常に高くなっている。要因「アナリストの経験と能力」も

同様の傾向を示しており、製造工程の区分「4 & 5」においては、工数密度に標準との有意差は見られないが、設計工程とテスト工程では、やはり、有意差が認められ、特にテスト工程においては、平均工数密度が標準の1.55倍にも達する。これらは、生産性変動要因と工数密度の間で想定される関係に強く反するものであり、該当するプロジェクトが他の生産性変動要因によって大きく影響を受けている可能性があるため、さらに詳しく分析する必要がある。

5 生産性変動要因が工数密度に与える影響

本稿では、ここまで、ソフトウェア開発における生産性を、生産性変動要因と工数密度の関係性という観点で見てきた。特に、分析対象であるソフトウェア開発プロジェクトを、生産性変動要因と開発工程で層別し、工数密度が標準よりも有意に高くなる、あるいは、低くなる条件（要因と工程の組み合わせ）を明らかにしてきた。工数密度が特定の工程で標準とは異なることになると予め分かっていたら、それに備えることで、生産性低下による工数超過といったリスクを回避することができる。さらに一步進めて、工数密度の具体的な値が予測できるのであれば、工数超過の回避に向けた具体的な対策の検討や準備が可能となり、リスク管理上、大変有益である。

試しに、10個の生産性変動要因を説明変数とし、テスト工程（結合テスト工程＋総合テスト工程）の工数密度を目的変数とした回帰分析を行ってみたが有意な結果は得られなかった。生産性変動要因は、対象とするソフトウェア開発プロジェクトの特性を表す「プロファイルデータ」である。プロジェクトデータは、多くの場合、プロジェクト開始時、もしくは、開始後の早い段階で得られることから、工数などを目的変数とする予測モデルにおいて説明変数となり得る。ただし、その値が、主観による5段階評価など粒度の粗いものだと、工数などを具体的な数値で予測するには限界がある。

粒度の粗いプロファイルデータを補う一つの方法は、開発開始後に開発作業やその成果物から得られるデータ（プロセス・プロダクトデータ）を併用することである。プロセス・プロダクトデータは、進行中のプ

ロジェクトから実際に収集されるデータであり、プロファイルデータとは別の観点でその実態を表していると考えられる。また、一般に、データとしての粒度は、プロファイルデータよりも細かく、目的変数と同種のデータ、例えば、ある工程の工数を予測するのであれば、それより以前に実施された工程の工数を利用することで、予測精度の向上が見込まれる。ただし、当然のことながら、プロジェクトがある程度進行してからでなければ得られないデータであり、プロジェクト開始直後の利用は難しい。4-3で示した3つの工程「設計」、「製造」、「テスト」で言えば、設計工程の工数密度の予測には適用できないが、製造工程やテスト工程の工数密度の予測であれば可能である。

そこで、ここでは、生産性変動要因に加えて、設計工程と製造工程の工数密度を併用し、テスト工程の工数密度を予測することとする。設計工程の工数密度だけを用いて、製造工程やテスト工程、もしくはそれらの2つを合わせた工程の工数密度を予測することも考えられるが、まずは、より多くのプロセス・プロダクトデータを利用することとした。また、設計、製造、テストという大きなくくりにおいては、ソフトウェアを作成する側面が強い「設計と製造」と、作成されたソフトウェアの品質を確認し、確保するという側面が強い「テスト」とに大別するのが妥当であると考えた。

分析対象の全141プロジェクトを対象に、設計工程の工数密度（設計工数密度）と製造工程の工数密度（製造工数密度）を説明変数 x_1 、 x_2 とし、テスト工程の工数密度（テスト工数密度）を目的変数 y として回帰分析を行った結果、

$$y = 0.46 x_1 + 0.27 x_2 + 0.037$$

y : テスト工数密度

x_1 : 設計工数密度、 x_2 : 製造工数密度

決定係数（寄与率） $R^2 = 0.66$

が得られた。

決定係数 R^2 の値は決して低くないが、改良の余地はある。生産性変動要因を説明変数に加えることも考えられるが、先に述べたとおり、テスト工数密度を目的変数とした場合、説明変数として加えても結果が

