

ソフトウェア開発データが語るメッセージ 2015

～プロジェクトや組織のマネジメントの指針と
信頼性・生産性の傾向～

2015 年 9 月

独立行政法人情報処理推進機構（I P A）

技術本部 ソフトウェア高信頼化センター（S E C）



Information-technology
Promotion
Agency, Japan

目次

1. はじめに	1
2. エンジニアリング編	19
2.1. レビューのコントロール	19
2.2. テストのコントロール	36
2.3. 上流工程での不具合摘出比率のコントロール	53
2.4. 成果物量のコントロールと見積り妥当性評価	56
3. マネジメント編	62
3.1. プロジェクト・マネジメント関連	62
3.1.1. ユーザの関与（ステークホルダー・マネジメント関連）	62
3.1.2. 成果物スコープの明確化（スコープ・マネジメント関連）	66
3.2. 組織の成熟度向上に向けたマネジメント関連	68
3.2.1. 組織体制の整備	68
3.2.2. ドメイン別マネジメント	70
3.2.3. 生産性変動要因の分析と活用	84
3.2.4. 信頼性変動要因の分析と活用	92
4. トレンド編	97
4.1. 信頼性のトレンド	97
4.2. 生産性のトレンド	100

1. はじめに

IPA/SEC では、ソフトウェア開発における定量的管理の普及促進を目的に、開発プロセスの標準化や見える化手法、定量的品質管理手法などの調査・検討を行っている。これらの活動の一環として、国内の多様で幅広い業種・業務から収集した多数のソフトウェア開発に関するプロジェクトデータを整理・分析し、「ソフトウェア開発データ白書」として取りまとめて書籍化し、2005 年から定期的に発行している。その最新版として「ソフトウェア開発データ白書 2014-2015」¹（以下、白書）を 2014 年 10 月 1 日に発行した。

IPA/SEC が構築・運用しているソフトウェア開発プロジェクトデータベース（以下、IPA/SEC データベース）は、次の特長を有する世界的にも比類なき貴重なデータベースであり、種々の分析に活用できるものである。

◇データ件数が多い。

「ソフトウェア開発データ白書 2014-2015」作成時点で、3,541 件のプロジェクトデータを蓄積している。

◇偏りの少ない多様なプロジェクトデータから成っている。

データ提供企業は日本を代表するソフトウェア開発ベンダ（現在約 30 社）であり、業種、業務を始め多様なプロジェクトデータが蓄積されている。

◇生産性、信頼性等の要因分析が可能である。

プロジェクトデータとして、プロジェクトのプロファイルだけでなく、品質要求、開発プロセス、ユーザの協力度合い、組織の成熟度、開発環境等の広範なエリアの要因を分析するためのデータを備えている。

◇経年変化の分析が可能である。

データ収集開始してから 10 年以上が経過していることと、データ提供企業各社から毎年データ収集していることから、信頼性、生産性等の経年変化（推移）の分析が可能である。

この度、信頼性向上のために有用と考えられる新たな項目を中心に分析を行い、それらの分析結果と定量的管理への助言（以下、メッセージ）とを纏めた「ソフトウェア開発データが語るメッセージ 2015」（以下、本書）を作成した。本書は、次のようにエンジニアリング編、マネジメント編及びトレンド編の三編から成っており、それぞれ主に開発プロジェクトのプロジェクト・マネジャー及び担当者、開発組織の管理層の方々、業界や国等の政策立案担当者の方々に活用して頂くことを意図している。ソフトウェアの信頼性及び生産性の向上に向けた定量的管理推進のご参考になれば幸いである。

◇エンジニアリング編

主に開発プロジェクトのプロジェクト・マネジャー及び担当者に向けて、開発プロセスの定量的コントロールに活用して頂くことを目的として、レビュー、テスト等に関するいくつかの新規分析結果とメッセージを提示する。

◇マネジメント編

主に開発組織の管理層の方々に向けて、ステークホルダー・マネジメント、スコープ・マネジメント等のプロジェクト・マネジメントへの活用や、組織の成熟度向上（体制の整備、標準類の改良、信頼性／生産性向上に向けた重点強化領域の特定等）への活用に資することを目的として、いくつかの新規分析結果とメッセージを提示する。

◇トレンド編

主に IT 業界や国等の政策立案担当者の方々に向けて、施策検討の際の参考情報として利用して頂くことを期待して、信頼性及び生産性のトレンドに関する新規分析結果とメッセージを提示する。

以下に、本書に掲載しているメッセージの抜粋と本書の位置付けを示す。メッセージについて、詳しくはそれぞれの掲載箇所をご覧ください。

¹ 「ソフトウェア開発データ白書 2014-2015」の Web 公開サイト：<https://www.ipa.go.jp/sec/publish/tn12-002.html>

<メッセージ抜粋>

(1) エンジニアリング編

対象者と活用シーン：主に開発プロジェクトのプロジェクト・マネジャー及び担当者に向けて、開発プロセスの定量的コントロールに活用して頂くことを目的としたメッセージ。

① レビューのコントロール

目的	信頼性確保のための、設計レビュー工数のコントロール
分析項目	設計レビュー工数比率と発生不具合密度との関係
メッセージ	<p>◇分析結果概要（読み取れる傾向）</p> <p>設計レビュー工数比率が低くなるに連れて、発生不具合密度が 0.05 件以上/KSLOC と高い（相対的に信頼性が低い）ものが増えており、設計レビュー工数が少ないと信頼性を確保できないリスクが高まる傾向が伺える。特に、設計レビュー工数比率が 2%未満の領域においてその傾向が顕著である。</p> <p>一方、設計レビュー工数比率が 7%以上の場合、信頼性が低いものは僅かである。具体的には、新規開発の場合、発生不具合密度が 0.05 件以上/KSLOC のものは見られない（全数 0 件/KSLOC である）。また、改良開発の場合、発生不具合密度が 0.05 件以上/KSLOC のものは少数である。</p> <p>◇助言</p> <p>信頼性を確保するためのレビューのコントロールと実現性の高い開発計画策定に向けて、設計レビュー工数比率ができるだけ 7%以上となるよう、設計レビューを強化することをお勧めする。</p> <p>◇参考文献</p> <p>定量データに基づく品質予測・評価については、次の IPA/SEC の文献を併せて参考にして頂けると幸いである。</p> <ul style="list-style-type: none"> ○「定量的品質予測のススメ」 http://www.ipa.go.jp/sec/publish/tn08-004.html ○「続 定量的品質予測のススメ」 http://www.ipa.go.jp/sec/publish/tn10-004.html
メトリクス	<p>◇設計レビュー工数比率：基本設計から製作までのレビュー工数÷開発 5 工程の総工数。</p> <p>◇発生不具合密度：稼働後（最長 6 ヶ月以内）に発生した不具合数÷開発規模（KSLOC）。</p>
考察/備考	<p>設計レビューの工数をかけるほど検出できる不具合数が増えるのは当然と言えるが、一方では設計レビューの工数をかけるに連れて検出能率（設計レビュー工数当りの検出不具合数）が低下する傾向が見られる。従って、後出する「(3) 設計レビュー効果と能率を勘案した、設計レビュー工数のコントロール」では、経済的に信頼性を確保するために、検出能率の低下が飽和しない範囲で（目安としては設計レビュー工数密度の 75 パーセンタイル値の辺りまで）、設計レビューを強化することをお勧めしている。</p> <p>ここで、IPA/SEC データベースに基づく設計レビュー工数密度の 75 パーセンタイル値は、新規開発の場合、約 9.8 人時/KSLOC である。SLOC 生産性の中央値（約 5.6SLOC/人時）でこの値を設計レビュー工数比率に換算すると、約 5.5%</p>

	に相当する。
掲載箇所	2.1 節「レビューのコントロール」

目的	設計レビュー効果を上げるための、設計文書量のコントロール
分析項目	設計文書化密度と設計レビュー指摘密度との関係
メッセージ	<p>◇分析結果概要（読み取れる傾向）</p> <p>設計文書化密度が高くなるに連れて、設計レビュー指摘密度が高くなる傾向が見られる。設計文書量を増やすことによって、設計レビュー効果が向上する傾向が伺える。具体的には、設計文書化密度が相対的に低いもの（中央値約 15.8 ページ/KSLOC 以下の集合）と高いもの（中央値より大の集合）とでは、設計レビュー指摘密度の中央値に明らかな差が見られる。（新規開発では約 1.8 倍の差、改良開発では約 5.1 倍の差が見られる。）</p> <p>◇助言</p> <p>信頼性を確保するための設計レビューのコントロールと実現性の高い開発計画策定に向けて、設計文書量を増強することをお勧めする。目安としては、設計文書化密度が中央値以上となるように増強することをお勧めする。</p> <p>設計文書化密度の中央値は、新規開発の場合約 15.8 ページ/KSLOC、改良開発の場合約 44.7 ページ/KSLOC である。</p>
メトリクス	<p>◇設計文書化密度：（基本設計書のページ数＋詳細設計書のページ数）÷開発規模（KSLOC）。</p> <p>◇設計レビュー指摘密度：基本設計から製作までのレビュー指摘数÷開発規模（KSLOC）。</p>
考察/備考	<p>設計文書量が増えると、より具体的で漏れの少ない設計文書になるので、設計レビュー効果が向上する（設計レビューで見つかる不具合数が増える）と考えられる。</p> <p>設計レビュー効果を向上させるには、設計文書量を増強する以外にも、次のような設計関連標準類の改良/整備が有効と考えられる。</p> <ul style="list-style-type: none"> ・設計書作成標準の改良/整備 ・レビューチェックリストの改良/整備 ・解析ツールの改良/整備 ・レビュー手法等の改良/整備
掲載箇所	2.1 節「レビューのコントロール」

目的	設計レビュー効果と能率を勘案した、設計レビュー工数のコントロール
分析項目	設計レビュー工数密度と設計レビュー指摘密度との関係 設計レビュー工数密度と設計レビュー検出能率との関係
メッセージ	<p>◇分析結果概要（読み取れる傾向）</p> <p>設計レビュー工数密度が高くなるに連れて、設計レビュー指摘密度が高くなる傾向が見られる。具体的には、新規開発の場合、設計レビュー工数密度が 75 パーセント値（約 9.8 人時/KSLOC）より大きい領域では、設計レビュー工数密度が 25 パーセント値（約 1.5 人時/KSLOC）以下の領域と比較して、設計レビュー指摘密度の中央値が約 4.8 倍大きくなっている。</p> <p>設計レビュー工数をかけるほど、設計レビューで指摘できる件数（ひいては検出できる不具合数）を増やすことができると言える。</p>

	<p>ただし、設計レビュー工数をかけるに連れて、検出能率（設計レビュー工数当りの指摘件数）は低下する傾向が見られる。具体的には、新規開発の場合、設計レビュー工数密度が 75 パーセンタイル値（約 9.8 人時/KSLOC）より大きい領域では、設計レビュー工数密度が 25 パーセンタイル値（約 1.5 人時/KSLOC）以下の領域と比較して、設計レビュー検出能率の中央値が約 1/2.7 倍に低下している。</p> <p>◇助言 経済的に信頼性を確保するためのレビューのコントロールと実現性の高い開発計画策定に向けて、検出能率の低下が飽和しない範囲で（目安としては設計レビュー工数密度の 75 パーセンタイル値の辺りまで）、設計レビューを強化することをお勧めする。 設計レビュー工数密度の 75 パーセンタイル値は、新規開発の場合約 9.8 人時/KSLOC であり、改良開発の場合約 11.3 人時/KSLOC である。</p> <p>◇参考文献 レビューの充実に関連して、次の IPA/SEC の文献を併せて参考にして頂けると幸いである。 ○「高信頼化ソフトウェアのための開発手法ガイドブック」 http://www.ipa.go.jp/files/000005144.pdf 3.2.2 項「レビュー手法の概要」、5 章「トレーサビリティ管理の手法」</p>
メトリクス	<p>◇設計レビュー工数密度：基本設計から製作までの設計レビュー工数÷開発規模（KSLOC）。</p> <p>◇設計レビュー指摘密度：基本設計から製作までのレビュー指摘件数÷開発規模（KSLOC）。</p> <p>◇設計レビュー検出能率：設計レビュー指摘件数÷設計レビュー工数。</p>
考察/備考	<p>◇設計時の不具合を設計レビューによって 100% 摘出するのが困難な現状では、設計レビューの工数をかけるほど検出できる不具合数が増えるのは当然と言える。また、設計レビューの工数をかけるに連れて検出能率（設計レビュー工数当りの検出不具合数）が低下することも経験的に明らかである。（確率モデルで考えると、例えば地雷原の地雷が少なくなれば地雷を踏む確率が低下するのに似ている。）</p> <p>開発現場では、無制限に設計レビューを増強するのは困難であり、経済的に（効率的に）設計レビューを遂行することが求められるので、当該メッセージを提示している。</p> <p>なお、経済的に（効率的に）レビューを遂行するための方策としては、レビュー工数面だけでなく、次のような方策も併せて検討することが望ましい。</p> <ul style="list-style-type: none"> ・設計関連標準類の改良/整備 設計書作成標準、レビューチェックリスト、解析ツール、レビュー手法等 ・設計文書量の増強 設計文書化密度（設計書ページ数÷開発規模）を高めることによって、設計レビューで見つかる不具合数を増やせる。白書等の分析結果からは、目安として約 16 ページ/KSLOC 以上とすることが望ましい（新規開発の場合）。 <p>◇設計レビューの収束性評価に関しては、「設計レビュー検出能率がある一定のレベルまで低下しているか否か」も、収束性評価項目の一つとして利用可能と考えられる。</p>

掲載箇所	2.1 節「レビューのコントロール」
------	--------------------

② テストのコントロール

目的	テスト評価の一方法（ゾーン分析によるテスト評価）
分析項目	テスト密度対テスト検出不具合密度のゾーン別発生不具合密度
メッセージ	<p>◇分析結果概要（読み取れる傾向）</p> <p>相対的にテスト密度が低くテスト検出不具合密度が高い集合と、相対的にテスト密度が高くテスト検出不具合密度が低い集合とでは、後者の方が発生不具合密度が低い（相対的に信頼性が高い）傾向が見られる。</p> <p>具体的には、新規開発の場合、テスト密度が中央値（約 30.8 ケース/KSLOC）以下でテスト検出不具合密度が中央値（約 1.60 件/KSLOC）より大の領域の発生不具合密度の中央値が 0.022 件/KSLOC であるのに対して、テスト密度が中央値より大でテスト検出不具合密度が中央値以下の領域の発生不具合密度の中央値は 0 件/KSLOC である。また、発生不具合密度の 75 パーセンタイル値には約 3.3 倍の開きが見られる。</p> <p>◇助言</p> <p>テスト密度が高くテスト検出不具合密度が低いのは相対的に信頼性が良い兆候の一つである。一方、テスト密度が低くてテスト検出不具合密度が高いのは相対的に信頼性が良くない兆候の一つである。信頼性を確保するためのテストのコントロールに向けて、この見方をテストの評価項目の一つとして採用することをお勧めする。</p> <p>新規開発の場合、テスト密度の中央値は約 30.8 ケース/KSLOC、テスト検出不具合密度の中央値は約 1.60 件/KSLOC である。改良開発の場合、テスト密度の中央値は約 63.9 ケース/KSLOC、テスト検出不具合密度の中央値は約 1.50 件/KSLOC である。</p> <p>◇参考文献</p> <p>定量データに基づく品質予測・評価については、次の IPA/SEC の文献を併せて参考にして頂けると幸いである。</p> <ul style="list-style-type: none"> ○「定量的品質予測のススメ」 http://www.ipa.go.jp/sec/publish/tn08-004.html ○「続 定量的品質予測のススメ」 http://www.ipa.go.jp/sec/publish/tn10-004.html
メトリクス	<p>◇テスト密度：(結合テストケース数+総合テスト（ベンダ確認）ケース数) ÷ 開発規模（KSLOC）。</p> <p>◇テスト検出不具合密度：(結合テスト検出不具合数+総合テスト（ベンダ確認）検出不具合数) ÷ 開発規模（KSLOC）。</p> <p>◇発生不具合密度：稼働後（最長 6 ヶ月以内）に発生した不具合数 ÷ 開発規模（KSLOC）。</p>
考察/備考	<p>「テスト密度が高くテスト検出不具合密度が低いのは相対的に信頼性が良い（稼働後の不具合発生が少ない）兆候の一つである」の主な要因として、次のことが考えられる。</p> <p>◇テスト密度が高いのにテスト検出不具合密度が低いものは、テスト開始時点の</p>

	<p>出来が良い（潜在不具合の密度が低い）、つまりいわゆる作込み品質が良いものと考えられる。</p> <p>◇テスト検出不具合密度が高いものは発生不具合密度（稼働後の不具合密度）も高く、テスト検出不具合密度が低いものは発生不具合密度（稼働後の不具合密度）も低い傾向が見られる。つまり、テストに至るまでの良し悪しがテストによって逆転するケースは少ないと言える。</p> <p>信頼性を高めるには、やはり作込み品質向上を目指すことが王道であり、テストによって挽回しようという作戦の成算は薄いと考えられる。</p>
掲載箇所	2.2 節「テストのコントロール」

目的	テスト評価の一方法（その2）（テスト検出能率によるテストの収束性評価）
分析項目	テスト検出能率と発生不具合密度との関係
メッセージ	<p>◇分析結果概要（読み取れる傾向）</p> <p>テスト検出能率が低い方が相対的に発生不具合密度が低い（信頼性が高い）傾向が見られる。</p> <p>具体的には、新規開発の場合、テスト検出能率が中央値（約 0.048 件／テストケース）以下の集合と中央値より大きい集合とでは、発生不具合密度の中央値に約 2.4 倍の開きが見られる。特に、テスト検出能率が P25（約 0.026 件／テストケース）以下の集合では、発生不具合密度が低い（信頼性が高い）傾向が顕著である。</p> <p>改良開発の場合、テスト検出能率が中央値（約 0.025 件／テストケース）以下の集合では発生不具合密度の中央値が 0 件／KSLOC であるのに対して、中央値より大きい集合では発生不具合密度の中央値が約 0.022 件／KSLOC である。特に、テスト検出能率が P25（約 0.010 件／テストケース）以下の集合では、発生不具合密度が低い（信頼性が高い）傾向が顕著である。</p> <p>◇助言</p> <p>信頼性を確保するためのテストの監視・コントロールの一環として、次のような評価をテストの収束性評価項目の一つとすることをお勧めする。</p> <p>【「テスト検出能率がある一定のレベルまで低下しているか否か」を、テストの収束性評価項目の一つとして利用する。特に、追加テストの収束性評価に利用することをお勧めする】</p> <p>◇参考文献</p> <p>定量データに基づく品質予測・評価については、次の IPA/SEC の文献を併せて参考にして頂けると幸いである。</p> <ul style="list-style-type: none"> ○「定量的品質予測のススメ」 http://www.ipa.go.jp/sec/publish/tn08-004.html ○「続 定量的品質予測のススメ」 http://www.ipa.go.jp/sec/publish/tn10-004.html
メトリクス	<p>◇テスト検出能率：テスト検出不具合数÷テストケース数</p> <p>◇発生不具合密度：稼働後（最長 6 ヶ月以内）に発生した不具合数÷開発規模</p>
考察／備考	<p>テスト検出能率が低下して行くということは、まだ潜在している（残存している）不具合が減少して行くということを意味する。テスト検出能率によってテストの収束性を評価することは、理に適っていると考えられる。</p> <p>IPA/SEC データベースの分析結果からは、テスト検出能率が次 25 パーセン</p>

	<p>マイル値を目安として下回ることが望ましいと考える。</p> <ul style="list-style-type: none"> ・新規開発の場合、約 0.025 件/テストケース 40 ケース以上のテストに対して 1 件の不具合検出に相当。 ・改良開発の場合、約 0.010 件/テストケース 100 ケース以上のテストに対して 1 件の不具合検出に相当。
掲載箇所	2.2 節「テストのコントロール」

目的	テスト効果と能率を勘案した、テスト量のコントロール
分析項目	<p>テスト密度とテスト検出不具合密度との関係</p> <p>テスト密度とテスト検出能率との関係</p>
メッセージ	<p>◇分析結果概要（読み取れる傾向）</p> <p>テスト密度が高くなるに連れて、テスト検出不具合密度が高くなる傾向が見られる。具体的には、新規開発の場合、テスト密度が 75 パーセンタイル値（約 58.6 ケース/KSLOC）より大きい領域では、テスト密度が 25 パーセンタイル値（約 13.0 ケース/KSLOC）以下の領域と比較して、テスト検出不具合密度の中央値が約 2.9 倍大きくなっている。</p> <p>ただし、テスト密度が高くなるに連れて、テスト検出能率が大幅に低下する傾向が見られる。具体的には、新規開発の場合、テスト密度が 75 パーセンタイル値より大きい領域では、テスト密度が 25 パーセンタイル値以下の領域と比較して、テスト検出能率の中央値が約 1/6.0 倍に低下している。</p> <p>◇助言</p> <p>テストケースを増やすほど、テストでより多くの不具合を検出することができる。ただし、テストケースを増やすに連れて、テスト検出能率（テストケース当りの検出不具合数）は大幅に低下する。</p> <p>経済的に信頼性を確保するためのテストのコントロールと実現性の高い開発計画策定に向けて、検出能率の低下が飽和しない範囲で（目安としてはテスト密度の 75 パーセンタイル値の辺りまで）、テストを強化することをお勧めする。テスト密度の 75 パーセンタイル値は、新規開発の場合約 58.6 ケース/KSLOC であり、改良開発の場合約 147.9 ケース/KSLOC である。</p> <p>◇参考文献</p> <p>テストの充実に関連して、次の IPA/SEC の文献を併せて参考にして頂けると幸いです。</p> <p>○「高信頼化ソフトウェアのための開発手法ガイドブック」 http://www.ipa.go.jp/files/000005144.pdf 3.2.3 項「テスト手法の概要」、6 章「テスト網羅性の高度化技法」</p>
メトリクス	<p>◇テスト密度：(結合テストケース数+総合テスト（ベンダ確認）ケース数) ÷ 開発規模（KSLOC）。</p> <p>◇テスト検出不具合密度：(結合テスト検出不具合数+総合テスト（ベンダ確認）検出不具合数) ÷ 開発規模（KSLOC）。</p> <p>◇テスト検出能率：(結合テスト検出不具合数+総合テスト（ベンダ確認）検出不具合数) ÷ (結合テストケース数+総合テスト（ベンダ確認）ケース数)。</p>
考察/備考	<p>潜在している不具合をテストによって 100%完全に検出するのが難しい現状では、テストケース数を増やすほど検出できる不具合数が増えるのは当然と言えます。</p>

	<p>る。また、テストケース数を増やすに連れて検出能率（テストケース数当りの検出不具合数）が低下することも経験的に明らかである。（確率モデルで考えると、例えば地雷原の地雷が少なくなれば、地雷を踏む確率が低下するのに似ている。）</p> <p>開発現場では、無制限にテストを増強するのは困難であり、経済的に（効率的に）テストを遂行することが求められるので、テスト密度に関する当該メッセージを提示している。</p> <p>なお、経済的に（効率的に）テストを遂行するための方策としては、テスト密度面だけでなく、次のようなテスト関連標準類の改良／整備も併せて検討することが望ましい。</p> <ul style="list-style-type: none"> ・テスト設計手法 ・テスト設計のチェックリスト ・静的／動的解析ツール ・リグレッションテスト支援ツール （再現実行ツール、テスト結果の自動判定ツール等） ・テスト仕様書作成標準
掲載箇所	2.2 節「テストのコントロール」

③ 上流工程での不具合摘出比率のコントロール

目的	信頼性確保のための、上流工程での不具合摘出比率のコントロール
分析項目	上流工程での不具合摘出比率と発生不具合密度との関係
メッセージ	<p>◇分析結果概要（読み取れる傾向）</p> <p>新規開発の場合、上流工程での不具合摘出比率が相対的に低い（中央値約 73.7% 以下の）集合と高い（中央値より大きい）集合とでは、後者の方が発生不具合密度が低い（相対的に信頼性が高い）傾向が見られる。具体的には、新規開発の場合、前者の発生不具合密度の中央値が約 0.05 件/KSLOC であるのに対して、後者の発生不具合密度の中央値は 0 件/KSLOC である。また、発生不具合密度の 75 パーセンタイル値については、約 4.4 倍の開きが見られる。</p> <p>（なお、改良開発の場合には、このような顕著な傾向は見られなかった。）</p> <p>◇助言</p> <p>上流工程での不具合摘出比率を高めることによって、信頼性向上を図ることができる。上流工程での不具合摘出比率が向上するように、設計レビューを強化することをお勧めする。</p> <p>◇参考文献</p> <p>上流工程での不具合摘出比率向上に向けて、ソフトウェア開発に係るプロセスを改良する方策を検討する場合、次の IPA/SEC の文献を参考にして頂けると幸いである。</p> <p>○「プロセス改善ナビゲーションガイド ～虎の巻編～」 http://www.ipa.go.jp/sec/publish/tn08-009.html</p>
メトリクス	<p>◇上流工程での不具合摘出比率：基本設計から製作までのレビュー指摘数÷（基本設計から製作までのレビュー指摘数+結合テストから総合テスト（ベンダ確認）までの検出不具合数）。</p> <p>◇発生不具合密度：稼働後（最長 6 ヶ月以内）に発生した不具合数÷開発規模</p>