大きく変わるとは考えにくい。そこで、説明変数ではなく、「テスト工数密度が高くなることが予想される」プロジェクトを選別（層別）する基準として生産性変動要因を用いることとし、回帰式（予測モデル）の改良が図られるかを確かめることとする。テスト工数密度がより高くなるプロジェクトにおいてより良い回帰式（予測モデル）が得られることは、リスク管理の観点から有益である。

4-3（図表8）において、区分「1 & 2」でテスト工数密度が標準よりも有意に高いと認められた生産性変動要因は、「信頼性」のみである。ここでは、もう少し可能性を広げるため、有意水準20%まで基準を緩め、区分「1 & 2」でテスト工数密度が標準よりも有意に高いと認められる要因として、「信頼性」、「開発スケジュール要求」、及び、「ユーザの参画割合」を追加する。計4つの要因それぞれについて、区分「1 & 2」に属するプロジェクトのみを対象とした。つまり、それら要因と区分で選別（層別）した上での回帰分析の結果を図表9にまとめる。比較のため、全141プロジェクトから得られた結果を再掲する。なお、選別（層別）によって分析対象となるプロジェクト件数が少なくなり、決定係数が不当に高くなる可能性があるため、自由度調整済決定係数を併記する。

同図表より、生産性変動要因「信頼性」、「開発スケジュール要求」、もしくは、「ユーザの参画割合」で対象プロジェクトを選別することにより、決定係数が0.8前後まで向上することがわかる。得られた3つの回帰式に共通するのは、製造工数密度 x_2 の係数が、もとの

0.27から0.43～0.50と大きくなっている点である。選別により製造工程での工数密度の影響がより明確に予測に反映され、回帰式（予測モデル）の改良につながったと見ることができる。

最後に、説明変数である設計工数密度と製造工数密度と目的変数であるテスト工数密度の関係を示す散布図上で、生産性変動要因による選別（層別）の効果をもう少し詳しく見てみることにする。ただし、それら3つの値に回帰式を重ねた散布図をそのまま図示するのは難しい。そこで、ここでは便宜的に、設計工数密度と製造工数密度の合計した値（設計製造工数密度）のみを説明変数として、テスト工数密度を目的変数とする回帰式を改めて求め、図示することとする。対象は、全141プロジェクト、および、図表9において決定係数の向上が見られた3つの生産性変動要因それぞれにより選別したプロジェクトの計4通りである。

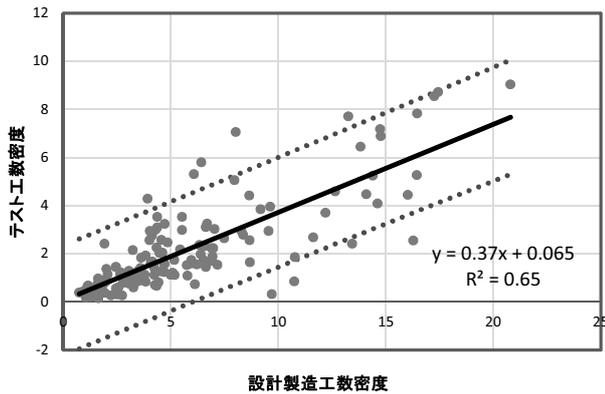
図表10に、設計製造工数密度とテスト工数密度の関係を示す散布図を示す。同図には、設計製造工数密度を説明変数 x とし、テスト工数密度を目的変数 y として得られた回帰式を示す。合わせて x から y の予測において、 y の予測値が95%の確率で存在することになる「95%予測区間」を示す。実線が回帰式、破線が予測区間の上限と下限である。また、回帰式には、決定係数 R^2 の値を併記している。図表9に示した値と比較すると、便宜的に説明変数を1つに減らしたが、決定係数の値にほとんど変化はないことが確認できる。