	(KSLOC)。
考察／備考	<p>上流工程での不具合摘出比率を向上させるような設計レビューの強化策としては、設計レビュー工数密度の向上等のレビュー工数面だけでなく、経済的に（効率的に）レビューを遂行するための次のような方策も併せて検討することが望ましいと考える。</p> <ul style="list-style-type: none"> 設計関連標準類の改良／整備 設計書作成標準、レビューチェックリスト、解析ツール、レビュー手法等 設計文書量の増強 設計文書化密度（設計書ページ数÷開発規模）を高めることによって、設計レビューで見つかる不具合数を増やせる。白書等の分析結果からは、目安として約 16 ページ／KSLOC 以上とすることが望ましい（新規開発の場合）。
掲載箇所	2.3 節「上流工程での不具合摘出比率のコントロール」

④ 成果物量のコントロールと見積り妥当性評価

目的	成果物量に着目して工程ごとに評価する方法
分析項目	開発規模と各工程の成果物量との相関 各工程の成果物量と各工程の工数との相関
メッセージ	<p>◇分析結果概要（読み取れる傾向） 新規開発の場合、開発規模（FP 規模又は SLOC 規模）と各工程の成果物量との間、及び各工程の成果物量と各工程の工数との間には、強い（又は中程度の）正相関が見られる。また、分析対象プロジェクトを特定の業種、例えば金融・保険業のプロジェクトに絞ると、より強い相関が見られる。</p> <p>◇助言 実現性の高い開発計画策定及び見積りの妥当性評価に向けて、次のような「成果物量に着目して工程ごとに評価する方法」を補完的に採用することをお勧めする。</p> <ul style="list-style-type: none"> ●開発規模当りの成果物量が妥当か（一定の範囲にあるか） ●成果物量当りの工数が妥当か（一定の範囲にあるか） <p>ここで、本書の次の基本統計量を目安として参考にして頂ければ幸いである。</p> <ul style="list-style-type: none"> 各工程の開発規模当りの成果物量（各工程の成果物量÷開発規模） 各工程の成果物量当りの工数（各工程の工数÷各工程の成果物量） <p>ただし、一定の範囲に収まっていないというだけで妥当でないとは評価するのは早計である。自組織の生産性変動要因による変動を勘案しながら評価することが望ましい。どのようなプロジェクトなのかによっては、変動要因による変動を始めとして一定の範囲に収まらなくなる合理的な理由が存在する可能性がある。評価対象プロジェクトに該当する変動要因によって生じる変動幅を勘案して、一定の妥当な範囲を上方修正／下方修正しながら妥当性評価することが望ましい。その結果においても妥当な範囲外となり、かつ変動要因以外の合理的理由がない場合には、見積りを見直すことが望ましい。</p>
考察／備考	◇工程別の成果物量に関する相関は、IPA/SEC のソフトウェア開発プロジェクトデータベースにおいて業種を始め様々なプロファイルのプロジェクトデータが混在した中でも見られる傾向であることから、各組織において同種のソフトウ

	<p>エア開発を行うドメインの中では、より強い相関が見られると考えられる。</p> <p>◇「成果物量に着目して工程ごとに評価する方法」のメリット</p> <p>成果物量には評価対象プロジェクトのスコープ、品質要求、制約条件等がある程度反映されるので、評価精度が高まると考えられる。(一方、開発規模から求めた総開発工数によって妥当性評価する方法では、これらを反映できない。)</p> <p>[例 1] 当プロジェクトでは品質要求が高いので、プロセス目標のうち、設計文書化密度(基本設計工程及び詳細設計工程の設計書ページ数÷開発規模)及びテスト密度(結合テスト工程及び総合テスト(ベンダ確認)工程のテストケース数÷開発規模)を高く設定する。</p> <p>[例 2] 当プロジェクトでは、関連システムとの接続テストを強化するので、総合テスト(ベンダ確認)工程のテスト密度(テストケース数÷開発規模)を高く設定する。</p> <p>また、工程ごとの成果物量及び工数を吟味することになるので、開発計画の実現可能性を高めることに繋がると考えられる。</p>
掲載箇所	2.4 節「成果物量のコントロールと見積り妥当性評価」

(2) マネジメント編

対象者と活用シーン：主に開発組織の管理層の方々に向けて、ステークホルダー・マネジメント、スコープ・マネジメント等のプロジェクト・マネジメントへの活用や、組織の成熟度向上(体制の整備、標準類の改良等)への活用に資することを目的としたメッセージ。

① プロジェクト・マネジメント関連

目的	ユーザ担当者の要求仕様関与度合いの向上 (ステークホルダー・マネジメントのユーザの関与関連)
分析項目	ユーザ担当者の要求仕様関与と SLOC 生産性との関係(新規開発) ユーザ担当者の要求仕様関与と発生不具合密度(信頼性)との関係(新規開発)
メッセージ	<p>◇分析結果概要(読み取れる傾向)</p> <p>ユーザ担当者の要求仕様関与が多い方が、少ない方と比較して相対的に SLOC 生産性が高く、発生不具合密度がやや低い(相対的に信頼性がやや高い)傾向が見られる。具体的には、両者の SLOC 生産性の中央値には、約 1.7 倍の開きが見られる。また、両者の発生不具合密度の中央値には、約 7.9 倍の開きが見られる(両者の発生不具合密度の 75 パーセンタイル値には、約 2.4 倍の開きが見られる)。</p> <p>◇助言</p> <p>ユーザ担当者の要求仕様関与が多い方が、少ない方と比較して相対的に生産性及び信頼性が良い傾向が見られる。生産性向上及び信頼性向上に向けて、ユーザ担当者の要求仕様関与度合いを高めるようユーザと調整することが望ましい。あるいは、ユーザ担当者の要求仕様関与度合いによる変動を目安として参考にしながら、開発計画を策定することが望ましい。</p> <p>◇参考文献</p> <p>ステークホルダー・マネジメントに関連して、次の IPA/SEC の文献を併せて参考にして頂けると幸いです。</p>

	<p>○「経営者が参画する要求品質の確保」 http://www.ipa.go.jp/sec/publish/tn05-002.html 要件定義までの超上流工程でのステークホルダーが、共通の認識を持ち、それぞれが果たす役割と責任を明確にして問題解決にあたるためのヒント</p> <p>○「機能要件の合意形成ガイド」 https://www.ipa.go.jp/sec/softwareengineering/reports/20100331.html 開発者が設計書を記述することのみではなく、発注者と開発者がシステム像を共有し、行き違いなく合意形成を行うための「コツ」</p>
メトリクス	<p>◇ユーザ担当者の要求仕様関与が多いとは「十分に関与又は概ね関与」を意味する。ユーザ担当者の要求仕様関与が少ないとは「関与が不十分又は未関与」を意味する。</p> <p>◇発生不具合密度：稼働後（最長 6 ヶ月以内）に発生した不具合数÷開発規模（KSLOC）。</p>
考察/備考	<p>プロジェクト・マネジメントのステークホルダー・マネジメントにおいては、ステークホルダー（特にユーザ）を効果的に関与させることが、プロジェクトを成功に導く重要な要素となっている。また、スコープ・マネジメントにおいては、成果物スコープ（成果物とその要件）のベースラインを確立しステークホルダー間で合意形成するプロセス（要件定義に相当）が重要な要素となっている。</p> <p>これらのことから、ユーザの関与の中で「ユーザ担当者の要求仕様への関与」が、プロジェクトを成功に導くために特に重要な要素と考えられるので、「ユーザ担当者の要求仕様への関与度合いと生産性及び信頼性との関係」を分析した。</p> <p>ユーザ担当者の要求仕様への関与度合いが高いと、要求仕様が早期に明確になる、要求仕様の変更/紆余曲折が少なくなり手戻りが少なくなる等の効果が期待できるので、生産性、信頼性両面で有利と考えられる。</p> <p>文書化、レビュー、テスト等の品質保証プロセスの工夫・努力によって信頼性向上を図ることは主に開発者自身が工夫・努力すれば良いことであるが、本件のようにユーザの理解と協力を得ることによってより一層の信頼性向上を図ることも重要と考える。</p>
掲載箇所	3.1.1 項「ユーザの関与（ステークホルダー・マネジメント関連）」

目的	要求仕様の明確化 (スコープ・マネジメントの成果物スコープ明確化関連)
分析項目	要求仕様の明確さと発生不具合密度（信頼性）との関係（新規開発）
メッセージ	<p>◇分析結果概要（読み取れる傾向） 要求仕様が明確な方が、あいまいな方と比較して発生不具合密度がやや低い（相対的に信頼性がやや高い）傾向が見られる。 具体的には、両者の発生不具合密度の中央値には、約 9.3 倍の開きが見られる（両者の発生不具合密度の 75 パーセンタイル値には、約 1.7 倍の開きが見られる）。</p> <p>◇助言 実現性の高い開発計画策定に向けて、要求仕様が早期に明確化するようユーザと調整することが望ましい。あるいは、要求仕様の明確さによる変動を目安として参考にしながら、開発計画を策定することが望ましい。</p> <p>◇参考文献 スコープ・マネジメントに関連して、次の IPA/SEC の文献を併せて参考にして</p>

	<p>頂けると幸いです。</p> <p>○「非機能要求グレード利用ガイド（2010年）」 http://www.ipa.go.jp/sec/softwareengineering/reports/20100416.html</p> <p>「非機能要求」について、発注者と受注者との認識の行き違いや異なる理解を防止することを目的とし、重要な項目から段階的に詳細化しながら非機能要求の確認を行うツール群。</p>
メトリクス	<p>◇要求仕様が明確とは「非常に明確又はかなり明確」を意味する。要求仕様があいまいとは「ややあいまい又は非常にあいまい」を意味する。</p> <p>◇発生不具合密度は、稼働後（最長6ヶ月以内）に発生した不具合数÷開発規模（KSLOC）。</p>
考察／備考	<p>要求仕様を早期に明確化すると、要求仕様の変更／紆余曲折が少なくなり手戻りが少なくなる等の効果が期待できるので、生産性、信頼性両面で有利と考えられる。</p> <p>スコープ・マネジメントにおいては、成果物スコープ（成果物とその要件）のベースラインを確立しステークホルダー間で合意形成するプロセス（要件定義に相当）が重要な要素となっている。また、ステークホルダー・マネジメントにおいては、ステークホルダー（特にユーザ）を効果的に関与させることが、プロジェクトを成功に導く重要な要素となっている。</p> <p>要求仕様の早期明確化に向けて、要件定義に注力するとともにユーザの要求仕様への関与度合いを高めることが重要と考えられる。</p>
掲載箇所	3.1.2 項「成果物スコープの明確化（スコープ・マネジメント関連）」

② 組織の成熟度向上に向けたマネジメント関連

目的	品質保証体制の強化 (組織体制の整備関連)
分析項目	品質保証体制別発生不具合密度（新規開発）
メッセージ	<p>◇分析結果概要（読み取れる傾向）</p> <p>品質保証体制として品質保証の専門スタッフが参加している集合の方が、参加していない集合と比較して発生不具合密度が低い（相対的に信頼性が高い）傾向が見られる。具体的には、両者の発生不具合密度の中央値には、約15.0倍の開きが見られる。両者の発生不具合密度の75パーセンタイル値には、約3.5倍の開きが見られる。</p> <p>◇助言</p> <p>信頼性向上に向けた組織体制整備の一環として、次のことをお勧めする。</p> <ul style="list-style-type: none"> ・より多くの開発プロジェクトに品質保証の専門スタッフが参加できるように、組織体制を強化する。 ・開発要員に対して、ソフトウェアの品質保証及びプロジェクト・マネジメントのスキル向上を図る。
メトリクス	◇発生不具合密度：稼働後（最長6ヶ月以内）に発生した不具合数÷開発規模（KSLOC）。
考察／備考	品質保証の専門スタッフがプロジェクトに参加することによって、専門スタッフのスキル／知識／経験が、主に次のような場面で、信頼性向上に向けて有効に

	<p>働くと考えられる。</p> <p>◇リスクの少ない（実現可能性の高い）開発計画の策定支援 種々の成功／失敗の事例を知っていることと、プロジェクト・マネジメントのスキルが高いことから、種々のプロジェクト・リスクを予見するとともにリスク対策について助言できる。</p> <p>◇レビュー及びテストに向けた効果的な助言 種々の不具合事例に基づいた弱点（陥りやすい誤り、漏れやすい箇所）に関する知識が豊富なので、レビュー及びテストに対して効果的に助言できる。</p> <p>◇精度の高い信頼性予測・評価 信頼性予測・評価スキルが高いので、開発時データ（文書化、レビュー、レビュー指摘、テスト、テスト時障害・誤り等のデータ）から、中間成果物及び成果物の信頼性／収束性を精度高く予測・評価できる。</p> <p>◇定量的管理の推進 定量的管理のスキルと豊富な経験から、プロジェクトの定量的管理を指導／支援できる。特に値を読み解く力があるので、定量データの分析結果をより適切なマネジメント・アクションに繋げることができる。</p>
掲載箇所	3.2.1 項「組織体制の整備」

目的	業種ドメイン別マネジメント
分析項目	発生不具合密度（信頼性）の業種間比較 KSLOC 当り工数（生産性）の業種間比較 生産性／信頼性の変動要因候補の業種間比較
メッセージ	<p>◇分析結果概要（読み取れる傾向） 発生不具合密度（信頼性）及び KSLOC 当り工数（生産性）は業種間で差異が見られる。また、信頼性要求レベルや開発プロセス関連の生産性／信頼性の変動要因候補においても業種間の差異が見られる。 特に金融・保険業では、他の業種より信頼性が高く生産性が低い傾向が見られる。また、その要因面では、他の業種より品質保証（設計文書化、レビュー、テスト）に多くの工数をかけている傾向が見られる。</p> <p>◇助言 マネジメントや分析を進める上で業種の影響を無視できないと判断できるので、業種によってドメインを分けてマネジメントすることをお勧めする。</p>
メトリクス	<p>◇発生不具合密度：稼働後（最長 6 ヶ月以内）に発生した不具合数÷開発規模（KSLOC）。</p> <p>◇開発プロセス関連の生産性／信頼性の変動要因候補として分析したもの： 設計文書化密度、設計レビュー工数密度、設計レビュー指摘密度、テスト密度、テスト検出不具合密度、上流工程での不具合摘出比率</p>
考察／備考	<p>◇特に金融・保険業では、他の業種より信頼性が高く生産性が低い傾向が見られる。また、その要因面では、他の業種より品質保証（設計文書化、レビュー、テスト）に多くの工数をかけている傾向が見られることから、システムリスクが高い（信頼性、公共性及び社会性の要求が高い）ソフトウェアの開発には、それ相応の品質保証（設計文書化、レビュー、テスト等）の工数が必要と考えられる。</p> <p>◇信頼性及び生産性は業種間で差異が見られるとともに、生産性／信頼性の変動</p>

	要因候補においても業種間の差異が見られることから、変動要因候補の多くが、生産性／信頼性の変動要因になっている可能性が高いと考えられる。
掲載箇所	3.2.2 項「ドメイン別マネジメント」

分析項目	生産性 (SLOC 規模当たり実績工数) の変動要因分析 (開発プロセス関連)
メッセージ	<p>◇分析結果概要 (読み取れる傾向)</p> <p>開発プロセス関連の生産性変動要因候補 (メトリクス欄参照) によって、生産性 (SLOC 規模当たり実績工数 (開発 5 工程)) が変動する傾向が見られる。</p> <p>◇助言</p> <ul style="list-style-type: none"> 生産性向上のための組織の重点強化領域の特定に向けて 生産性向上を進めて行くためには、個々のプロジェクトのマネジメントよりも、組織の改善に向けたマネジメント・サイクルを回すことが一層重要であろう。重点的に強化すると効果的な領域を特定し、適切な方策を立てることが望まれる。そのために、組織の生産性変動要因群を把握することが効果的であり、お勧めする。 開発計画の実現可能性検討、工数見積りの妥当性評価等に向けて 生産性の目標値とそれを達成するための開発プロセスの目標値あるいは見積り工数が、一定の妥当な範囲 (例えば管理指標の 25 パーセンタイル値～P75 パーセンタイル値の範囲) に収まっているか否かで評価するのが基本的な妥当性評価方法ではあるが、その範囲に収まっていないというだけで NG (妥当でない) と評価するのは早計である。どのようなプロジェクトなのかによっては、生産性変動要因による変動を始めとして一定の範囲に収まらなくなる合理的な理由が存在する可能性がある。評価対象プロジェクトに該当する生産性変動要因によって生じる変動幅を勘案して、一定の妥当な範囲を上方修正／下方修正しながら妥当性評価することをお勧めする。その結果においても妥当な範囲外となり、かつ生産性変動要因以外の合理的な理由がない場合には、計画や見積りを見直すことが望ましい。
メトリクス	<p>開発プロセス関連の生産性変動要因候補として分析したものは次の通り。</p> <p>◇設計文書化密度：(基本設計書のページ数+詳細設計書のページ数) ÷ 開発規模 (KSLOC)</p> <p>◇設計レビュー工数密度：基本設計から製作までの設計レビュー工数 ÷ 開発規模 (KSLOC)</p> <p>◇設計レビュー指摘密度：基本設計から製作までのレビュー指摘数 ÷ 開発規模 (KSLOC)</p> <p>◇テスト密度：(結合テストケース数+総合テスト (ベンダ確認) ケース数) ÷ 開発規模 (KSLOC)</p> <p>◇テスト検出不具合密度：(結合テスト検出不具合数+総合テスト (ベンダ確認) 検出不具合数) ÷ 開発規模 (KSLOC)</p> <p>◇上流工程での不具合摘出比率：基本設計から製作までのレビュー指摘数 ÷ (基本設計から製作までのレビュー指摘数+結合テストから総合テスト (ベンダ確認) までの検出不具合数)</p>
考察/備考	IPVSEC データベースにおいて業種を始め様々なプロファイルのプロジェクトデータが混在した中では、開発プロセス関連の生産性変動要因候補 (メトリクス欄参照) によって生産性が変動する傾向が見られる。各組織の同種のソフトウェア

	ア開発を行うドメインの中では、同様な傾向が見られるとは限らないが、より強い傾向を示す変動要因が見つかることと期待される。自組織の生産性変動要因群を分析・把握することが重要かつ効果的である。
掲載箇所	3.2.3 項「生産性変動要因の分析と活用」

分析項目	信頼性（発生不具合密度）の変動要因分析（開発プロセス関連）
メッセージ	<p>◇分析結果概要（読み取れる傾向）</p> <p>発生不具合密度が 0.02 件/KLSLOC 未満のプロジェクトを良群、0.02/KLSLOC 以上のプロジェクトを否群に分けて、両者に開発プロセス関連の信頼性変動要因候補（メトリクス欄参照）の値にどの程度の差異があるかを分析した。</p> <p>その結果、信頼性変動要因候補のいくつかによって、信頼性（発生不具合密度）が変動する傾向が見られる。</p> <p>具体的には、良群に次の傾向が見られる。</p> <ul style="list-style-type: none"> ・レビュー工数密度が高い ・上流工程での不具合摘出比率が高い ・テスト検出不具合密度が低い ・テスト検出能率（テスト項目当りの検出不具合数）が低い <p>◇助言</p> <ul style="list-style-type: none"> ・自組織の各ドメインにおいて自組織の定量データを用いて、信頼性変動要因を分析することをお勧めする。組織の信頼性向上方策の検討にあたっては、信頼性変動要因に着目して重点的に強化すると効果的な領域を特定し、その領域の信頼性向上方策を検討することをお勧めする。 ・IPA/SEC データベースに基づいた当分析結果からは、「レビュー工数密度を高めると信頼性が向上する」、「上流工程での不具合摘出比率を高めると信頼性が向上する」等から、設計及びレビューのプロセス改善が効果的と考えられる。 （例）レビュー工数密度を高める。 レビュー効率を高める（設計文書化密度を上げる、設計書作成標準やチェックリストを改良する等）
メトリクス	<p>◇発生不具合密度：稼働後（最長 6 ヶ月以内）に発生した不具合数 ÷ 開発規模 (KSLOC)。</p> <p>開発プロセス関連の信頼性変動要因候補として分析したものは次の通り。</p> <p>◇設計レビュー工数密度：基本設計から製作までの設計レビュー工数 ÷ 開発規模 (KSLOC)</p> <p>◇設計レビュー指摘密度：基本設計から製作までのレビュー指摘数 ÷ 開発規模 (KSLOC)</p> <p>◇上流工程での不具合摘出比率：基本設計から製作までのレビュー指摘数 ÷ (基本設計から製作までのレビュー指摘数 + 結合テストから総合テスト（ベンダ確認）までの検出不具合数)</p> <p>◇テスト密度：(結合テストケース数 + 総合テスト（ベンダ確認）ケース数) ÷ 開発規模 (KSLOC)</p> <p>◇テスト検出不具合密度：(結合テスト検出不具合数 + 総合テスト（ベンダ確認）検出不具合数) ÷ 開発規模 (KSLOC)</p> <p>◇テスト検出能率：(結合テスト検出不具合数 + 総合テスト（ベンダ確認）検出不具合数) ÷ (結合テストケース数 + 総合テスト（ベンダ確認）ケース数)</p>

考察／備考	IPA/SEC データベースにおいて業種を始め様々なプロファイルのプロジェクトデータが混在した中では、開発プロセス関連の信頼性変動要因候補（メトリクス欄参照）のいくつかによって信頼性が変動する傾向が見られる。各組織の同種のソフトウェア開発を行うドメインの中では、同様な傾向が見られるとは限らないが、より強い傾向を示す変動要因が見つかることと期待される。自組織の信頼性変動要因群を分析・把握することが重要かつ効果的である。
掲載箇所	3.2.4 項「信頼性変動要因の分析と活用」

(3) トレンド編

対象者と活用シーン：主に IT 業界や国等の政策立案担当者の方々に向けて、施策検討の際の参考情報としての利用を期待するメッセージ。

① 信頼性のトレンド

目的	信頼性向上に向けた政策の検討 (マネジメント活動や諸施策の総合的な効果の確認、品質マネジメントの年度計画や中長期計画における信頼性目標の設定等を含む)
分析項目	発生不具合密度（信頼性）の経年変化（新規開発）
メッセージ	◇分析結果概要（読み取れる傾向） IPA/SEC のソフトウェア開発プロジェクトデータベースにおいては、プロジェクトデータを収集開始してから最近まで（プロジェクトの実績終了年が 2004 年から 2012 年まで）の発生不具合密度の中央値の経年変化に低下傾向（信頼性向上の傾向）が見られる。 具体的には、新規開発の場合、発生不具合密度の中央値が 0.02 件/KSLOC から 0 件/KSLOC に低下している。2011 年及び 2012 年の中央値は 0 件/KSLOC。ただし、発生不具合密度の 75 パーセンタイル値については、顕著な低下傾向は見られない。いまだに、ばらつきは小さくないと言える。 ◇助言 上記の IPA/SEC データベースに基づく発生不具合密度（信頼性）の経年変化の分析結果を、次のようなシーンで参考にして頂けると幸いである。 ・信頼性向上に向けたマネジメント活動や諸施策の総合的な効果の確認 ・品質マネジメントの年度計画や中長期計画における目標と施策の検討等
メトリクス	◇発生不具合密度：稼働後（最長 6 ヶ月以内）に発生した不具合数÷開発規模（KSLOC）。
考察／備考	信頼性変動要因に関連する IPA/SEC データベースの今までの分析結果から、次の方策も信頼性向上に寄与すると考えられる。ヒントとして参考にされたい。 ◇レビュー工数を増強する。 目安としては、設計レビュー工数密度（設計レビュー工数÷開発規模）が 10 人時/KSLOC 程度となるよう増強する。 ◇上流（設計及び製作）工程での不具合摘出比率（上流工程での不具合摘出数÷開発完了までの不具合摘出総数）を高める。 目安としては、上流工程での不具合摘出比率を 74%以上とする。 設計レビュー効果を上げるために、設計文書化密度（設計書ページ数÷開発規

	<p>模) を約 16 ページ/KSLOC 以上とする (新規開発の場合)。 ◇ユーザ担当者の要求仕様関与度合いを高めるよう、ユーザと調整する。 ◇要求仕様を早期に明確化するよう、ユーザと調整する。</p>
掲載箇所	4.1 節「信頼性のトレンド」

② 生産性のトレンド

目的	<p>生産性向上に向けた政策の検討 (マネジメント活動や諸施策の総合的な効果の確認、生産性マネジメントの年度計画や中長期計画における生産性目標の設定等を含む)</p>
分析項目	SLOC 生産性の経年変化
メッセージ	<p>◇分析結果概要 (読み取れる傾向) IPA/SEC のソフトウェア開発プロジェクトデータベースにおいては、プロジェクトデータを収集開始してから最近まで (プロジェクトの実績終了年が 2004 年から 2012 年まで) の SLOC 生産性の中央値の経年変化に、生産性向上 (または生産性低下) の傾向は見られない。 ソフトウェア開発プロジェクトは、低価格と短納期の強いプレッシャーに晒されることが多々ある。また、予算管理や価格交渉等の場面で、年率 5% の開発コスト削減や前年度比 10% のライン単価低減等が要求されるケースが散見される。しかし、データによる裏付け等の根拠が希薄であると、このような状況は開発プロジェクトのリスク増大や品質低下を招く恐れがある。</p> <p>◇助言 上記の IPA/SEC データベースに基づく SLOC 生産性の経年変化の分析結果を、次のようなシーンで参考にして頂けると幸いである。</p> <ul style="list-style-type: none"> ・生産性向上に向けたマネジメント活動や諸施策の総合的な効果の確認 ・生産性マネジメントの年度計画や中長期計画における目標と施策の検討 ・予算管理、単価の改定/交渉
メトリクス	◇SLOC 生産性 : $\text{SLOC 規模} \div \text{開発工数}$ (開発 5 工程)
考察/備考	<p>生産性に向上傾向が見られない要因の一つとして、次のことが考えられる。 「複雑さ・難しさが増している一方で、開発要員のスキルレベルが追いついていない。開発プロジェクトにとっての相対的な難易度は高まっていることが考えられる。」 少数精鋭部隊で開発していた時代は去っている。次のような状況下でも、如何にして要員のスキルレベルを確保/向上させるか、また効率的な開発方法を実現するかが課題の一つと考える。</p> <p>◇開発すべきソフトウェアの総量が増大するのに伴って、ますます大勢の開発要員が必要になっている (要員の資質の平均レベルが低下する)。 ◇大半の開発プロジェクトが改良開発 (改修・保守や拡張) になって来ており、新規開発 (特にスクラッチ開発) によってスキルアップするチャンスが少なくなっている。</p>
掲載箇所	4.2 節「生産性のトレンド」

<本書の位置付け>

- (1) 白書は客観的な統計情報等を提供しているのに対して、本書は、データに基づくメッセージを発信することを主眼にしたものである。メッセージには、IPA/SECの見識や観点が反映されていることにご留意頂きたい。
- (2) 主に信頼性向上のために有用と考えられる新たな分析項目を検討し、白書作成用のIPA/SECデータベースを用いてそれらを分析した。それらの分析結果のうち傾向が読み取れるものについて、メッセージを考察している。なお、メッセージの基となっている分析結果については、今後さまざまな方面からの議論を行った上で、次回以降の「ソフトウェア開発データ白書」に反映したいと考えている。
- (3) 本書で提示したメッセージを「ヒント」として参考にしながら、定量データに基づくソフトウェア品質／生産性マネジメントを一層進めて頂ければ幸いである。また、白書と併せて、本書の指標、分析方法及び分析結果を定量的管理の「目安」や「ヒント」として参考にしながら、自組織のソフトウェア開発プロジェクトデータを用いた定量的管理を一層進めて頂ければ幸いである。
- (4) 白書及び本書の統計値を参考にする場合には、次の点に留意して「目安」として参考にされたい。
 - ① 白書及び本書の分析結果は、基本的に種々の組織、業種、プロファイル等が混在したサンプルデータ集合から得られたものである。信頼性、生産性等の評価にあたっては、妥当な水準／範囲が業種、品質要求レベルを始めとする種々の変動要因によって変動することを勘案する必要がある。一般に、システムはその重要性や停止時の影響の大きさによりクラス分けされる。本書のメッセージに示される推奨値については、システムのクラスに応じて適宜補正した上で、「目安」として利用されたい。
 - ② 白書及び本書で用いているサンプルデータ集合には、定量的管理が行われている組織で定量的管理されたプロジェクトのデータが多く含まれていることから、白書及び本書の統計情報は定量的管理されたプロジェクトの統計情報と見るのが妥当である。従って、白書及び本書の生産性／信頼性に関する統計値は世間相場観よりも良い値を示している可能性がある。
 - ③ 今後、新しい開発プロジェクトのデータが追加されて行くことに伴って、統計値や傾向が変化する可能性がある。特にサンプルデータ数がまだ少ない分析項目については、統計値や傾向が変化する可能性が高い。
- (5) 白書及び本書の分析結果は、基本的に種々の組織、業種、プロファイル等が混在したサンプルデータ集合から得られたものである。白書及び本書で見られる傾向はこのようなサンプルデータ集合においても見られる傾向であることから、各組織の同種のドメインにおいては、より顕著な傾向が見られると期待される。

2. エンジニアリング編

主に開発プロジェクトのプロジェクト・マネジャー及び担当者に向けて、開発プロセスの定量的コントロールに活用して頂くことを目的として、レビュー、テスト等に関するいくつかの新規分析結果とメッセージを提示する。

2.1. レビューのコントロール

(1) 信頼性確保のための、設計レビュー工数のコントロール

【概要】

分析項目	設計レビュー工数比率と発生不具合密度との関係
メッセージ	<p>◇分析結果概要（読み取れる傾向）</p> <p>設計レビュー工数比率が低くなるに連れて、発生不具合密度が 0.05 件以上/KSLOC と高い（相対的に信頼性が低い）ものが増えており、設計レビュー工数が少ないと信頼性を確保できないリスクが高まる傾向が伺える。特に、設計レビュー工数比率が 2%未満の領域においてその傾向が顕著である。</p> <p>一方、設計レビュー工数比率が 7%以上の場合、信頼性が低いものは僅かである。具体的には、新規開発の場合、発生不具合密度が 0.05 件以上/KSLOC のものは見られない（全数 0 件/KSLOC である）。また、改良開発の場合、発生不具合密度が 0.05 件以上/KSLOC のものは少数である。</p> <p>◇助言</p> <p>信頼性を確保するためのレビューのコントロールと実現性の高い開発計画策定に向けて、設計レビュー工数比率ができるだけ 7%以上となるよう、設計レビューを強化することをお勧めする。</p> <p>◇参考文献</p> <p>定量データに基づく品質予測・評価については、次の IPA/SEC の文献を併せて参考にして頂けると幸いである。</p> <ul style="list-style-type: none"> ○「定量的品質予測のススメ」 http://www.ipa.go.jp/sec/publish/tn08-004.html ○「続 定量的品質予測のススメ」 http://www.ipa.go.jp/sec/publish/tn10-004.html
メトリクス	<p>◇設計レビュー工数比率：基本設計から製作までのレビュー工数÷開発 5 工程の総工数。</p> <p>◇発生不具合密度：稼働後（最長 6 ヶ月以内）に発生した不具合数÷開発規模 (KSLOC)。</p>
考察/備考	<p>設計レビューの工数をかけるほど検出できる不具合数が増えるのは当然と言えるが、一方では設計レビューの工数をかけるに連れて検出能率（設計レビュー工数当りの検出不具合数）が低下する傾向が見られる。従って、後出する「(3) 設計レビュー効果と能率を勘案した、設計レビュー工数のコントロール」では、経済的に信頼性を確保するために、検出能率の低下が飽和しない範囲で（目安としては設計レビュー工数密度の 75 パーセンタイル値の辺りまで）、設計レビューを強化することをお勧めしている。</p> <p>ここで、IPA/SEC データベースに基づく設計レビュー工数密度の 75 パーセンタ</p>

	イル値は、新規開発の場合、約 9.8 人時/KSLOC である。SLOC 生産性の中央値 (約 5.6SLOC/人時) でこの値を設計レビュー工数比率に換算すると、約 5.5% に相当する。
--	---

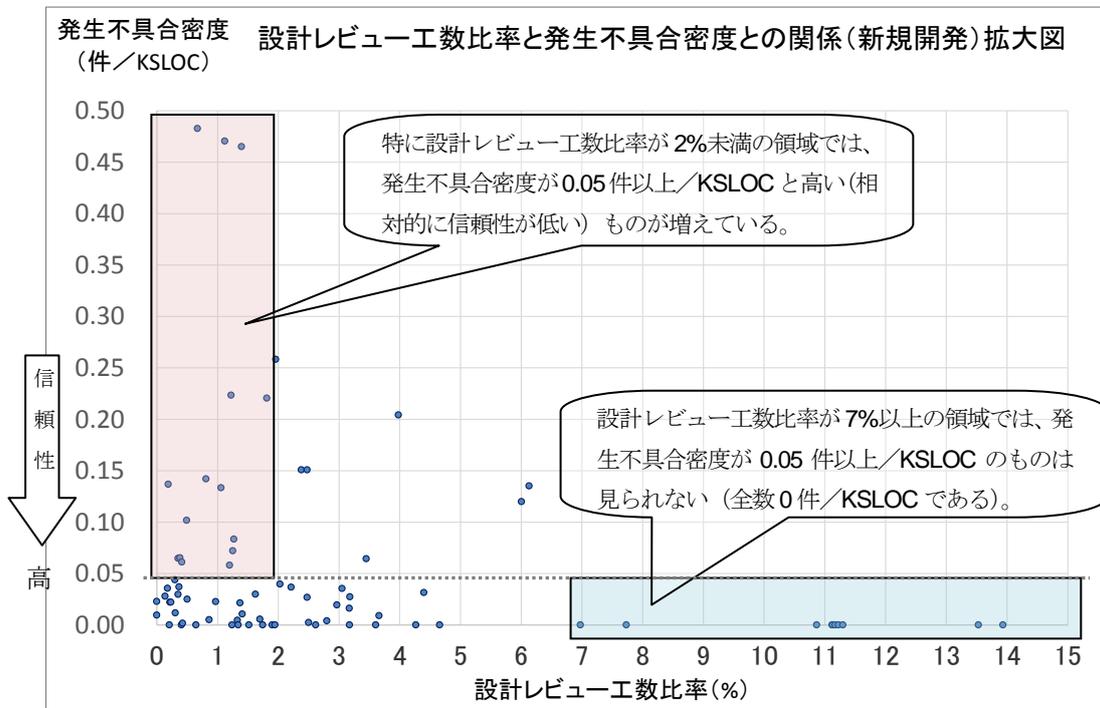
【検索条件】

- ◇ 開発 5 工程が揃っているプロジェクト
- ◇ 開発プロジェクトの種別が新規開発又は改良開発 (改修・保守又は拡張)
- ◇ 開発規模 (SLOC 規模) > 0
- ◇ 主開発言語が COBOL, C, VB, Java のいずれか
- ◇ 開発 5 工程の総工数 > 0
- ◇ 稼働後の発生不具合数 >= 0
- ◇ 基本設計から製作までのレビュー工数 > 0

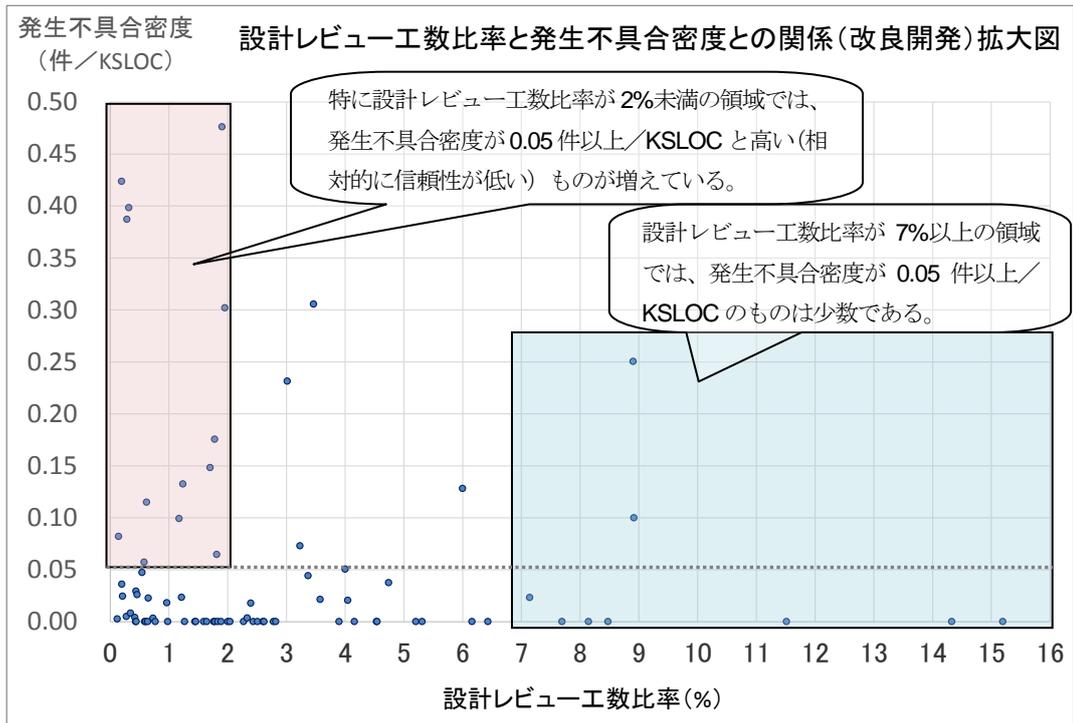
【分析結果】

設計レビュー工数比率が低くなるに連れて、発生不具合密度が 0.05 件以上/KSLOC と高い (信頼性が相対的に低い) ものが増えている。特に、設計レビュー工数比率が 2%未満の領域においてその傾向が顕著である。一方、設計レビュー工数比率が 7%以上の領域では、稼働後の発生不具合密度が 0.05 件以上/KSLOC のものは少数である。具体的には、新規開発の場合、発生不具合密度が 0.05 件以上/KSLOC のものは見られない (全数 0 件/KSLOC である)。また、改良開発の場合、発生不具合密度が 0.05 件以上/KSLOC のものは少数である。

<新規開発の場合> 散布図



<改良開発の場合> 散布図



(2) 設計レビュー効果を上げるための、設計文書量のコントロール

【概要】

分析項目	設計文書化密度と設計レビュー指摘密度との関係
メッセージ	<p>◇分析結果概要（読み取れる傾向）</p> <p>設計文書化密度が高くなるに連れて、設計レビュー指摘密度が高くなる傾向が見られる。設計文書量を増やすことによって、設計レビュー効果が向上する傾向が伺える。具体的には、設計文書化密度が相対的に低いもの（中央値約 15.8 ページ/KSLOC 以下の集合）と高いもの（中央値より大の集合）とでは、設計レビュー指摘密度の中央値に明らかな差が見られる。（新規開発では約 1.8 倍の差、改良開発では約 5.1 倍の差が見られる。）</p> <p>◇助言</p> <p>信頼性を確保するための設計レビューのコントロールと実現性の高い開発計画策定に向けて、設計文書量を増強することをお勧めする。目安としては、設計文書化密度が中央値以上となるように増強することをお勧めする。</p> <p>設計文書化密度の中央値は、新規開発の場合約 15.8 ページ/KSLOC、改良開発の場合約 44.7 ページ/KSLOC である。</p>
メトリクス	<p>◇設計文書化密度：（基本設計書のページ数＋詳細設計書のページ数）÷開発規模（KSLOC）。</p> <p>◇設計レビュー指摘密度：基本設計から製作までのレビュー指摘数÷開発規模（KSLOC）。</p>
考察／備考	<p>設計文書量が増えると、より具体的で漏れの少ない設計文書になるので、設計レビュー効果が向上する（設計レビューで見つかる不具合数が増える）と考えられる。</p> <p>設計レビュー効果を向上させるには、設計文書量を増強する以外にも、次のような設計関連標準類の改良／整備が有効と考えられる。</p> <ul style="list-style-type: none"> ・設計書作成標準の改良／整備 ・レビューチェックリストの改良／整備 ・解析ツールの改良／整備 ・レビュー手法等の改良／整備

【検索条件】

- ◇ 開発 5 工程が揃っているプロジェクト
- ◇ 開発プロジェクトの種別が新規開発又は改良開発（改修・保守又は拡張）
- ◇ 開発規模（SLOC 規模） > 0
- ◇ 主開発言語が COBOL, C, VB, Java のいずれか
- ◇ 開発 5 工程の総工数 > 0
- ◇ 稼働後の発生不具合数 >= 0
- ◇ 基本設計書のページ数 > 0 かつ 詳細設計書のページ数 > 0
- ◇ 基本設計から製作までのレビュー指摘数 > 0

【分析結果】

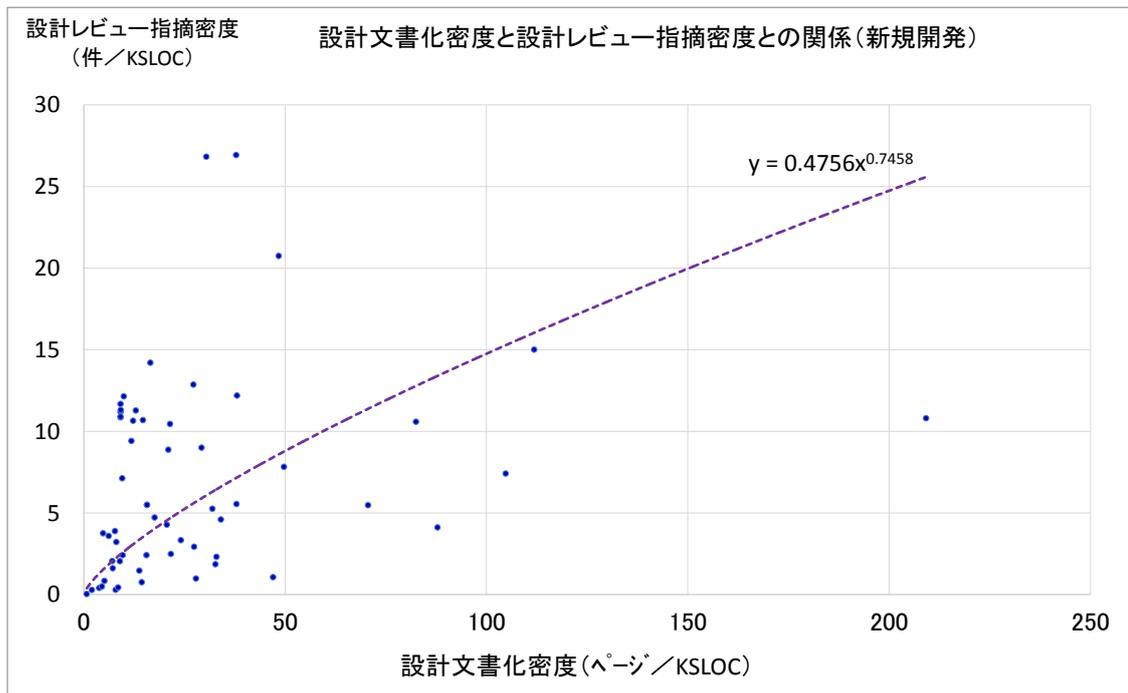
設計文書化密度が高くなるに連れて、設計レビュー指摘密度が高くなる傾向が見られる。
また、設計文書化密度が相対的に低いもの（中央値以下の集合）と高いもの（中央値より大の集合）とでは、

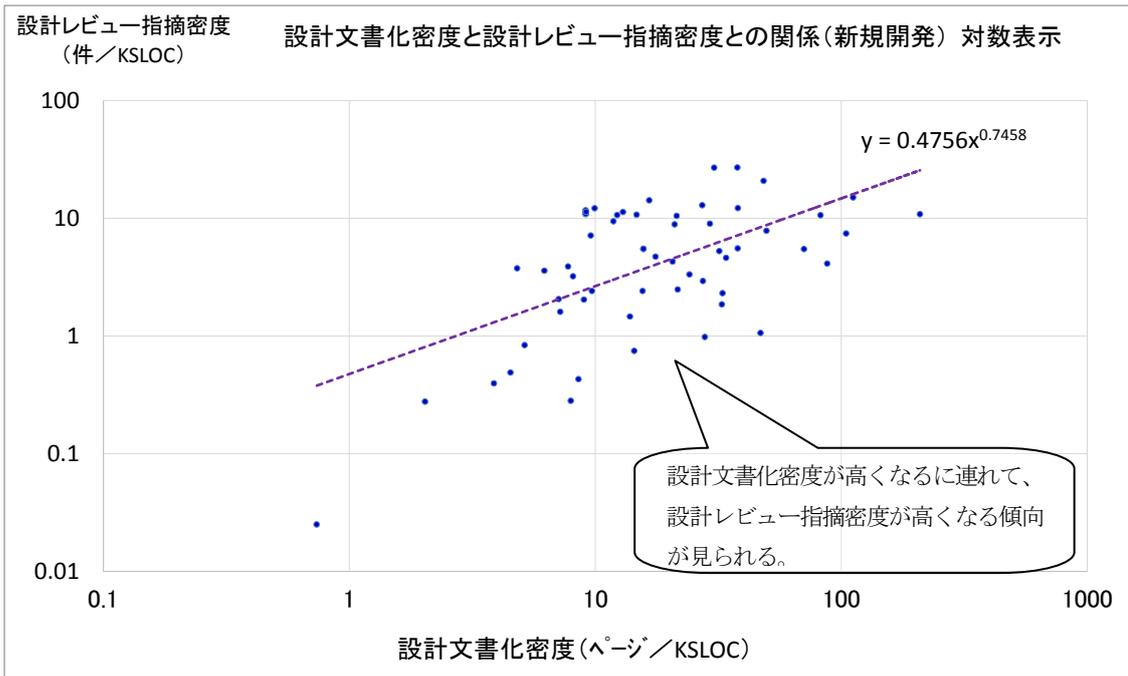
設計レビュー指摘密度に次のように明らかな差が見られる。

- ・新規開発の場合、中央値で約 1.8 倍の差が見られる。
(対数化後のデータを用いた Welch の t 検定によれば、1%有意で差が認められる。)
- ・改良開発の場合、中央値で約 5.1 倍の差が見られる。
(対数化後のデータを用いた Welch の t 検定によれば、1%有意で差が認められる。)

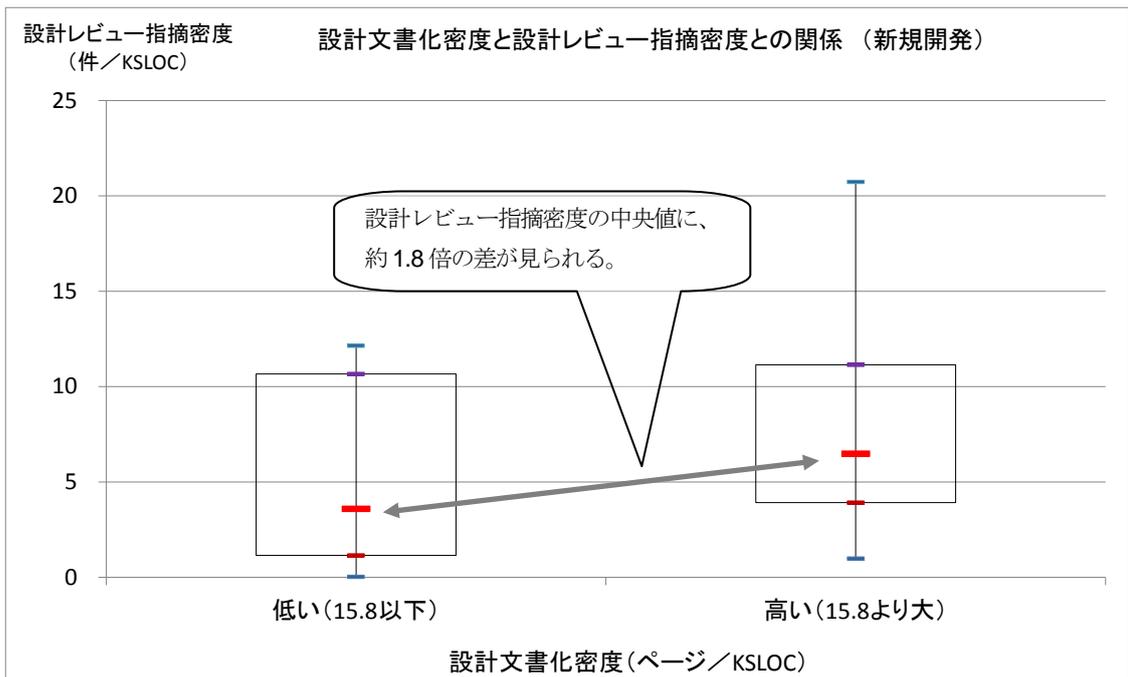
ここで、設計文書化密度の中央値は、新規開発の場合約 15.8 ページ/KSLOC、改良開発の場合約 44.7 ページ/KSLOC である。