全141プロジェクトから得られた結果と、3つの生産性変動要因それぞれでプロジェクトを選別して得ら

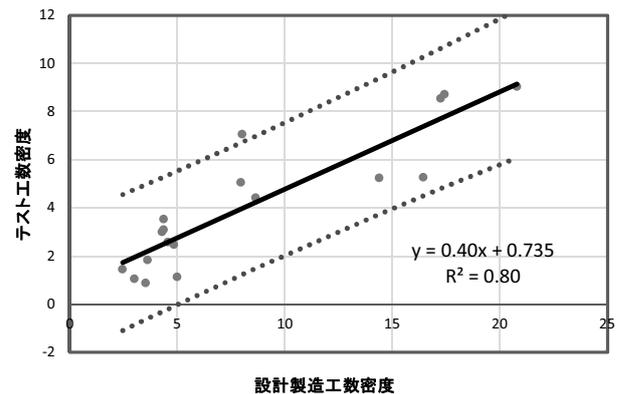
図表9 設計工数密度と製造工数密度を説明変数、テスト工数密度を目的変数とする回帰分析の結果

生産性変動要因	区分	プロジェクト件数	回帰式 y: テスト工数密度 x ₁ : 設計工数密度 x ₂ : 製造工数密度	決定係数 R^2
				自由度調整済決定係数 R^2
すべて	すべて	141	$y = 0.46 x_1 + 0.27 x_2 + 0.037$	0.66 0.65
機能性	1 & 2	48	$y = 0.54 x_1 + 0.24 x_2 + 0.061$	0.59 0.57
信頼性	1 & 2	18	$y = 0.37 x_1 + 0.43 x_2 + 0.767$	0.80 0.77
開発スケジュール要求	1 & 2	21	$y = 0.29 x_1 + 0.50 x_2 - 0.148$	0.83 0.81
ユーザの参画割合	1 & 2	21	$y = 0.40 x_1 + 0.48 x_2 - 0.074$	0.88 0.86

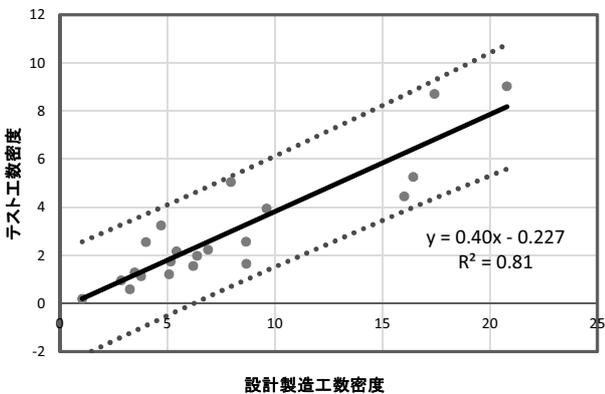
図表10 「設計製造工数密度」と「テスト工数密度」の関係



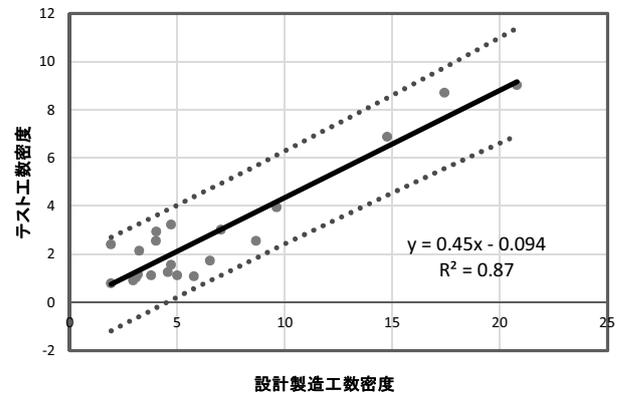
(a) 全141プロジェクト



(b) 要因「信頼性」の区分「1&2」に属するプロジェクト



(c) 要因「開発スケジュール要求」の区分「1&2」に属するプロジェクト



(d) 要因「ユーザの参画割合」の区分「1&2」に属するプロジェクト

れた結果を比較してみると、まず、決定係数に関しては、先述の通り、要因による選別により向上が見られる。一方、予測区間に関しては、選別に用いた要因により結果が異なる。すなわち、要因「ユーザの参画割合」による選別では、予測区間の幅が、全プロジェクトを対象とした場合よりも若干小さくなっているが、要因「開発スケジュール要求」による選別では、大きな違いはなく、要因「信頼性」による選別では、わずかではあるがかって大きくなっている。これは選別によって、回帰分析の対象プロジェクト数が約7分の1から8分の1にまで大きく減少したためである。