<新規開発の場合> 散布図



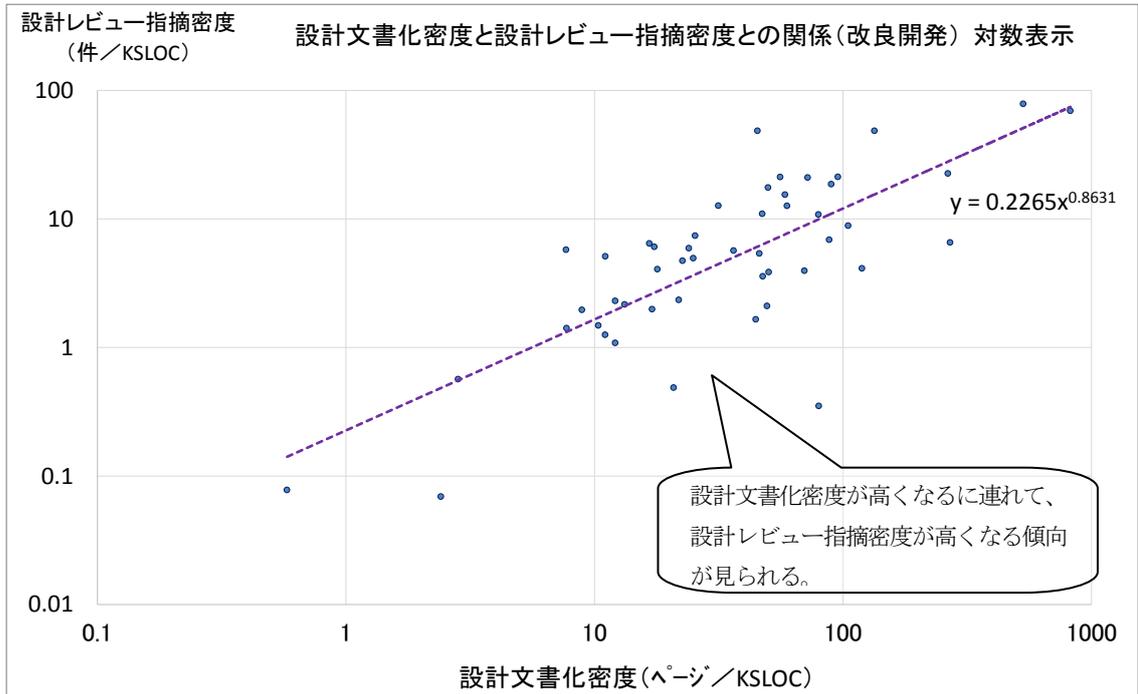
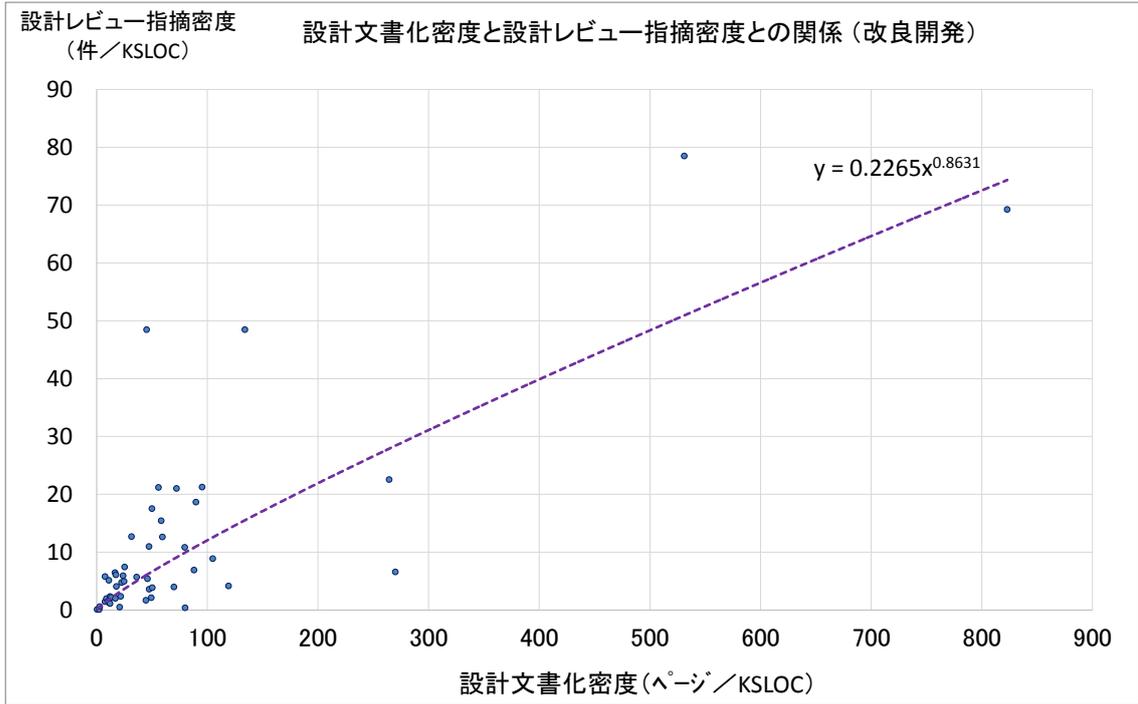


箱ひげ図

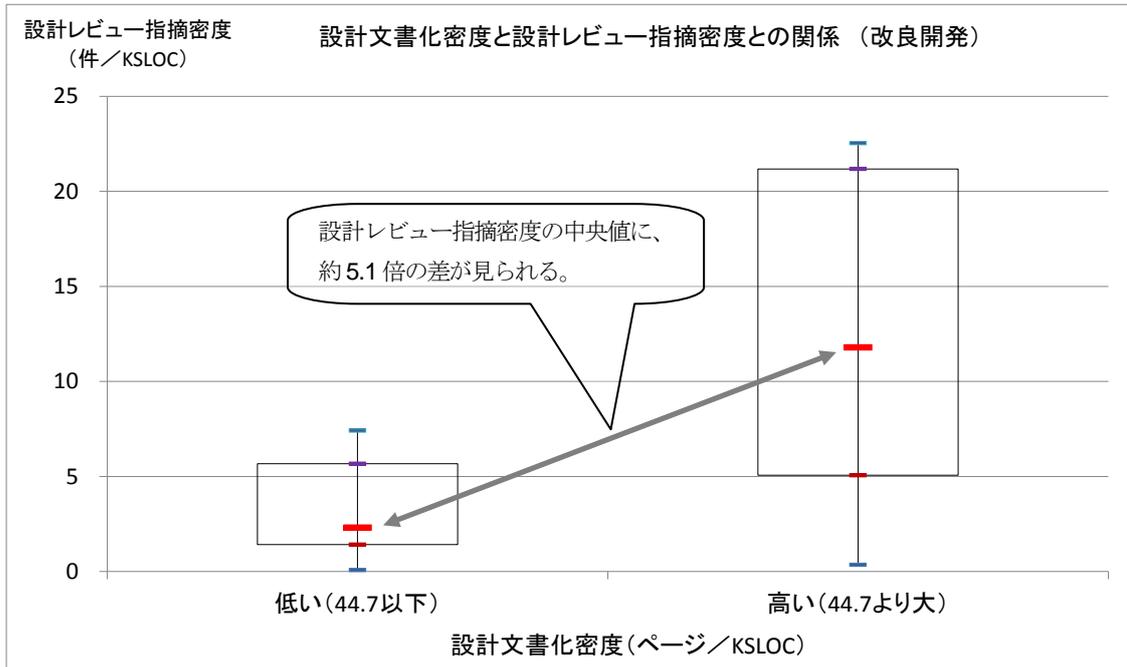


設計文書化密度	N	設計レビュー指摘密度 (件/KSLOC)						
		最小	P25	中央	P75	最大	平均	標準偏差
低い (15.8以下)	31	0.03	1.15	3.58	10.66	12.14	5.09	4.47
高い (15.8より大)	28	0.98	3.91	6.47	11.14	26.92	8.66	7.02

<改良開発の場合> 散布図



箱ひげ図



設計文書化密度	N	設計レビュー指摘密度 (件/KSLOC)						
		最小	P25	中央	P75	最大	平均	標準偏差
低い (44.7以下)	25	0.07	1.41	2.30	5.67	12.67	3.50	2.96
高い (44.7より大)	24	0.35	5.06	11.78	21.18	78.46	19.25	21.00

(3) 設計レビュー効果と能率を勘案した、設計レビュー工数のコントロール

【概要】

分析項目	設計レビュー工数密度と設計レビュー指摘密度との関係 設計レビュー工数密度と設計レビュー検出能率との関係
メッセージ	<p>◇分析結果概要（読み取れる傾向）</p> <p>設計レビュー工数密度が高くなるに連れて、設計レビュー指摘密度が高くなる傾向が見られる。具体的には、新規開発の場合、設計レビュー工数密度が 75 パーセンタイル値（約 9.8 人時/KSLOC）より大きい領域では、設計レビュー工数密度が 25 パーセンタイル値（約 1.5 人時/KSLOC）以下の領域と比較して、設計レビュー指摘密度の中央値が約 4.8 倍大きくなっている。</p> <p>設計レビュー工数をかけるほど、設計レビューで指摘できる件数（ひいては検出できる不具合数）を増やすことができると言える。</p> <p>ただし、設計レビュー工数をかけるに連れて、検出能率（設計レビュー工数当りの指摘件数）は低下する傾向が見られる。具体的には、新規開発の場合、設計レビュー工数密度が 75 パーセンタイル値（約 9.8 人時/KSLOC）より大きい領域では、設計レビュー工数密度が 25 パーセンタイル値（約 1.5 人時/KSLOC）以下の領域と比較して、設計レビュー検出能率の中央値が約 1/2.7 倍に低下している。</p> <p>◇助言</p> <p>経済的に信頼性を確保するためのレビューのコントロールと実現性の高い開発計画策定に向けて、検出能率の低下が飽和しない範囲で（目安としては設計レビュー工数密度の 75 パーセンタイル値の辺りまで）、設計レビューを強化することをお勧めする。</p> <p>設計レビュー工数密度の 75 パーセンタイル値は、新規開発の場合約 9.8 人時/KSLOC であり、改良開発の場合約 11.3 人時/KSLOC である。</p> <p>◇参考文献</p> <p>レビューの充実に関連して、次の IPA/SEC の文献を併せて参考にして頂けると幸いである。</p> <p>○「高信頼化ソフトウェアのための開発手法ガイドブック」 http://www.ipa.go.jp/files/000005144.pdf 3.2.2 項「レビュー手法の概要」、5 章「トレーサビリティ管理の手法」</p>
メトリクス	<p>◇設計レビュー工数密度： 基本設計から製作までの設計レビュー工数÷開発規模 (KSLOC)。</p> <p>◇設計レビュー指摘密度： 基本設計から製作までのレビュー指摘件数÷開発規模 (KSLOC)。</p> <p>◇設計レビュー検出能率： 設計レビュー指摘件数÷設計レビュー工数。</p>
考察/備考	<p>◇設計時の不具合を設計レビューによって 100%抽出するのが困難な現状では、設計レビューの工数をかけるほど検出できる不具合数が増えるのは当然と言える。また、設計レビューの工数をかけるに連れて検出能率（設計レビュー工数当りの検出不具合数）が低下することも経験的に明らかである。（確率モデルで考えると、例えば地雷原の地雷が少なくなれば地雷を踏む確率が低下するのに似ている。）</p> <p>開発現場では、無制限に設計レビューを増強するのは困難であり、経済的に（効</p>

	<p>率的に) 設計レビューを遂行することが求められるので、当該メッセージを提示している。</p> <p>なお、経済的に (効率的に) レビューを遂行するための方策としては、レビュー工数面だけでなく、次のような方策も併せて検討することが望ましい。</p> <ul style="list-style-type: none"> ・ 設計関連標準類の改良/整備 設計書作成標準、レビューチェックリスト、解析ツール、レビュー手法等 ・ 設計文書量の増強 設計文書化密度 (設計書ページ数÷開発規模) を高めることによって、設計レビューで見つかる不具合数を増やせる。白書等の分析結果からは、目安として約 16 ページ/KSLOC 以上とすることが望ましい (新規開発の場合)。 <p>◇設計レビューの収束性評価に関しては、「設計レビュー検出能率がある一定のレベルまで低下しているか否か」も、収束性評価項目の一つとして利用可能と考えられる。</p>
--	--

【検索条件】

- ◇ 開発 5 工程が揃っているプロジェクト
- ◇ 開発プロジェクトの種別が新規開発又は改良開発 (改修・保守又は拡張)
- ◇ 開発規模 (SLOC 規模) > 0
- ◇ 主開発言語が COBOL, C, VB, Java のいずれか
- ◇ 開発 5 工程の総工数 > 0
- ◇ 稼働後の発生不具合数 >= 0
- ◇ 基本設計から製作までのレビュー工数 > 0
- ◇ 基本設計から製作までのレビュー指摘数 > 0

【分析結果】

設計レビュー工数密度が高くなるに連れて、設計レビュー指摘密度が高くなる傾向が見られる。

具体的には、新規開発の場合、設計レビュー工数密度が **75** パーセンタイル値 (約 **9.8** 人時/KSLOC) より大きい領域では、設計レビュー工数密度が **25** パーセンタイル値 (約 **1.5** 人時/KSLOC) 以下の領域と比較して、設計レビュー指摘密度の中央値が約 **4.8** 倍大きくなっている。

ただし、設計レビュー工数密度が高くなるに連れて、設計レビュー検出能率が低下する傾向が見られる。

具体的には、新規開発の場合、設計レビュー工数密度が **75** パーセンタイル値 (約 **9.8** 人時/KSLOC) より大きい領域では、設計レビュー工数密度が **25** パーセンタイル値 (約 **1.5** 人時/KSLOC) 以下の領域と比較して、設計レビュー検出能率の中央値が約 **1/2.7** 倍に低下している。

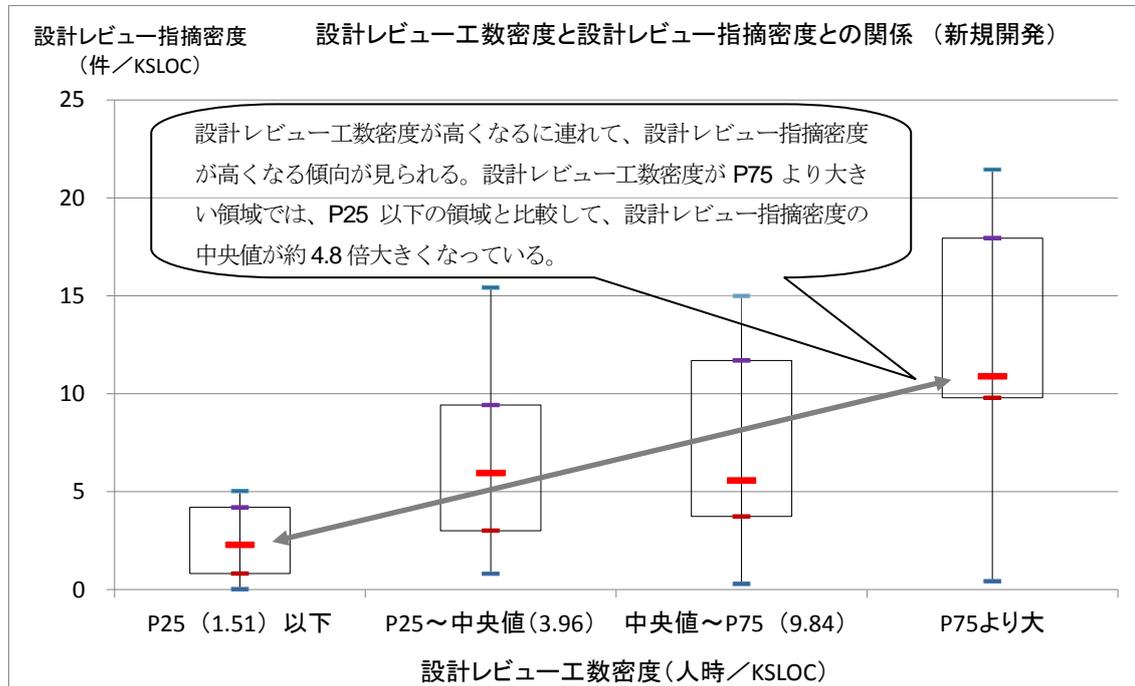
改良開発の場合、設計レビュー工数密度が **75** パーセンタイル値 (約 **11.3** 人時/KSLOC) より大きい領域では、設計レビュー工数密度が **25** パーセンタイル値 (約 **2.4** 人時/KSLOC) 以下の領域と比較して、設計レビュー指摘密度の中央値が約 **6.0** 倍大きくなっている。

ただし、設計レビュー工数密度が高くなるに連れて、設計レビュー検出能率が低下する傾向が見られる。

具体的には、改良開発の場合、設計レビュー工数密度が **75** パーセンタイル値 (約 **11.3** 人時/KSLOC) より大きい領域では、設計レビュー工数密度が **25** パーセンタイル値 (約 **2.4** 人時/KSLOC) 以下の領域と比較して、設計レビュー検出能率の中央値が約 **1/3.0** 倍に低下している。

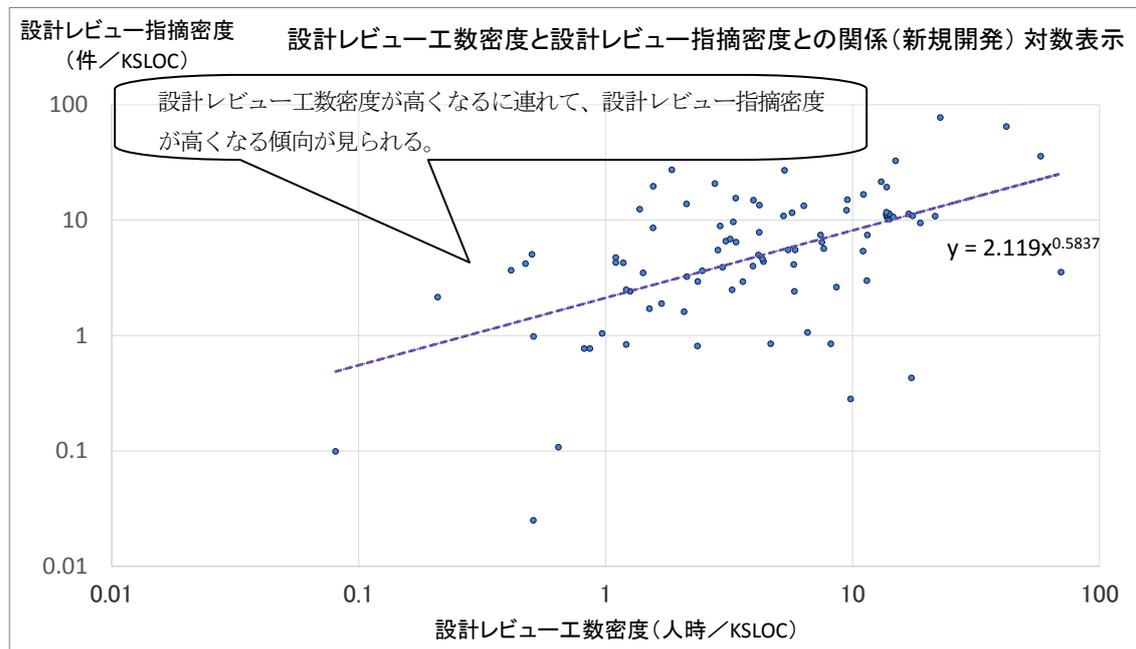
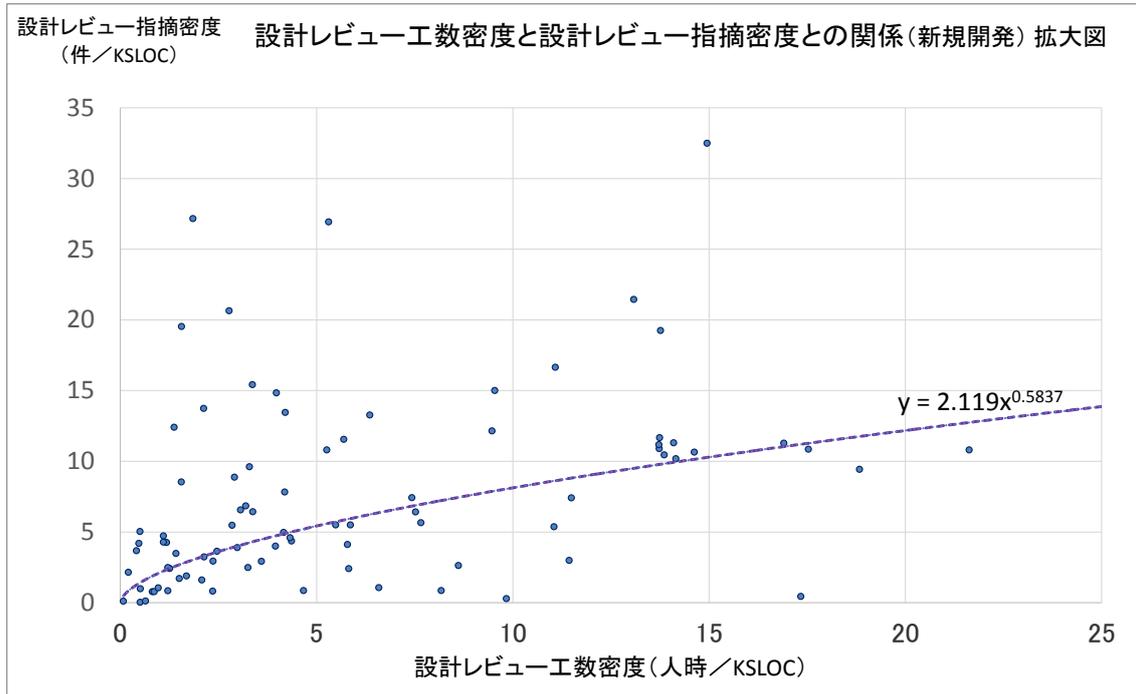
<新規開発の場合>

◇箱ひげ図（設計レビュー工数密度と設計レビュー指摘密度との関係）

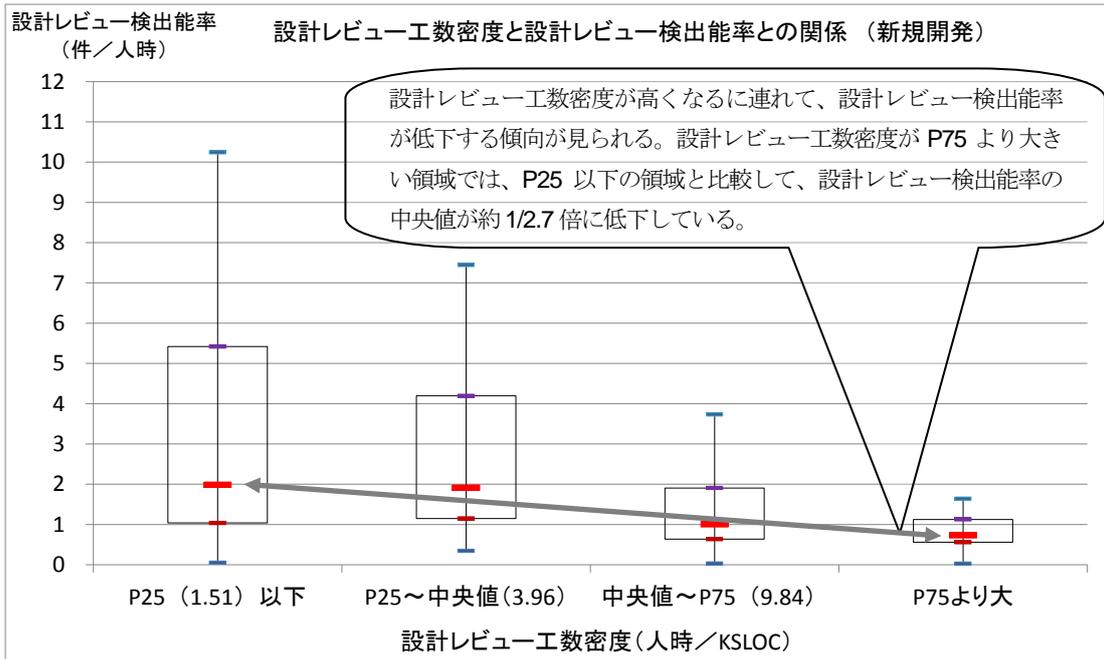


設計レビュー工数密度	N	最小	P25	中央	P75	設計レビュー指摘密度 (件/KSLOC)		
						最大	平均	標準偏差
P25 (1.51) 以下	28	0.03	0.82	2.27	4.20	12.39	2.77	2.82
P25~中央値 (3.96)	27	0.81	3.01	5.94	9.42	27.16	8.01	7.07
中央値~P75 (9.84)	27	0.28	3.74	5.57	11.69	26.92	7.59	6.20
P75より大	27	0.43	9.79	10.88	17.95	77.07	17.62	18.79

◇散佈図（設計レビュー工数密度と設計レビュー指摘密度との関係）

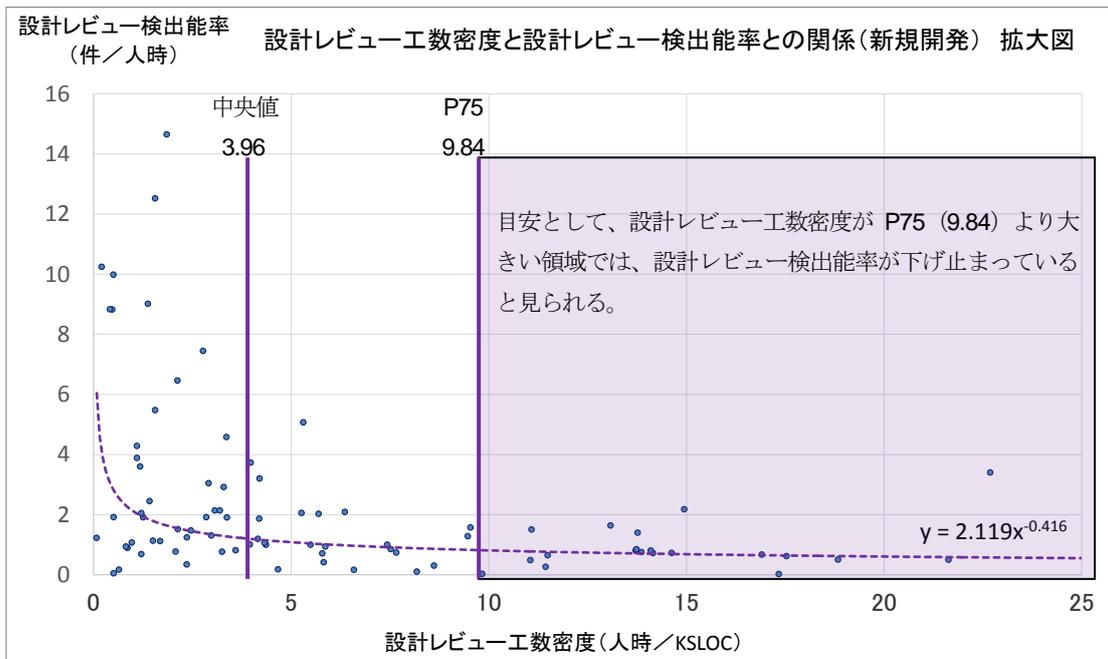


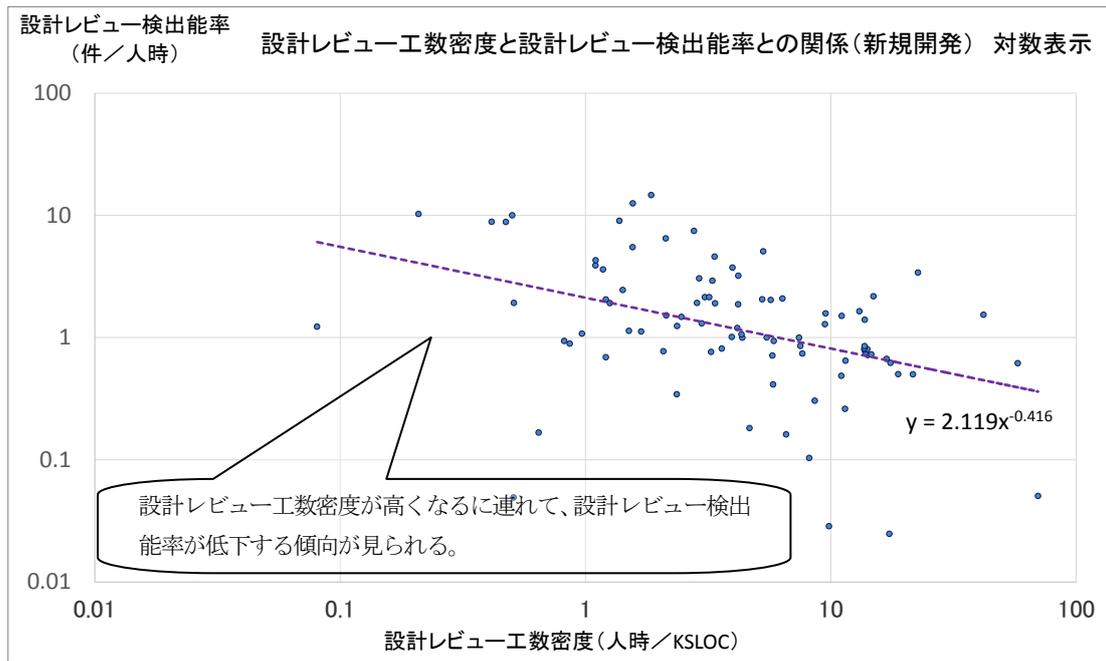
◇箱ひげ図 (設計レビュー工数密度と設計レビュー検出能率との関係)



設計レビュー工数密度	N	最小	P25	中央	P75	設計レビュー検出能率 (件/ケース)		
						最大	平均	標準偏差
P25 (1.51) 以下	28	0.05	1.04	1.98	5.42	10.24	3.66	3.58
P25~中央値 (3.96)	27	0.34	1.15	1.91	4.20	14.64	3.43	3.82
中央値~P75 (9.84)	27	0.03	0.64	1.00	1.91	5.07	1.36	1.22
P75より大	27	0.02	0.56	0.73	1.12	3.40	0.93	0.74

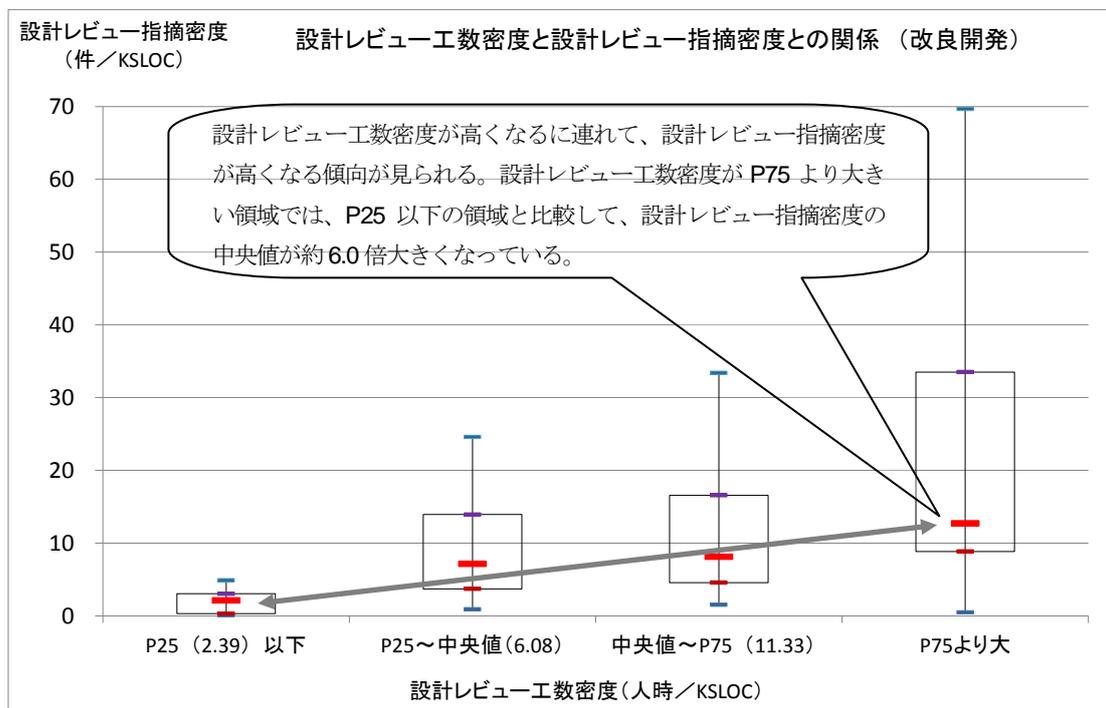
◇散布図 (設計レビュー工数密度と設計レビュー検出能率との関係)





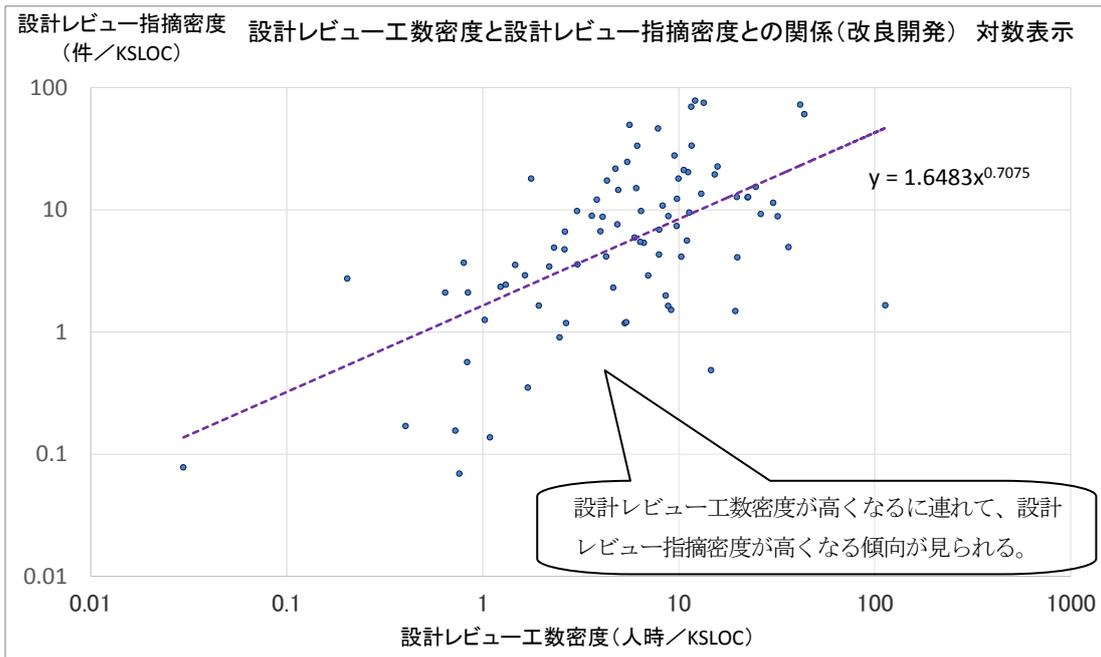
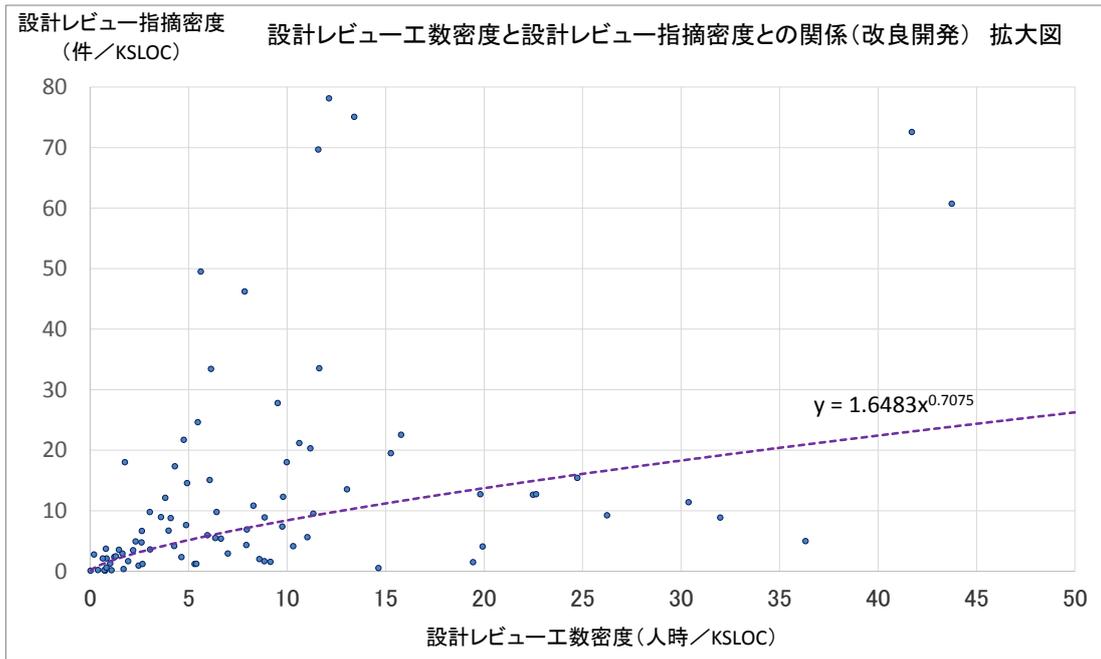
<改良開発の場合>

◇箱ひげ図 (設計レビュー工数密度と設計レビュー指摘密度との関係)

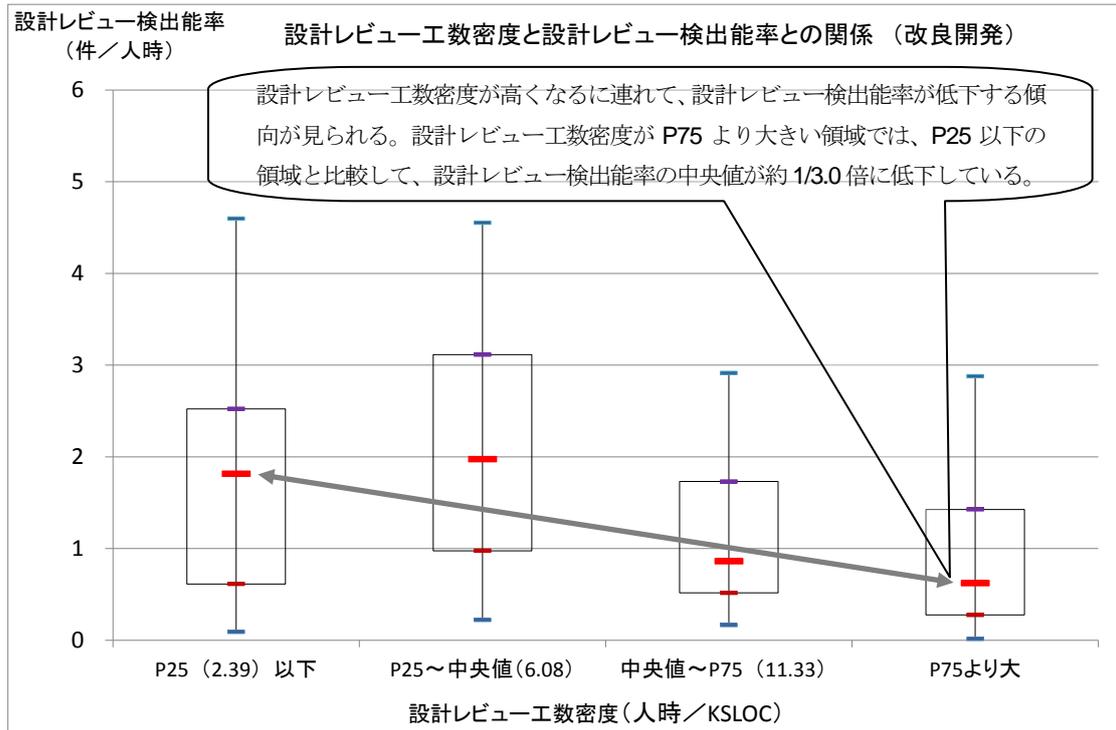


設計レビュー工数密度	N	設計レビュー指摘密度 (件/KSLOC)						
		最小	P25	中央	P75	最大	平均	標準偏差
P25 (2.39) 以下	24	0.07	0.30	2.10	3.04	18.00	2.63	3.90
P25~中央値 (6.08)	23	0.90	3.71	7.13	13.94	49.48	10.37	10.99
中央値~P75 (11.33)	23	1.51	4.57	8.12	16.57	46.19	12.05	11.51
P75より大	23	0.49	8.85	12.71	33.51	78.11	25.74	27.26

◇散布図（設計レビュー工数密度と設計レビュー指摘密度との関係）

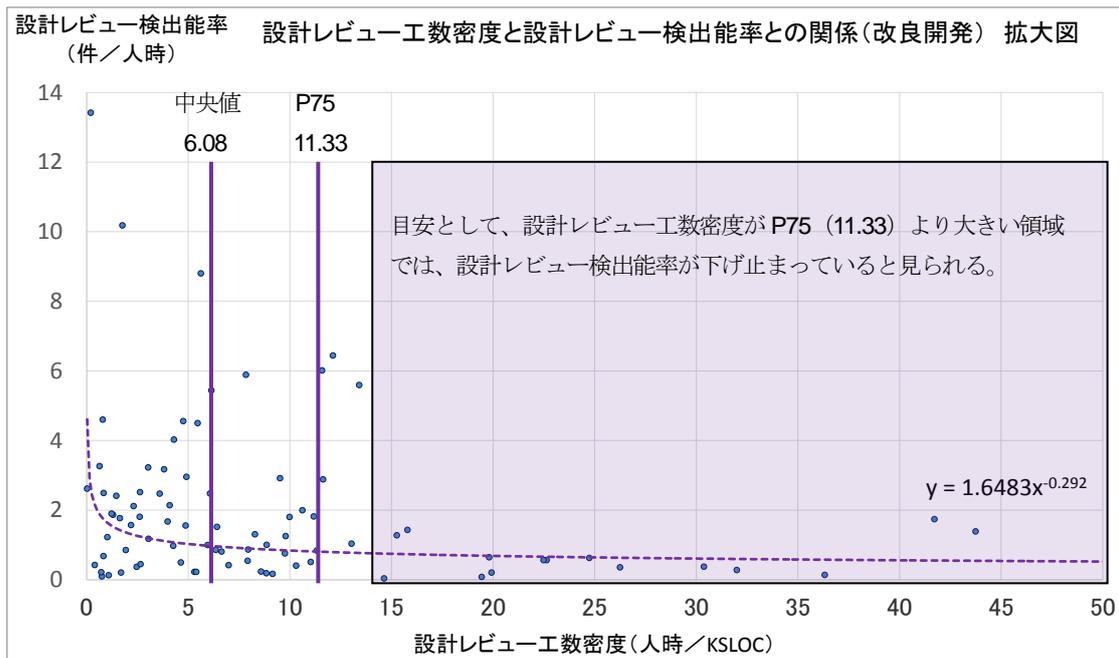


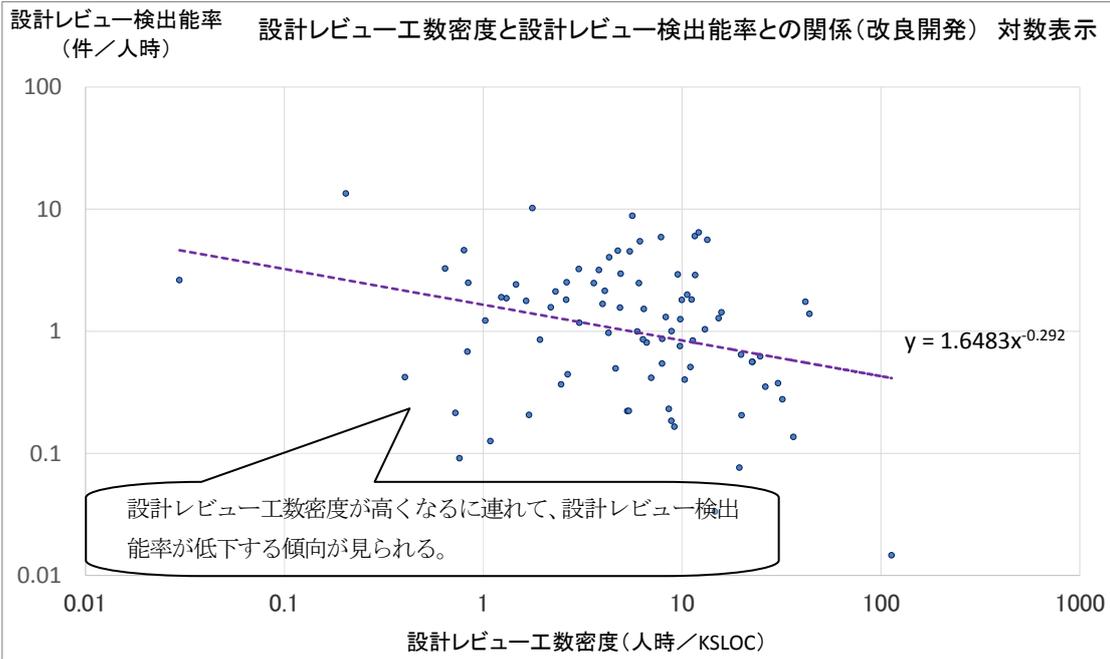
◇箱ひげ図（設計レビュー工数密度と設計レビュー検出能率との関係）



設計レビュー工数密度	N	最小	P25	中央	P75	設計レビュー検出能率 (件/ケース)		
						最大	平均	標準偏差
P25 (2.39) 以下	24	0.09	0.61	1.81	2.52	13.41	2.60	3.40
P25~中央値 (6.08)	23	0.22	0.98	1.97	3.11	8.80	2.31	1.98
中央値~P75 (11.33)	23	0.17	0.52	0.86	1.73	5.88	1.43	1.53
P75より大	23	0.01	0.28	0.62	1.43	6.44	1.51	2.01

◇散布図（設計レビュー工数密度と設計レビュー検出能率との関係）





2.2. テストのコントロール

(1) テスト評価の一方法

■ゾーン分析によるテスト評価

【概要】

分析項目	テスト密度対テスト検出不具合密度のゾーン別発生不具合密度
メッセージ	<p>◇分析結果概要（読み取れる傾向）</p> <p>相対的にテスト密度が低くテスト検出不具合密度が高い集合と、相対的にテスト密度が高くテスト検出不具合密度が低い集合とでは、後者の方が発生不具合密度が低い（相対的に信頼性が高い）傾向が見られる。</p> <p>具体的には、新規開発の場合、テスト密度が中央値（約 30.8 ケース/KSLOC）以下でテスト検出不具合密度が中央値（約 1.60 件/KSLOC）より大の領域の発生不具合密度の中央値が 0.022 件/KSLOC であるのに対して、テスト密度が中央値より大でテスト検出不具合密度が中央値以下の領域の発生不具合密度の中央値は 0 件/KSLOC である。また、発生不具合密度の 75 パーセンタイル値には約 3.3 倍の開きが見られる。</p> <p>◇助言</p> <p>テスト密度が高くテスト検出不具合密度が低いのは相対的に信頼性が良い兆候の一つである。一方、テスト密度が低くてテスト検出不具合密度が高いのは相対的に信頼性が良くない兆候の一つである。信頼性を確保するためのテストのコントロールに向けて、この見方をテストの評価項目の一つとして採用することをお勧めする。</p> <p>新規開発の場合、テスト密度の中央値は約 30.8 ケース/KSLOC、テスト検出不具合密度の中央値は約 1.60 件/KSLOC である。改良開発の場合、テスト密度の中央値は約 63.9 ケース/KSLOC、テスト検出不具合密度の中央値は約 1.50 件/KSLOC である。</p> <p>◇参考文献</p> <p>定量データに基づく品質予測・評価については、次の IPA/SEC の文献を併せて参考にして頂けると幸いである。</p> <ul style="list-style-type: none"> ○「定量的品質予測のススメ」 http://www.ipa.go.jp/sec/publish/tn08-004.html ○「続 定量的品質予測のススメ」 http://www.ipa.go.jp/sec/publish/tn10-004.html
メトリクス	<p>◇テスト密度：(結合テストケース数+総合テスト（ベンダ確認）ケース数) ÷ 開発規模（KSLOC）。</p> <p>◇テスト検出不具合密度：(結合テスト検出不具合数+総合テスト（ベンダ確認）検出不具合数) ÷ 開発規模（KSLOC）。</p> <p>◇発生不具合密度：稼働後（最長 6 ヶ月以内）に発生した不具合数 ÷ 開発規模（KSLOC）。</p>
考察/備考	<p>「テスト密度が高くテスト検出不具合密度が低いのは相対的に信頼性が良い（稼働後の不具合発生が少ない）兆候の一つである」の主な要因として、次のことが考えられる。</p> <p>◇テスト密度が高いのにテスト検出不具合密度が低いものは、テスト開始時点の</p>