したがって、現状では生産性変動要因によるプロジェクトの選別(層別)は、テスト工数密度の予測精度の向上にはあまり寄与していないと言わざるを得ない。ただし、別の見方をすれば、選別により、予測精度を保ったまま、ごく少数のプロジェクトで構成されるグループを複数形成することに成功したとも言える。グループに属するプロジェクトが少ないことから、プロジェ

クトデータの増加による精度向上の可能性は大きく、また、グループごとにより適した(説明変数やその係数が異なる)予測モデルが構築されることにもつながる。

6 まとめ

本稿では、経済調査会が平成19年度から26年度に実施した「ソフトウェア開発に関する調査」で収集されたプロジェクトデータ(ソフトウェア開発データリポジトリ)を分析することで、ソフトウェア開発における生産性変動要因(委託者側要因6個、受託者側要因4個)が、基本5工程(基本設計、詳細設計、PG設計製造、結合テスト、総合テスト)における工数密度に与える影響を調べ、工数密度が標準よりも有意に高くなる、あるいは、低くなる条件(要因と工程の組み合わせ)を明らかにした。さらに、生産性低下による工数超過といった開発リスクの回避に向けた具体的な対策の検討や準備を可能とすべく、生産性変動要因と設計工数密

度と製造工数密度からテスト工数密度を予測することを試みた。得られた主な知見は次のとおりである。

- (1) 工程別工数密度の平均値は、PG設計製造工程が他の工程の2倍程度と突出しているが、変動係数が大きいのは、テスト工程（結合テスト工程＋総合テスト工程）である。
- (2) 「ソフトウェア障害が人命や社会基盤にかかわる損失を招く、もしくは、財政上の大規模な損失を招くと想定された」プロジェクト及び「開発スケジュールが当初の想定よりも10%以上短縮された」プロジェクトにおいて、5工程全体の工数密度が標準よりも高くなる。
- (3) 工程別にみると、「ソフトウェア障害が人命や社会基盤にかかわる損失を招く、もしくは、財政上の大規模な損失を招くと想定された」プロジェクトでは、詳細設計工程以外で工数密度が高くなり、特に結合テスト工程と総合テスト工程において、工数密度が標準の2倍程度と非常に高くなる。「開発スケジュールが当初の想定よりも10%以上短縮された」プロジェクトでは、PG設計製造工程と結合テスト工程において工数密度が高くなる。
- (4) 工数密度が標準よりも有意に低いと認められたのは、「ソフトウェア障害が簡単に復旧可能な小規模な損失しか招かない、もしくは、軽微な損失しか招かないと想定された」プロジェクトの詳細設計工程のみである。
- (5) (3)に加えて、設計工程（基本設計工程＋詳細設計工程）においては、「SE・プログラマの経験や能力が十分ではない」プロジェクトにおける生産性低下を、製造工程（PG設計製造工程）においては、「機能が過度に要求された」プロジェクトにおける生産性低下に特に注意し、工数超過に備える必要がある。
- (6) 生産性変動要因「信頼性」、「開発スケジュール要求」、もしくは、「ユーザの参画割合」で対象プロジェクトを選別（層別）することで、設計工数密度と製造工数密度からテスト工数密度を予測するモデルの決定係数（寄与率）を高めることができる。
- (7) 生産性変動要因「信頼性」、「開発スケジュール要

求」、もしくは、「ユーザの参画割合」で分析対象プロジェクトを選別（層別）しても、設計製造工数密度からテスト工数密度の予測精度（予測区間幅）は概ね保たれる。

分析の最後に試みた「テスト工数密度の予測」では、必ずしも予測精度の高いモデルが得られた訳ではないが、工数密度との関係性が確認された生産性変動要因であれば、それによって分析対象プロジェクトを選別（層別）し、その数を大幅に減らしても、モデルの予測精度（予測区間幅）は概ね保たれることが分かった。選別で減少した分析対象プロジェクトを、プロジェクトデータの更なる収集により増加させれば、選別を行わない場合よりも高い予測精度を実現することが出来る。このことは、プロジェクトデータの収集を継続し、ソフトウェア開発データリポジトリの整備を推進することへの強い動機付けとなる。