	<p>出来が良い（潜在不具合の密度が低い）、つまりいわゆる作込み品質が良いものと考えられる。</p> <p>◇テスト検出不具合密度が高いものは発生不具合密度（稼働後の不具合密度）も高く、テスト検出不具合密度が低いものは発生不具合密度（稼働後の不具合密度）も低い傾向が見られる。つまり、テストに至るまでの良し悪しがテストによって逆転するケースは少ないと言える。</p> <p>信頼性を高めるには、やはり作込み品質向上を目指すことが王道であり、テストによって挽回しようという作戦の成算は薄いと考えられる。</p>
--	--

【検索条件】

- ◇ 開発 5 工程が揃っているプロジェクト
- ◇ 開発プロジェクトの種別が新規開発又は改良開発（改修・保守又は拡張）
- ◇ 開発規模（SLOC 規模） > 0
- ◇ 主開発言語が COBOL, C, VB, Java のいずれか
- ◇ 開発 5 工程の総工数 > 0
- ◇ 稼働後の発生不具合数 >= 0
- ◇ 結合テストケース数+総合テスト（ベンダ確認）ケース数 > 0
- ◇ 結合テスト検出不具合数+総合テスト（ベンダ確認）検出不具合数 > 0

【分析結果】

相対的にテスト密度が低くテスト検出不具合密度が高い集合と、相対的にテスト密度が高くテスト検出不具合密度が低い集合とでは、後者の方が発生不具合密度が低い（相対的に信頼性が高い）傾向が見られる。

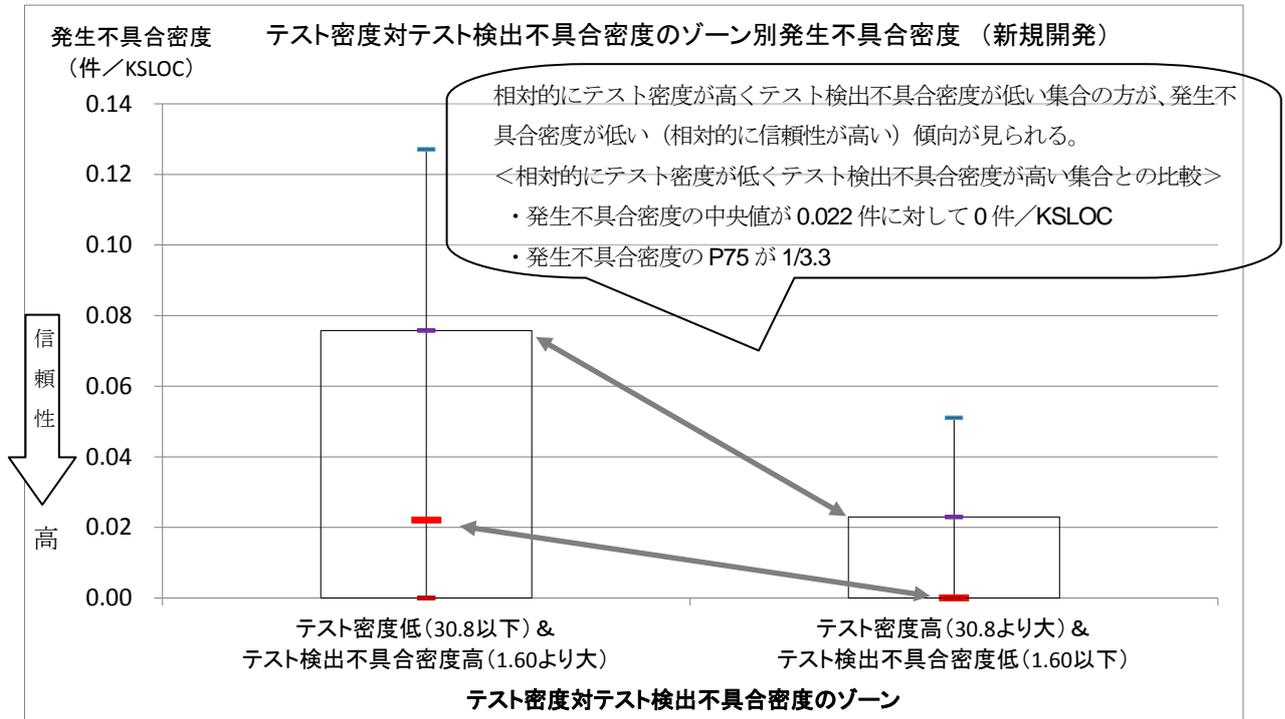
具体的には、新規開発の場合、テスト密度が中央値（約 30.8 ケース/KSLOC）以下でテスト検出不具合密度が中央値（約 1.60 件/KSLOC）より大の領域の発生不具合密度の中央値が 0.022 件/KSLOC であるのに対して、テスト密度が中央値より大でテスト検出不具合密度が中央値以下の領域の発生不具合密度の中央値は 0 件/KSLOC である。

また、発生不具合密度の 75 パーセンタイル値には約 3.3 倍の開きが見られる。（対数化しないデータを用いた Welch の t 検定によれば、ほぼ 10%有意で差が認められる。）

改良開発の場合、テスト密度が中央値（約 63.9 ケース/KSLOC）以下でテスト検出不具合密度が中央値（約 1.50 件/KSLOC）より大の領域の発生不具合密度の中央値が 0.040 件/KSLOC であるのに対して、テスト密度が中央値より大でテスト検出不具合密度が中央値以下の領域の発生不具合密度の中央値は 0 件/KSLOC である。また、発生不具合密度の 75 パーセンタイル値には約 6.5 倍の開きが見られる。（対数化しないデータを用いた Welch の t 検定によれば、5%有意で差が認められる。）

<新規開発の場合>

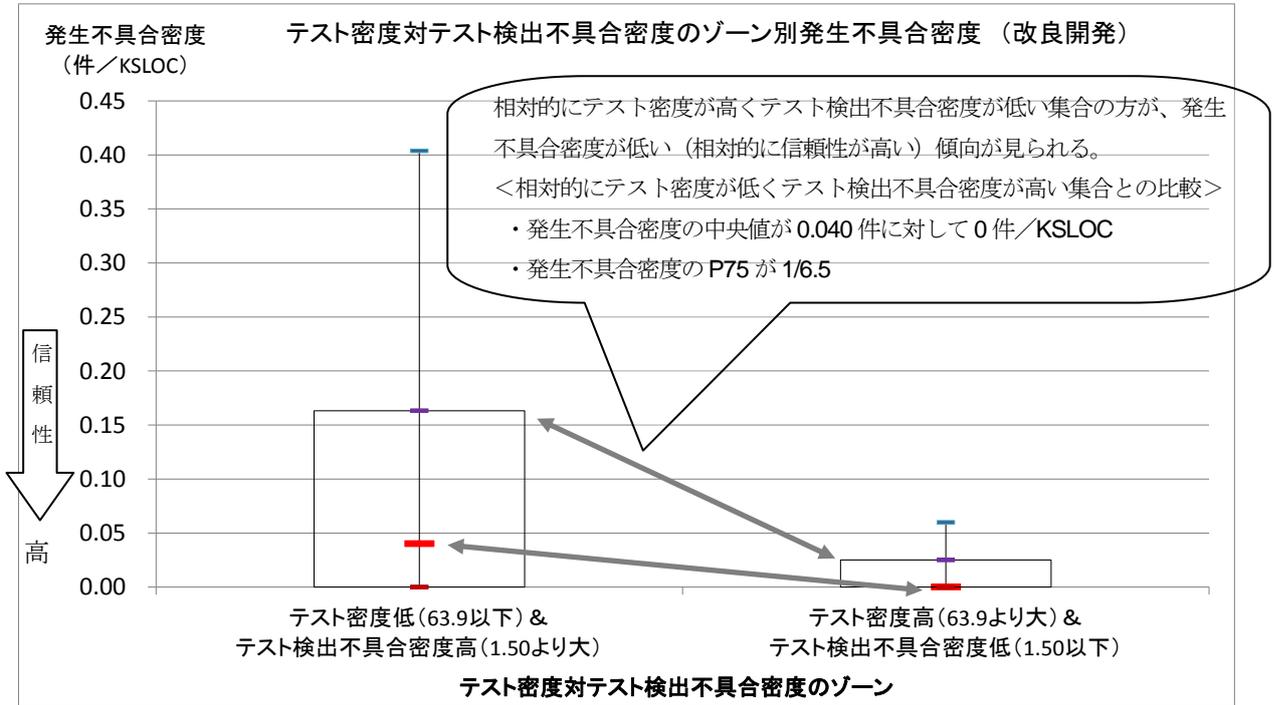
◇箱ひげ図



テスト密度対テスト検出不具合密度のゾーン	N	最小	P25	中央	P75	最大	発生不具合密度 (件/KSLOC)	
							平均	標準偏差
テスト密度低 (30.8以下) & テスト検出不具合密度高 (1.60より大)	75	0.000	0.000	0.022	0.076	4.708	0.159	0.566
テスト密度高 (30.8より大) & テスト検出不具合密度低 (1.60以下)	68	0.000	0.000	0.000	0.023	1.269	0.051	0.171

<改良開発の場合>

◇箱ひげ図



テスト密度対テスト検出不具合密度のゾーン	N	最小	P25	中央	P75	最大	発生不具合密度 (件/KSLOC)	
							平均	標準偏差
テスト密度低 (63.9以下) & テスト検出不具合密度高 (1.50より大)	55	0.000	0.000	0.040	0.163	3.194	0.223	0.523
テスト密度高 (63.9より大) & テスト検出不具合密度低 (1.50以下)	48	0.000	0.000	0.000	0.025	0.595	0.041	0.114

(2) テスト評価の一方法 (その2)

■テスト検出能率によるテストの収束性評価

【概要】

分析項目	テスト検出能率と発生不具合密度との関係
メッセージ	<p>◇分析結果概要 (読み取れる傾向)</p> <p>テスト検出能率が低い方が相対的に発生不具合密度が低い (信頼性が高い) 傾向が見られる。</p> <p>具体的には、新規開発の場合、テスト検出能率が中央値 (約 0.048 件/テストケース) 以下の集合と中央値より大きい集合とでは、発生不具合密度の中央値に約 2.4 倍の開きが見られる。特に、テスト検出能率が P25 (約 0.026 件/テストケース) 以下の集合では、発生不具合密度が低い (信頼性が高い) 傾向が顕著である。</p> <p>改良開発の場合、テスト検出能率が中央値 (約 0.025 件/テストケース) 以下の集合では発生不具合密度の中央値が 0 件/KSLOC であるのに対して、中央値より大きい集合では発生不具合密度の中央値が約 0.022 件/KSLOC である。特に、テスト検出能率が P25 (約 0.010 件/テストケース) 以下の集合では、発生不具合密度が低い (信頼性が高い) 傾向が顕著である。</p> <p>◇助言</p> <p>信頼性を確保するためのテストの監視・コントロールの一環として、次のような評価をテストの収束性評価項目の一つとすることをお勧めする。</p> <p>【「テスト検出能率がある一定のレベルまで低下しているか否か」を、テストの収束性評価項目の一つとして利用する。特に、追加テストの収束性評価に利用することをお勧めする】</p> <p>◇参考文献</p> <p>定量データに基づく品質予測・評価については、次の IPA/SEC の文献を併せて参考にして頂けると幸いである。</p> <ul style="list-style-type: none"> ○「定量的品質予測のススメ」 http://www.ipa.go.jp/sec/publish/tn08-004.html ○「続 定量的品質予測のススメ」 http://www.ipa.go.jp/sec/publish/tn10-004.html
メトリクス	<p>◇テスト検出能率：テスト検出不具合数÷テストケース数</p> <p>◇発生不具合密度：稼働後 (最長 6 ヶ月以内) に発生した不具合数÷開発規模</p>
考察/備考	<p>テスト検出能率が低下して行くということは、まだ潜在している (残存している) 不具合が減少して行くということを意味する。テスト検出能率によってテストの収束性を評価することは、理に適っていると考えられる。</p> <p>IPA/SEC データベースの分析結果からは、テスト検出能率が次 25 パーセント値を目安として下回ることが望ましいと考える。</p> <ul style="list-style-type: none"> ・新規開発の場合、約 0.025 件/テストケース 40 ケース以上のテストに対して 1 件の不具合検出に相当。 ・改良開発の場合、約 0.010 件/テストケース 100 ケース以上のテストに対して 1 件の不具合検出に相当。

[検索条件等]

- ・開発 5 工程が揃っているプロジェクト
- ・開発プロジェクトの種別が新規開発又は改良開発（改修・保守又は拡張）
- ・開発規模（SLOC 規模） > 0
- ・主開発言語が COBOL, C, VB, Java のいずれか
- ・開発 5 工程の総工数 > 0
- ・稼働後の発生不具合数 >= 0
- ・結合テストケース数+総合テスト（ベンダ確認）ケース数 > 0
- ・結合テスト検出不具合数+総合テスト（ベンダ確認）検出不具合数 > 0

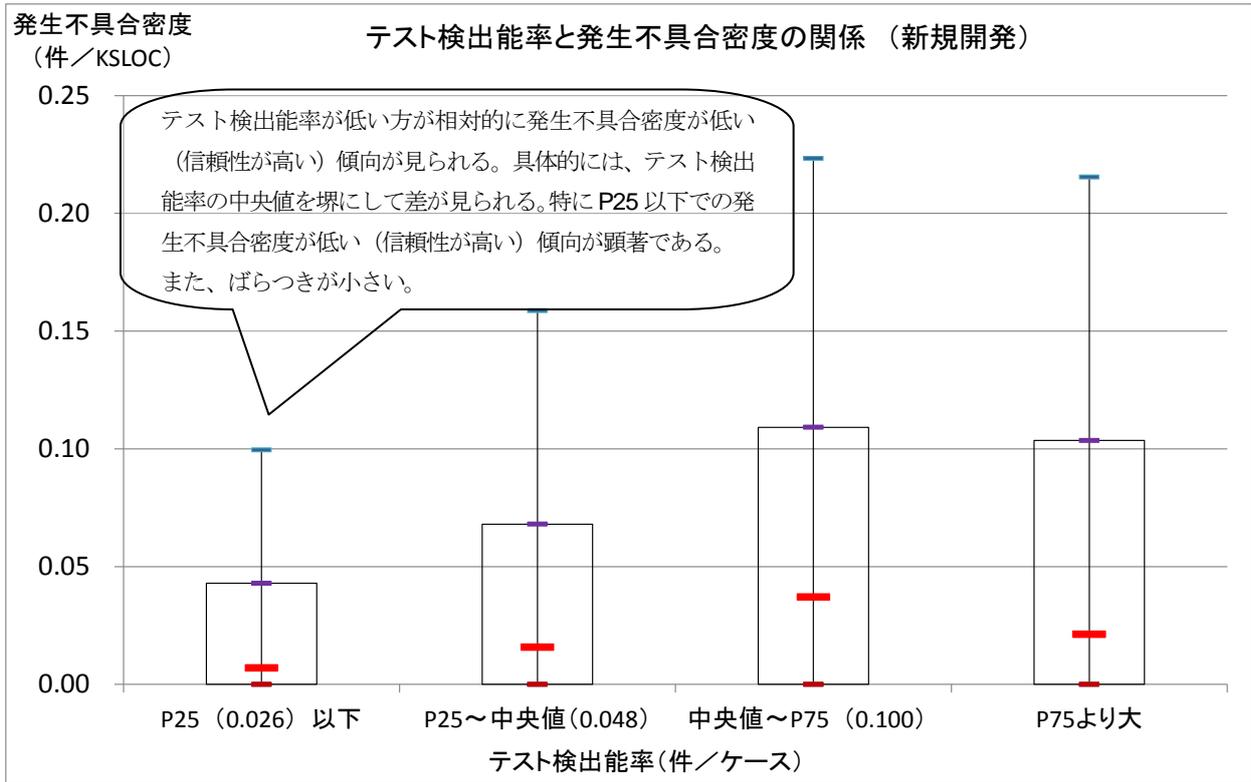
[分析結果 1：結合テスト、総合テスト（ベンダ確認）総合での分析結果]

テスト検出能率が低い方が相対的に発生不具合密度が低い（信頼性が高い）傾向が見られる。具体的には、新規開発の場合、テスト検出能率が中央値（約 0.048 件/テストケース）以下の集合と中央値より大きい集合とでは、発生不具合密度の中央値に約 2.4 倍の開きが見られる。（対数化しないデータを用いた Welch の t 検定によれば、5%有意で差が認められる。）特に、テスト検出能率が P25（約 0.026 件/テストケース）以下の集合では、発生不具合密度が低い（信頼性が高い）傾向が顕著である。

改良開発の場合、テスト検出能率が中央値（約 0.025 件/テストケース）以下の集合では発生不具合密度の中央値が 0 件/KSLOC であるのに対して、中央値より大きい集合では発生不具合密度の中央値が約 0.022 件/KSLOC である。（対数化しないデータを用いた Welch の t 検定によれば、1%有意で差が認められる。）特に、テスト検出能率が P25（約 0.010 件/テストケース）以下の集合では、発生不具合密度が低い（信頼性が高い）傾向が顕著である。

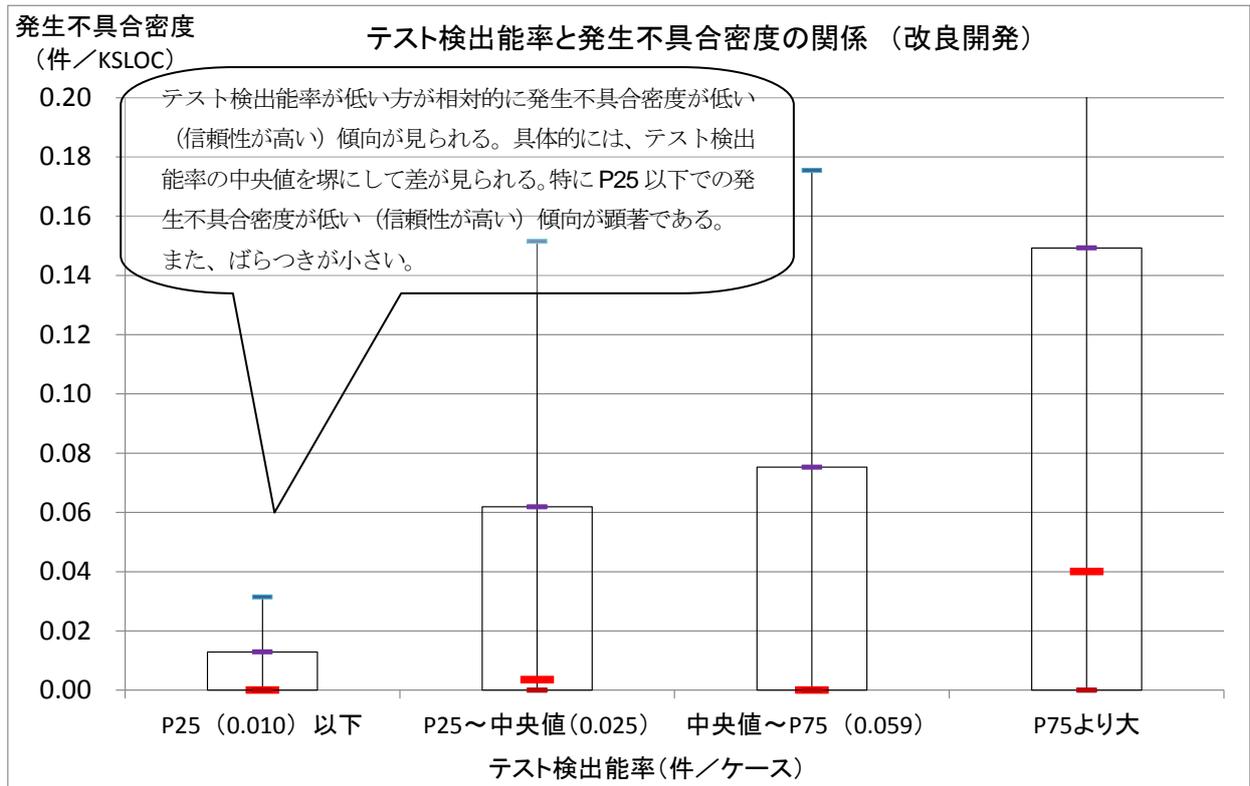
ここで、テスト検出能率は、 $(\text{結合テスト検出不具合数} + \text{総合テスト（ベンダ確認）検出不具合数}) \div (\text{結合テストケース数} + \text{総合テスト（ベンダ確認）ケース数})$ 。発生不具合密度は、稼働後（最長 6 ヶ月以内）に発生した不具合数 ÷ 開発規模（KSLOC）。

◇新規開発の場合



テスト検出能率	N	最小	P25	中央	P75	最大	発生不具合密度 (件/KSLOC)	
							平均	標準偏差
P25 (0.026) 以下	94	0.0000	0.0000	0.0069	0.0429	1.2690	0.0643	0.1842
P25~中央値 (0.048)	93	0.0000	0.0000	0.0157	0.0680	0.5747	0.0646	0.1139
中央値~P75 (0.100)	93	0.0000	0.0000	0.0370	0.1091	4.8392	0.1911	0.6156
P75より大	94	0.0000	0.0000	0.0212	0.1036	4.7082	0.1433	0.5114

◇改良開発の場合



テスト検出能率	N	発生不具合密度 (件/KSLOC)						
		最小	P25	中央	P75	最大	平均	標準偏差
P25 (0.010) 以下	73	0.0000	0.0000	0.0000	0.0129	0.4947	0.0263	0.0784
P25~中央値 (0.025)	73	0.0000	0.0000	0.0035	0.0619	0.5946	0.0515	0.1033
中央値~P75 (0.059)	73	0.0000	0.0000	0.0000	0.0753	5.8451	0.1772	0.7029
P75より大	73	0.0000	0.0000	0.0400	0.1493	3.1944	0.2081	0.5370

(3) テスト効果と能率を勘案した、テスト量のコントロール

【概要】

分析項目	テスト密度とテスト検出不具合密度との関係 テスト密度とテスト検出能率との関係
メッセージ	<p>◇分析結果概要（読み取れる傾向）</p> <p>テスト密度が高くなるに連れて、テスト検出不具合密度が高くなる傾向が見られる。具体的には、新規開発の場合、テスト密度が 75 パーセンタイル値（約 58.6 ケース/KSLOC）より大きい領域では、テスト密度が 25 パーセンタイル値（約 13.0 ケース/KSLOC）以下の領域と比較して、テスト検出不具合密度の中央値が約 2.9 倍大きくなっている。</p> <p>ただし、テスト密度が高くなるに連れて、テスト検出能率が大幅に低下する傾向が見られる。具体的には、新規開発の場合、テスト密度が 75 パーセンタイル値より大きい領域では、テスト密度が 25 パーセンタイル値以下の領域と比較して、テスト検出能率の中央値が約 1/6.0 倍に低下している。</p> <p>◇助言</p> <p>テストケースを増やすほど、テストでより多くの不具合を検出することができる。ただし、テストケースを増やすに連れて、テスト検出能率（テストケース当りの検出不具合数）は大幅に低下する。</p> <p>経済的に信頼性を確保するためのテストのコントロールと実現性の高い開発計画策定に向けて、検出能率の低下が飽和しない範囲で（目安としてはテスト密度の 75 パーセンタイル値の辺りまで）、テストを強化することをお勧めする。</p> <p>テスト密度の 75 パーセンタイル値は、新規開発の場合約 58.6 ケース/KSLOC であり、改良開発の場合約 147.9 ケース/KSLOC である。</p> <p>◇参考文献</p> <p>テストの充実に関連して、次の IPA/SEC の文献を併せて参考にして頂けると幸いである。</p> <p>○「高信頼化ソフトウェアのための開発手法ガイドブック」 http://www.ipa.go.jp/files/000005144.pdf 3.2.3 項「テスト手法の概要」、6 章「テスト網羅性の高度化技法」</p>
メトリクス	<p>◇テスト密度：(結合テストケース数+総合テスト（ベンダ確認）ケース数) ÷ 開発規模（KSLOC）。</p> <p>◇テスト検出不具合密度：(結合テスト検出不具合数+総合テスト（ベンダ確認）検出不具合数) ÷ 開発規模（KSLOC）。</p> <p>◇テスト検出能率：(結合テスト検出不具合数+総合テスト（ベンダ確認）検出不具合数) ÷ (結合テストケース数+総合テスト（ベンダ確認）ケース数)。</p>
考察/備考	<p>潜在している不具合をテストによって 100% 完全に検出するのが難しい現状では、テストケース数を増やすほど検出できる不具合数が増えるのは当然と言える。また、テストケース数を増やすに連れて検出能率（テストケース数当りの検出不具合数）が低下することも経験的に明らかである。（確率モデルで考えると、例えば地雷原の地雷が少なくなれば、地雷を踏む確率が低下するのに似ている。）</p> <p>開発現場では、無制限にテストを増強するのは困難であり、経済的に（効率的に）テストを遂行することが求められるので、テスト密度に関する当該メッセージを提示している。</p>

	<p>なお、経済的に（効率的に）テストを遂行するための方策としては、テスト密度面だけでなく、次のようなテスト関連標準類の改良／整備も併せて検討することが望ましい。</p> <ul style="list-style-type: none"> ・テスト設計手法 ・テスト設計のチェックリスト ・静的／動的解析ツール ・リグレッションテスト支援ツール （再現実行ツール、テスト結果の自動判定ツール等） ・テスト仕様書作成標準
--	---

【検索条件】

- ◇ 開発 5 工程が揃っているプロジェクト
- ◇ 開発プロジェクトの種別が新規開発又は改良開発（改修・保守又は拡張）
- ◇ 開発規模（SLOC 規模） > 0
- ◇ 主開発言語が COBOL, C, VB, Java のいずれか
- ◇ 開発 5 工程の総工数 > 0
- ◇ 稼働後の発生不具合数 ≥ 0
- ◇ 結合テストケース数+総合テスト（ベンダ確認）ケース数 > 0
- ◇ 結合テスト検出不具合数+総合テスト（ベンダ確認）検出不具合数 > 0

【分析結果】

テスト密度が高くなるに連れて、テスト検出不具合密度が高くなる傾向が見られる。具体的には、新規開発の場合、テスト密度が 75 パーセンタイル値（約 58.6 ケース/KSLOC）より大きい領域では、テスト密度が 25 パーセンタイル値（約 13.0 ケース/KSLOC）以下の領域と比較して、テスト検出不具合密度の中央値が約 2.9 倍大きくなっている。

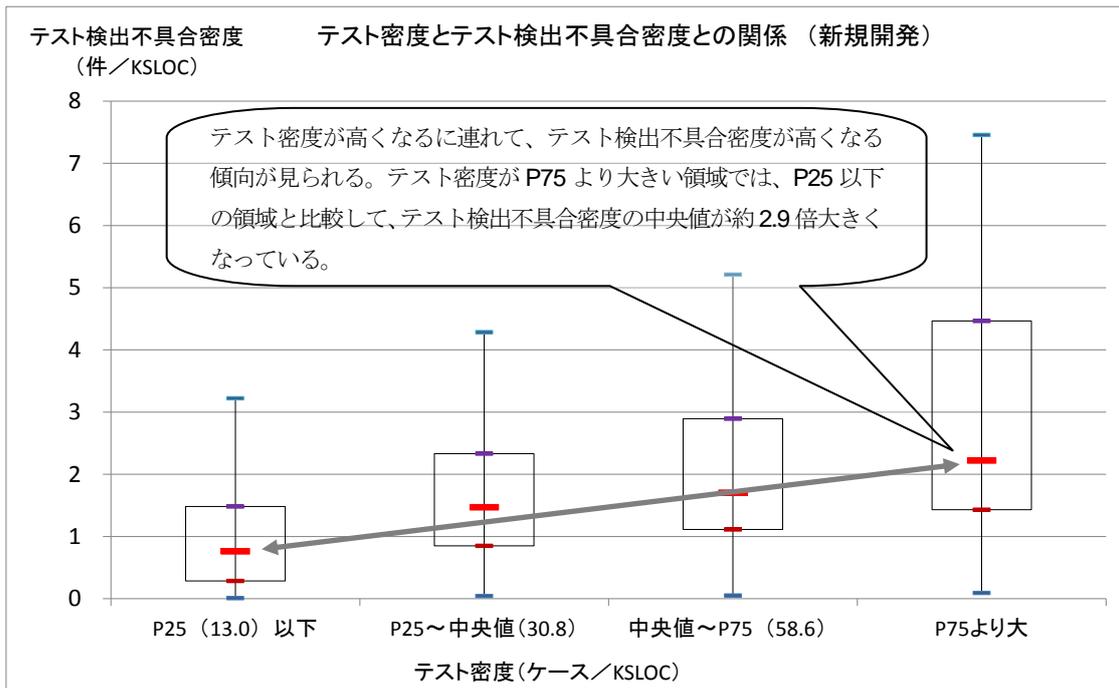
ただし、テスト密度が高くなるに連れて、テスト検出能率が大幅に低下する傾向が見られる。具体的には、新規開発の場合、テスト密度が 75 パーセンタイル値より大きい領域では、テスト密度が 25 パーセンタイル値以下の領域と比較して、テスト検出能率の中央値が約 1/6.0 倍に低下している。

改良開発の場合、テスト密度が 75 パーセンタイル値（約 147.9 ケース/KSLOC）より大きい領域では、テスト密度が 25 パーセンタイル値（約 24.9 ケース/KSLOC）以下の領域と比較して、テスト検出不具合密度の中央値が約 4.2 倍大きくなっている。

ただし、テスト密度が高くなるに連れて、テスト検出能率が大幅に低下する傾向が見られる。具体的には、改良開発の場合、テスト密度が 75 パーセンタイル値より大きい領域では、テスト密度が 25 パーセンタイル値以下の領域と比較して、テスト検出能率の中央値が約 1/7.2 倍に低下している。

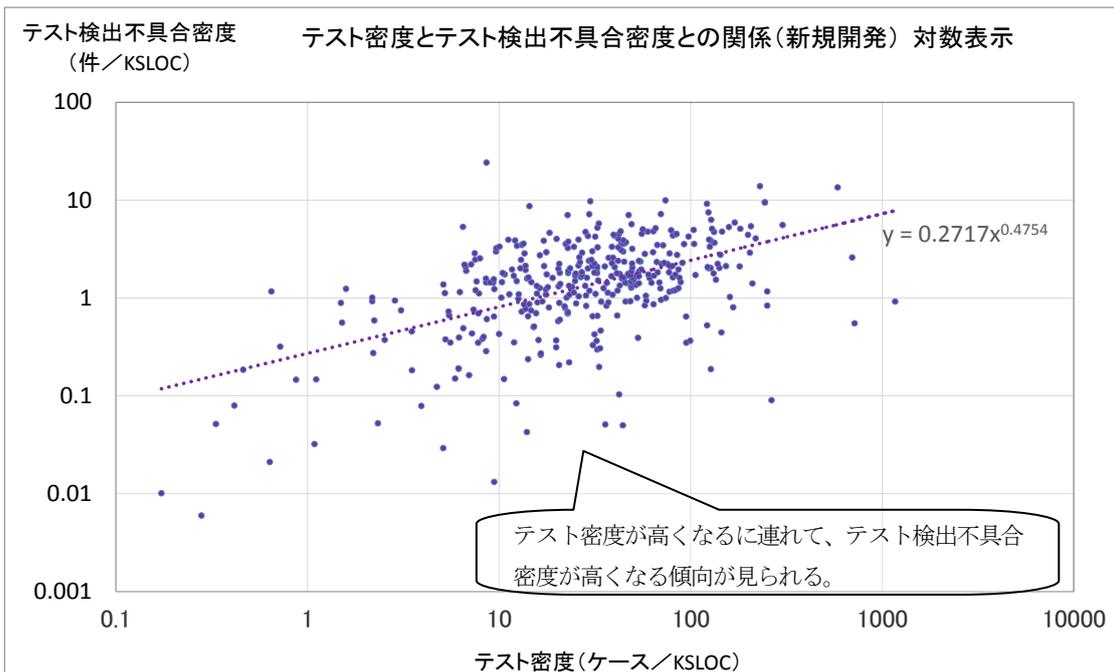
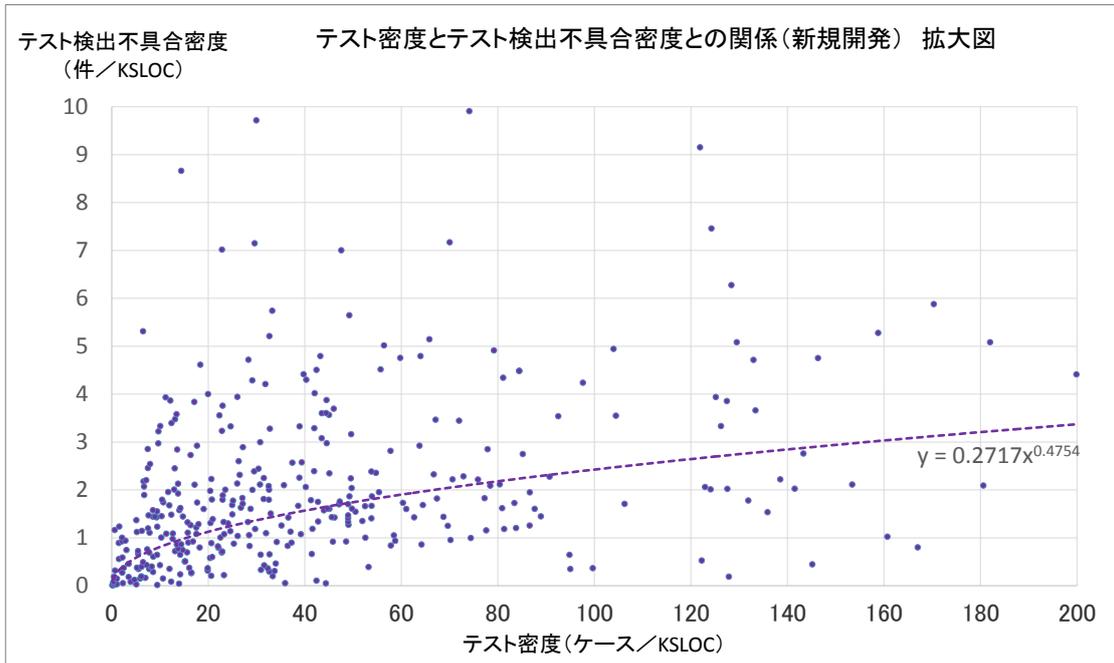
<新規開発の場合>

◇箱ひげ図（テスト密度とテスト検出不具合密度との関係）

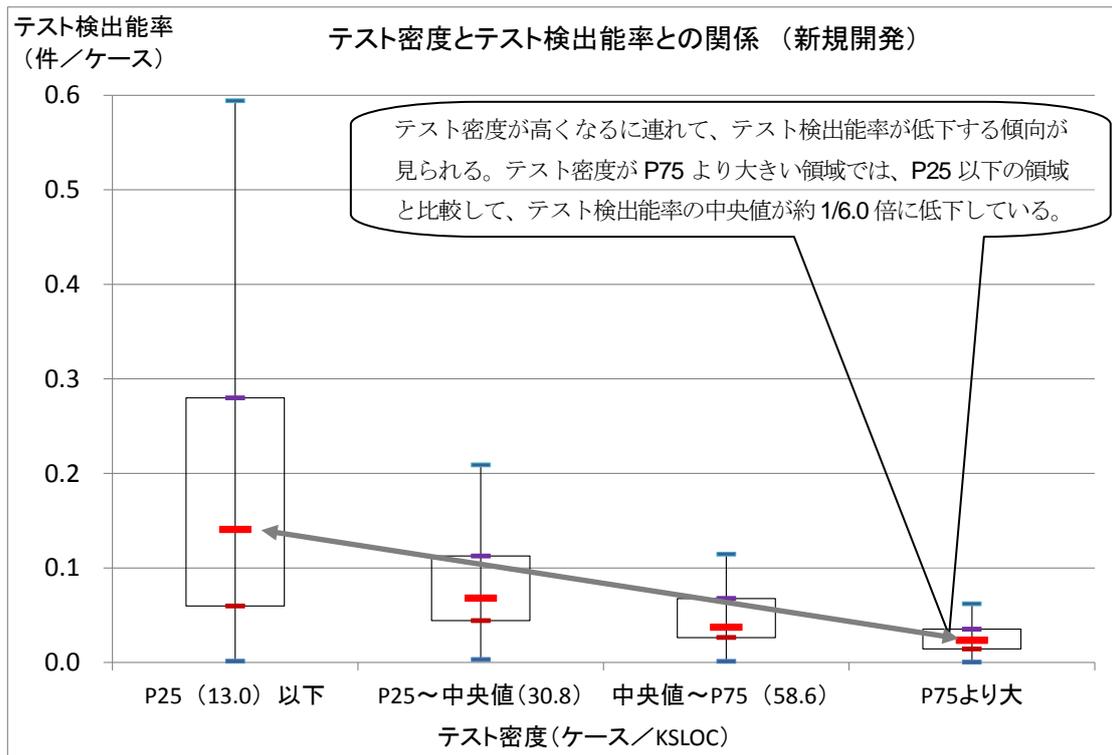


テスト密度	N	テスト検出不具合密度 (件/KSLOC)						
		最小	P25	中央	P75	最大	平均	標準偏差
P25 (13.0) 以下	97	0.006	0.285	0.758	1.481	24.135	1.331	2.661
P25~中央値 (30.8)	97	0.042	0.849	1.467	2.332	9.714	1.923	1.706
中央値~P75 (58.6)	97	0.050	1.112	1.700	2.894	7.000	2.090	1.437
P75より大	97	0.090	1.429	2.218	4.465	13.846	3.170	2.670

◇散布図 (テスト密度とテスト検出不具合密度との関係)

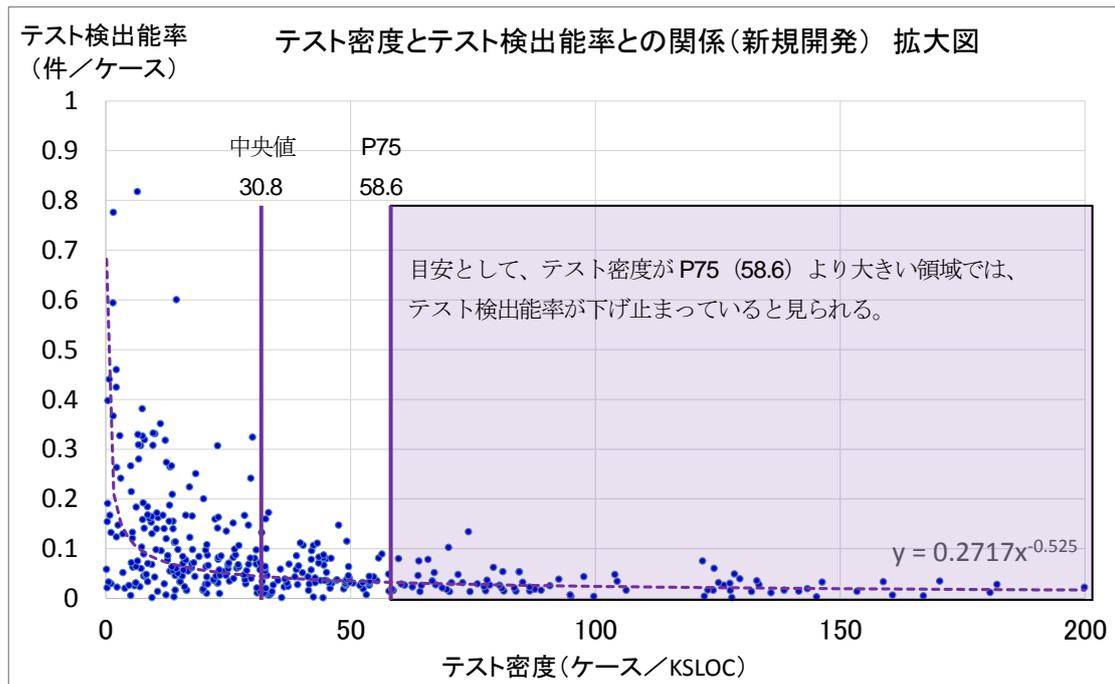


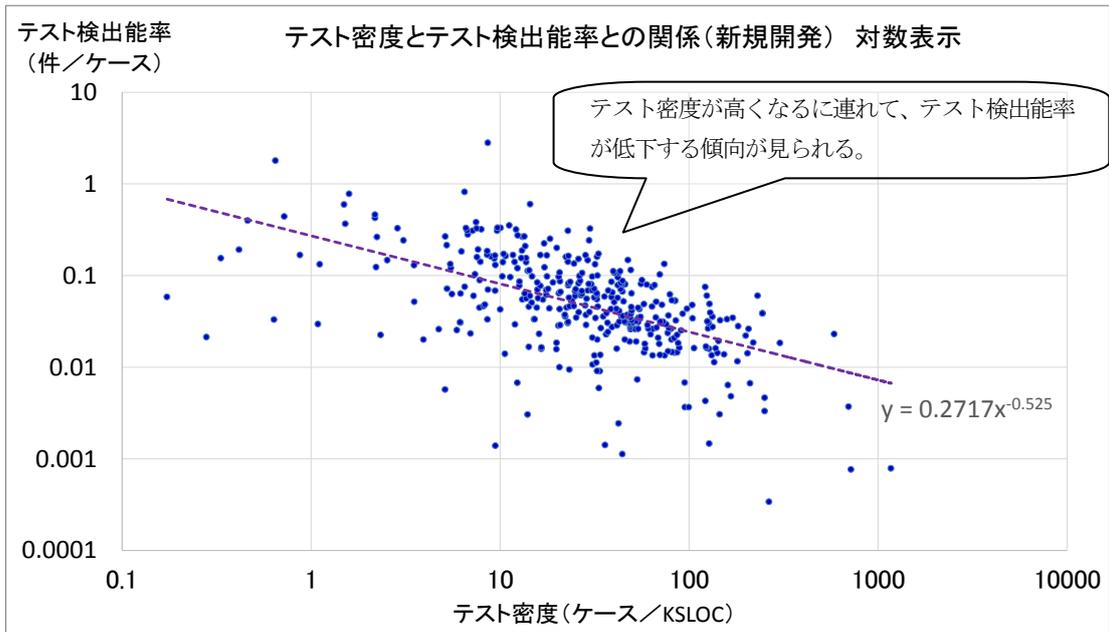
◇箱ひげ図 (テスト密度とテスト検出能率との関係)



テスト密度	N	テスト検出能率 (件/ケース)						
		最小	P25	中央	P75	最大	平均	標準偏差
P25 (13.0) 以下	97	0.0014	0.0597	0.1406	0.2798	2.8013	0.2246	0.3613
P25~中央値 (30.8)	97	0.0030	0.0443	0.0680	0.1126	0.6000	0.0941	0.0860
中央値~P75 (58.6)	97	0.0011	0.0264	0.0372	0.0677	0.1724	0.0497	0.0357
P75より大	97	0.0003	0.0143	0.0233	0.0353	0.1336	0.0283	0.0227

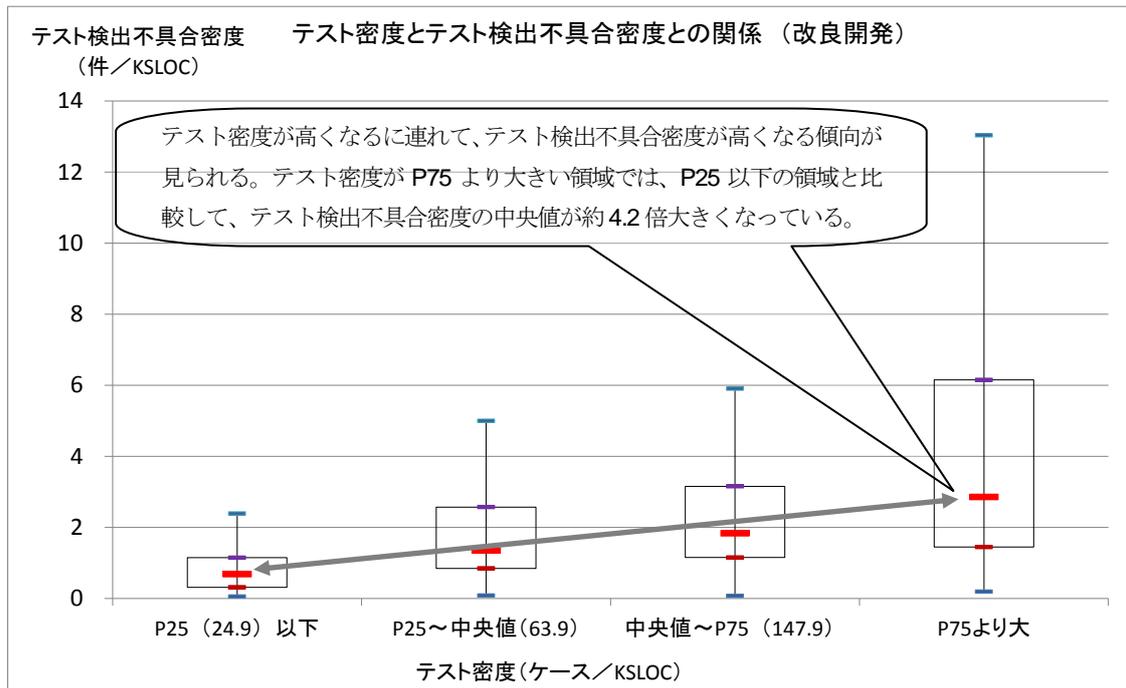
◇散布図 (テスト密度とテスト検出能率との関係)