謝辞

工数密度の有意差検定においては、福岡工業大学情報科学科戸田航史先生より貴重なアドバイスを頂きました。ここに感謝の意を表します。

参考文献

- [1] 戸田航史、松本健一、大岩佐和子、押野智樹：“ソフトウェア開発における工程別生産性に関する分析”、経済調査会レビュー、Vol.8、財団法人経済調査会、2011年
- [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（IPA/SEC）：“ソフトウェア開発データ白書2010-2011”、2010年
- [3] 財団法人経済調査会：“ソフトウェア開発データリポジトリの分析”、2010年版、2010年

※経済調査会が、毎年実施する「ソフトウェア開発に関する調査」で収集したプロジェクトデータのデータベース。平成13年度から平成18年度までの6年度分で、延べ1,203プロジェクトの蓄積データの収集・分析結果を冊子「ソフトウェア開発データリポジトリの分析（2010年版）」（234ページ）として、平成22年7月に発行した。なお、改訂増補版として、平成13年度から平成24年度までの12年度分で、延べ1,853プロジェ

クトの蓄積データの収集・分析結果をとりまとめ、冊子「ソフトウェア開発データリポジトリの分析(2015年版)」(364ページ)として、平成27年5月に発行した。同冊子のPDF版は経済調査会のWebサイト <http://www.zai-keicho.or.jp/> の「研究成果」で確認できる。

付録 生産性変動要因の定義

生産性変動要因	生産性変動要因の説明	生産性区分
委託者側の要因	機能性	1 全体的に過度に要求された
		2 部分的に過度に要求された
		3 適度に要求された
		4 部分的に要求された
		5 要求されなかった
	信頼性	1 ソフトウェア障害は人命にかかわる損失が想定された
		2 ソフトウェア障害は財政上の大規模な損失が想定された
		3 ソフトウェア障害は復旧可能な中規模の損失が想定された
		4 ソフトウェア障害は簡単に復旧可能な小規模な損失が想定された
		5 ソフトウェア障害は軽微な損失が想定された
	プラットフォーム	1 全体的に適合性に欠けた
		2 部分的に適合性に欠けた
		3 適合性があった
		4 かなり適合性があった
		5 非常に適合性があった
開発スケジュール要求	1 スケジュールは当初の想定より25%以上短縮された	
	2 スケジュールは当初の想定より10%～25%前後短縮された	
	3 スケジュールは当初の想定どおりであった	
	4 スケジュールは当初の想定より10%～50%前後許容された	
	5 スケジュールは当初の想定より50%以上許容された	
発注要件の明確度と安定度	1 全体的に不明確・不安定であった	
	2 部分的に不明確・不安定であった	
	3 明確で安定していた	
	4 非常に明確で安定していた	
	5 先行モデルがあり非常に明確で常に安定していた	
ユーザの参画割合	1 全く参画しなかった	
	2 一部参画した	
	3 主要な工程について適度に参画した	
	4 全工程について適度に参画した	
	5 全工程について適切な担当者が適度に参画した	
受託者側の要因	先行モデルの流用と標準モデルの採用	1 先行モデルは存在しなかったまたは適合する標準モデルは存在しなかった
		2 先行モデルは存在したがほとんど流用出来なかったまたは適合する標準モデルは存在したが採用しなかった
		3 先行モデルは部分的に流用出来たまたは適合する標準モデルは部分的に採用出来た
		4 先行モデルはかなり流用出来たまたは適合する標準モデルはかなり採用出来た
		5 先行モデルは全面的に流用出来たまたは適合する標準モデルは全面的に採用出来た
	プロジェクト管理者の経験と能力	1 経験無かった
		2 少数の小規模プロジェクトの管理を経験していた
		3 多数の小規模プロジェクトの管理を経験していた
		4 少数の中大規模プロジェクトの管理を経験していた
		5 多数の中大規模プロジェクトの管理を経験していた
	アナリストの経験と能力	1 経験無かった
		2 少数の小規模プロジェクトのアナリストを経験していた
		3 多数の小規模プロジェクトのアナリストを経験していた
		4 少数の中大規模プロジェクトのアナリストを経験していた
		5 多数の中大規模プロジェクトのアナリストを経験していた
SE・プログラマの経験と能力	1 要員の過半数は経験無かった	
	2 半数はある程度経験していた	
	3 過半数はある程度経験していた	
	4 半数は十分な経験残り半数はある程度経験していた	
	5 過半数は十分な経験をしていた	