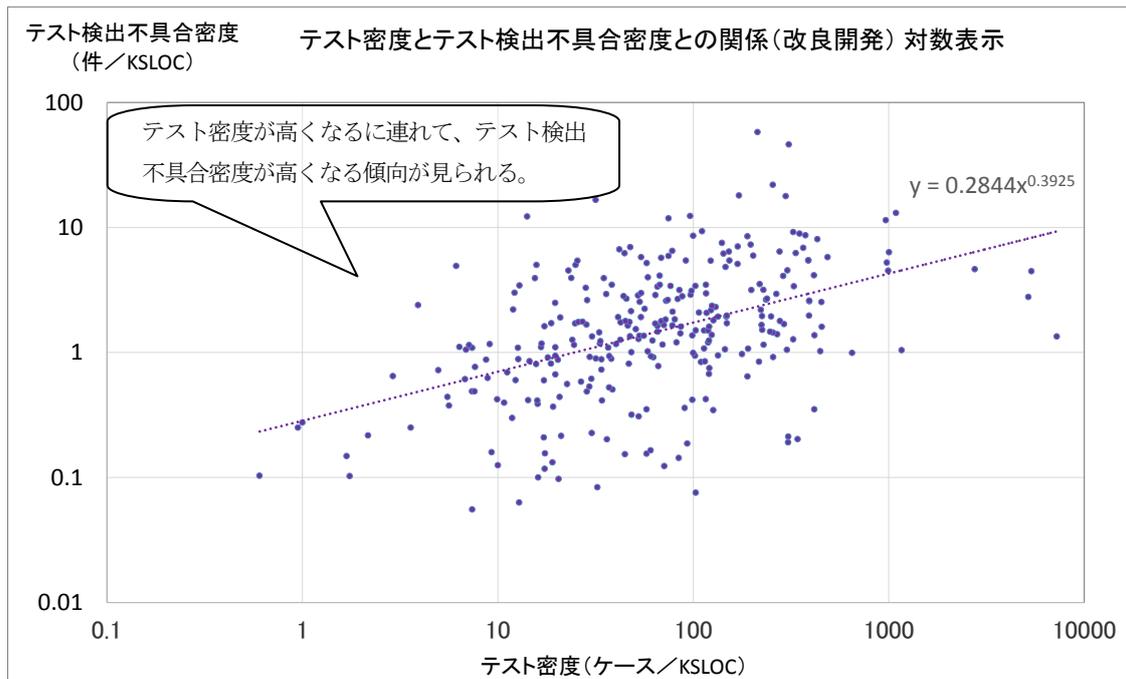
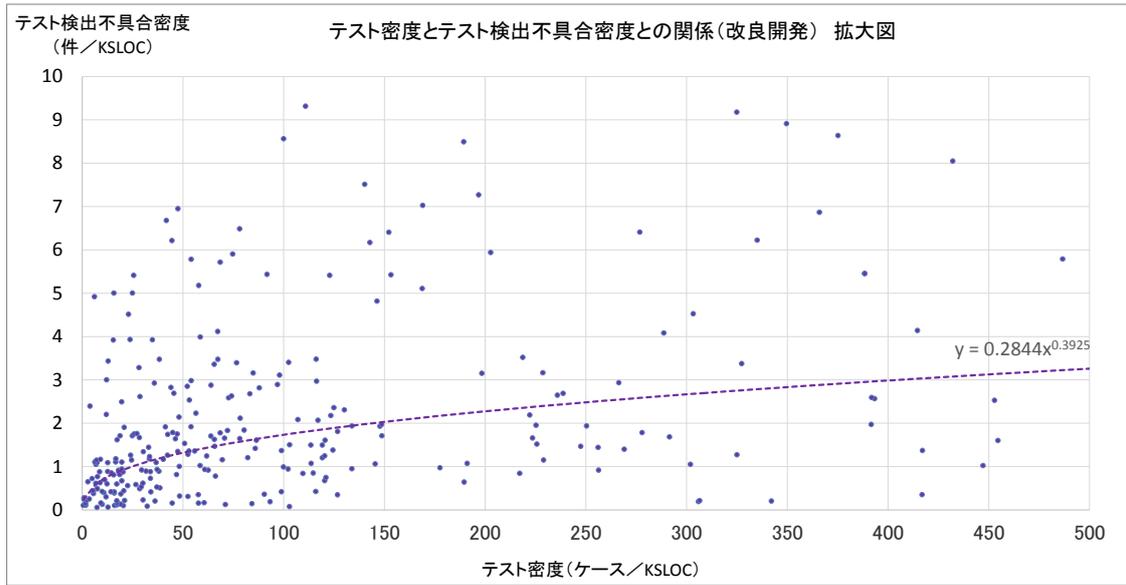
<改良開発の場合>

◇箱ひげ図 (テスト密度とテスト検出不具合密度との関係)

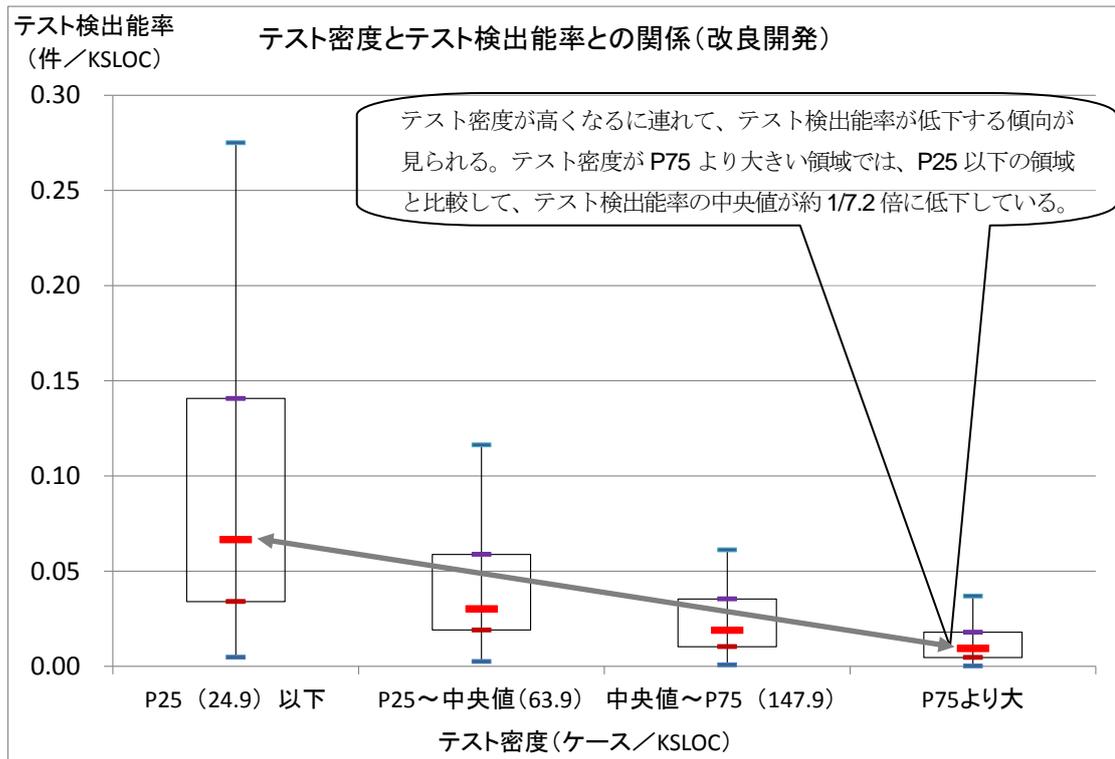


テスト密度	N	テスト検出不具合密度 (件/KSLOC)						
		最小	P25	中央	P75	最大	平均	標準偏差
P25 (24.9) 以下	77	0.06	0.32	0.68	1.15	12.22	1.20	1.74
P25~中央値 (63.9)	76	0.08	0.85	1.35	2.57	16.59	2.06	2.40
中央値~P75 (147.9)	76	0.08	1.15	1.83	3.16	12.35	2.67	2.53
P75より大	76	0.19	1.45	2.85	6.15	57.83	5.56	8.89

◇散布図 (テスト密度とテスト検出不具合密度との関係)

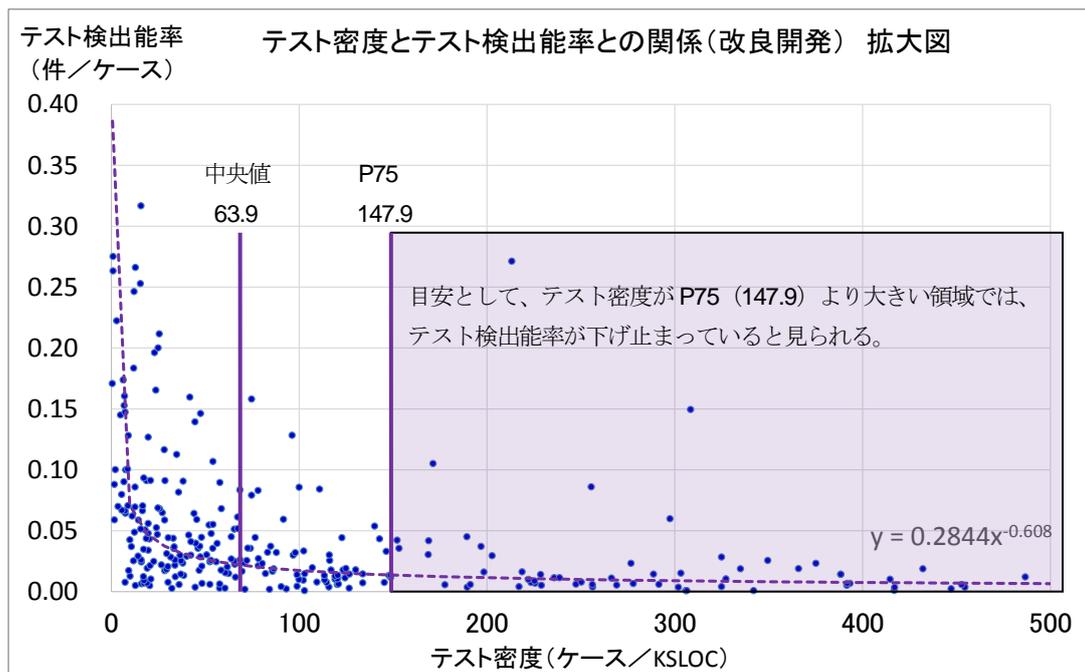


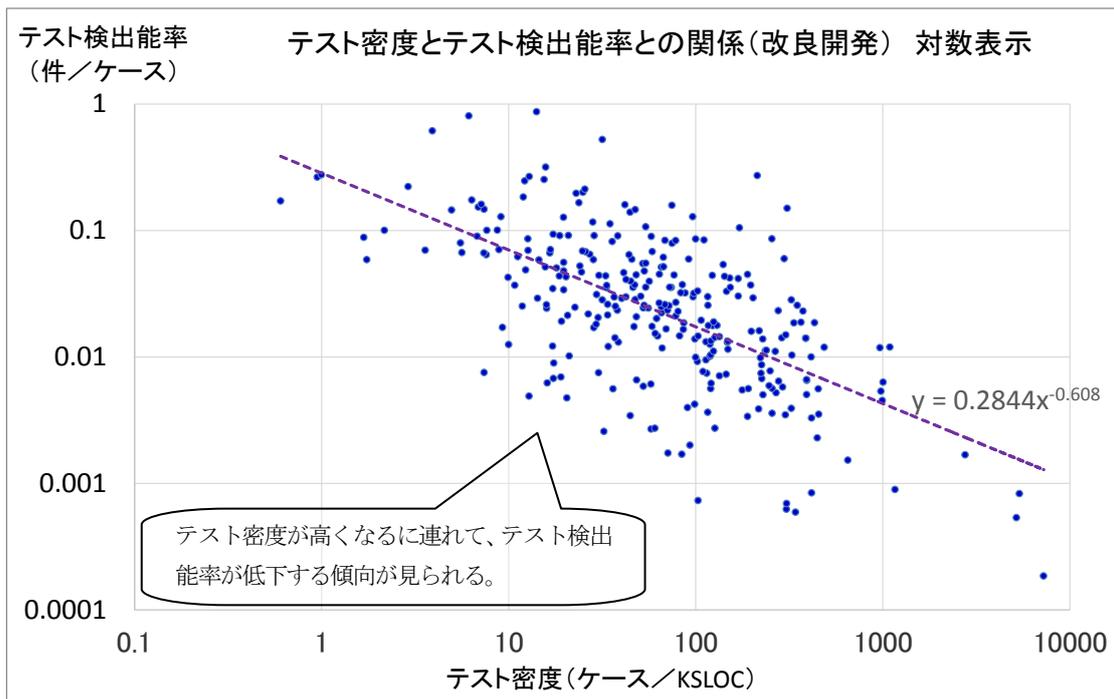
◇箱ひげ図 (テスト密度とテスト検出能率との関係)



テスト密度	N	最小	P25	中央	P75	最大	平均	標準偏差
P25 (24.9) 以下	77	0.0047	0.0340	0.0665	0.1407	0.8652	0.1140	0.1547
P25~中央値 (63.9)	76	0.0026	0.0191	0.0301	0.0588	0.5232	0.0529	0.0719
中央値~P75 (147.9)	76	0.0007	0.0103	0.0189	0.0354	0.1579	0.0288	0.0289
P75より大	76	0.0002	0.0046	0.0092	0.0179	0.2711	0.0198	0.0381

◇散布図 (テスト密度とテスト検出能率との関係)





2.3. 上流工程での不具合摘出比率のコントロール

(1) 信頼性確保のための、上流工程での不具合摘出比率のコントロール

【概要】

分析項目	上流工程での不具合摘出比率と発生不具合密度との関係
メッセージ	<p>◇分析結果概要（読み取れる傾向）</p> <p>新規開発の場合、上流工程での不具合摘出比率が相対的に低い（中央値約 73.7% 以下の）集合と高い（中央値より大きい）集合とでは、後者の方が発生不具合密度が低い（相対的に信頼性が高い）傾向が見られる。具体的には、新規開発の場合、前者の発生不具合密度の中央値が約 0.05 件/KSLOC であるのに対して、後者の発生不具合密度の中央値は 0 件/KSLOC である。また、発生不具合密度の 75 パーセンタイル値については、約 4.4 倍の開きが見られる。</p> <p>（なお、改良開発の場合には、このような顕著な傾向は見られなかった。）</p> <p>◇助言</p> <p>上流工程での不具合摘出比率を高めることによって、信頼性向上を図ることができる。上流工程での不具合摘出比率が向上するように、設計レビューを強化することをお勧めする。</p> <p>◇参考文献</p> <p>上流工程での不具合摘出比率向上に向けて、ソフトウェア開発に係るプロセスを改良する方策を検討する場合、次の IPA/SEC の文献を参考にして頂けると幸いである。</p> <p>○「プロセス改善ナビゲーションガイド ～虎の巻編～」</p> <p>http://www.ipa.go.jp/sec/publish/trn08-009.html</p>
メトリクス	<p>◇上流工程での不具合摘出比率：基本設計から製作までのレビュー指摘数÷（基本設計から製作までのレビュー指摘数+結合テストから総合テスト（ベンダ確認）までの検出不具合数）。</p> <p>◇発生不具合密度：稼働後（最長 6 ヶ月以内）に発生した不具合数÷開発規模（KSLOC）。</p>
考察/備考	<p>上流工程での不具合摘出比率を向上させるような設計レビューの強化策としては、設計レビュー工数密度の向上等のレビュー工数面だけでなく、経済的に（効率的に）レビューを遂行するための次のような方策も併せて検討することが望ましいと考える。</p> <ul style="list-style-type: none"> ・設計関連標準類の改良/整備 設計書作成標準、レビューチェックリスト、解析ツール、レビュー手法等 ・設計文書量の増強 設計文書化密度（設計書ページ数÷開発規模）を高めることによって、設計レビューで見つかる不具合数を増やせる。白書等の分析結果からは、目安として約 16 ページ/KSLOC 以上とすることが望ましい（新規開発の場合）。

【検索条件】

- ◇ 開発 **5** 工程が揃っているプロジェクト
- ◇ 開発プロジェクトの種別が新規開発又は改良開発（改修・保守又は拡張）
- ◇ 開発規模（SLOC 規模） > 0

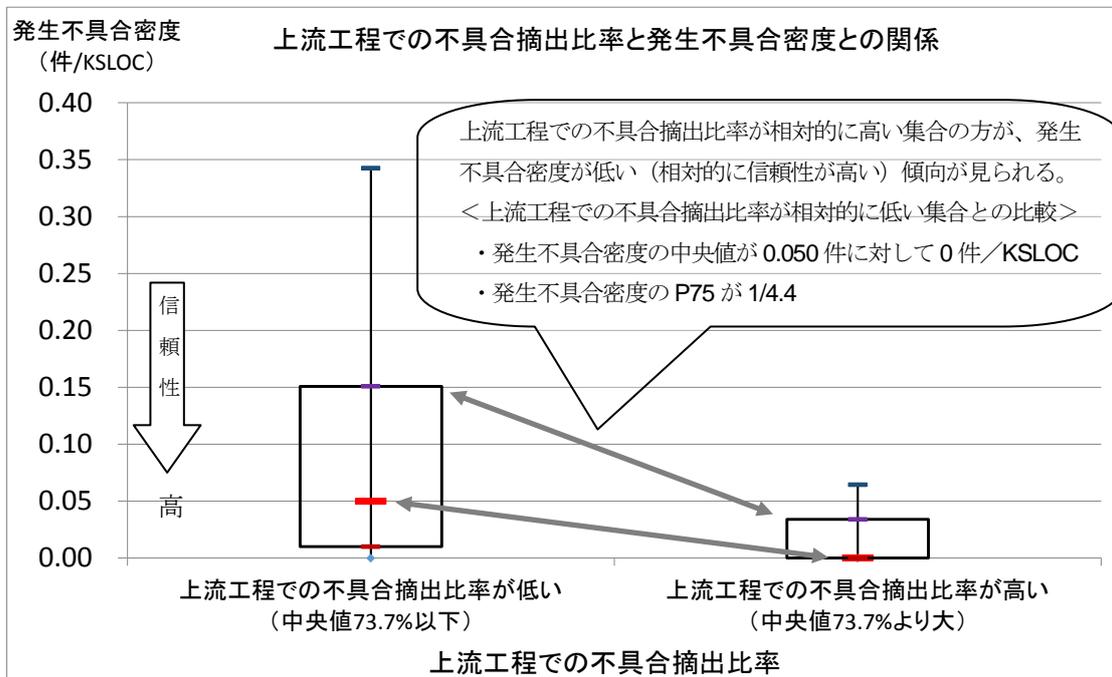
- ◇ 主開発言語が COBOL, C, VB, Java のいずれか
- ◇ 開発 5 工程の総工数 > 0
- ◇ 稼働後の発生不具合数 >= 0
- ◇ 基本設計から製作までのレビュー指摘数 > 0
- ◇ 結合テスト検出不具合数+総合テスト（ベンダ確認）検出不具合数 > 0

【分析結果】

新規開発の場合、上流工程での不具合摘出比率が相対的に低い（中央値約 73.7%以下の）集合と高い（中央値より大きい）集合とでは、後者の方が発生不具合密度が低い（相対的に信頼性が高い）傾向が見られる。具体的には、新規開発の場合、前者の発生不具合密度の中央値が約 0.05 件/KSLOC であるのに対して、後者の発生不具合密度の中央値は 0 件/KSLOC である。また、発生不具合密度の 75 パーセンタイル値については、約 4.4 倍の開きが見られる。（対数化後のデータを用いた Welch の t 検定によれば、10%有意で差が認められる。）（なお、改良開発の場合には、このような顕著な傾向は見られなかった。）

<新規開発の場合>

◇箱ひげ図（上流工程での不具合摘出比率と発生不具合密度との関係）



上流工程での不具合摘出比率	N	発生不具合密度 (件/KSLOC)						
		最小	P25	中央	P75	最大	平均	標準偏差
上流工程での不具合摘出比率が低い (中央値73.7%以下)	47	0.000	0.010	0.050	0.151	1.093	0.149	0.227
上流工程での不具合摘出比率が高い (中央値73.7%より大)	46	0.000	0.000	0.000	0.034	1.117	0.077	0.222

<改良開発の場合>

顕著な傾向が見られなかったので、割愛する。

2.4. 成果物量のコントロールと見積り妥当性評価

(1) 成果物量に着目して工程ごとに評価する方法

【概要】

分析項目	開発規模と各工程の成果物量との相関 各工程の成果物量と各工程の工数との相関
メッセージ	<p>◇分析結果概要（読み取れる傾向）</p> <p>新規開発の場合、開発規模（FP 規模又は SLOC 規模）と各工程の成果物量との間、及び各工程の成果物量と各工程の工数との間には、強い（又は中程度の）正相関が見られる。また、分析対象プロジェクトを特定の業種、例えば金融・保険業のプロジェクトに絞ると、より強い相関が見られる。</p> <p>◇助言</p> <p>実現性の高い開発計画策定及び見積りの妥当性評価に向けて、次のような「成果物量に着目して工程ごとに評価する方法」を補完的に採用することをお勧めする。</p> <ul style="list-style-type: none"> ●開発規模当りの成果物量が妥当か（一定の範囲にあるか） ●成果物量当りの工数が妥当か（一定の範囲にあるか） <p>ここで、本書の次の基本統計量を目安として参考にして頂ければ幸いである。</p> <ul style="list-style-type: none"> ・各工程の開発規模当りの成果物量（各工程の成果物量÷開発規模） ・各工程の成果物量当りの工数（各工程の工数÷各工程の成果物量） <p>ただし、一定の範囲に収まっていないというだけで妥当でないとは評価するのは早計である。自組織の生産性変動要因による変動を勘案しながら評価することが望ましい。どのようなプロジェクトなのかによっては、変動要因による変動を始めとして一定の範囲に収まらなくなる合理的な理由が存在する可能性がある。評価対象プロジェクトに該当する変動要因によって生じる変動幅を勘案して、一定の妥当な範囲を上方修正／下方修正しながら妥当性評価することが望ましい。その結果においても妥当な範囲外となり、かつ変動要因以外の合理的理由がない場合には、見積りを見直すことが望ましい。</p>
考察／備考	<p>◇工程別の成果物量に関する相関は、IPA/SEC のソフトウェア開発プロジェクトデータベースにおいて業種を始め様々なプロファイルのプロジェクトデータが混在した中でも見られる傾向であることから、各組織において同種のソフトウェア開発を行うドメインの中では、より強い相関が見られると考えられる。</p> <p>◇「成果物量に着目して工程ごとに評価する方法」のメリット</p> <p>成果物量には評価対象プロジェクトの範囲、品質要求、制約条件等がある程度反映されるので、評価精度が高まると考えられる。（一方、開発規模から求めた総開発工数によって妥当性評価する方法では、これらを反映できない。）</p> <p>[例 1] 当プロジェクトでは品質要求が高いため、プロセス目標のうち、設計文書化密度（基本設計工程及び詳細設計工程の設計書ページ数÷開発規模）及びテスト密度（結合テスト工程及び総合テスト（ベンダ確認）工程のテストケース数÷開発規模）を高く設定する。</p> <p>[例 2] 当プロジェクトでは、関連システムとの接続テストを強化するので、総合テスト（ベンダ確認）工程のテスト密度（テストケース数÷開発規模）を高く設定する。</p>

	また、工程ごとの成果物量及び工数を吟味することになるので、開発計画の実現可能性を高めることに繋がると考えられる。
--	--

【検索条件等】

- ◇ 開発プロジェクトの種別が新規開発又は改良開発（改修・保守又は拡張）
- ◇ 開発規模（FP 規模又は SLOC 規模） > 0
- ◇ 分析対象開発工程の成果物量 > 0
- ◇ 分析対象開発工程の開発工数 > 0
- ◇ 開発規模が FP 規模の場合、FP 計測手法が IFPUG, SPR, NESMA のいずれか
- ◇ 開発規模が SLOC 規模の場合、主開発言語が COBOL, C, VB, Java のいずれか
- ◇ 常用対数化したデータを用いて相関係数を求める。
- ◇ 各開発工程の成果物量

開発工程	成果物量（実績）
要件定義	要件定義書ページ数
基本設計	基本設計書ページ数
詳細設計	詳細設計書ページ数
製作	コード行数（SLOC）
結合テスト	結合テストケース数
総合テスト（ベンダ確認）	総合テスト（ベンダ確認）ケース数

【分析結果】

<相関係数>

新規開発の場合、開発規模（FP 規模又は SLOC 規模）と各工程の成果物量との間、及び各工程の成果物量と各工程の工数との間には、強い（又は中程度の）正相関が見られる。また、分析対象プロジェクトを特定の業種、例えば金融・保険業のプロジェクトに絞ると、より強い相関が見られる。

（なお、改良開発の場合には、新規開発の場合より弱い相関に留まっており、分析結果は割愛する。）

以下の表に、新規開発の場合の上記の相関係数を示す。0.7 以上の相関係数は、黄色の網掛けで示す。

◇ FP 規模の場合（対数変換後）

開発工程	要件定義	基本設計	詳細設計	製作	結合テスト	総合テスト（ベンダ確認）
データ数	86	127	93	184	192	229
FP規模と各工程の成果物量との相関	0.56	0.68	0.72	0.77	0.64	0.43
各工程の成果物量と各工程の工数との相関	0.67	0.81	0.80	0.79	0.65	0.42

◇ FP 規模、金融・保険業の場合（対数変換後）

開発工程	要件定義	基本設計	詳細設計	製作	結合テスト	総合テスト（ベンダ確認）
データ数	31	49	26	63	51	67
FP規模と各工程の成果物量との相関	0.62	0.75	0.81	0.84	0.68	0.43
各工程の成果物量と各工程の工数との相関	0.63	0.85	0.86	0.77	0.70	0.27

◇ SLOC 規模の場合（対数変換後）

開発工程	要件定義	基本設計	詳細設計	製作	結合テスト	総合テスト (ベンダ確認)
データ数	74	137	131	535	324	328
SLOC規模と各工程の成果物量との相関	0.69	0.68	0.67		0.60	0.47
各工程の成果物量と各工程の工数との相関	0.70	0.71	0.78	0.75	0.63	0.44

◇ SLOC 規模、金融・保険業の場合（対数変換後）

開発工程	要件定義	基本設計	詳細設計	製作	結合テスト	総合テスト (ベンダ確認)
データ数	30	54	49	160	104	113
SLOC規模と各工程の成果物量との相関	0.80	0.73	0.67		0.61	0.44
各工程の成果物量と各工程の工数との相関	0.80	0.76	0.82	0.76	0.68	0.40

<各工程の成果物量に関する基本統計量>

新規開発における次の基本統計量を以下に示す。

- ・各工程の開発規模当りの成果物量（各工程の成果物量÷開発規模）
- ・各工程の成果物量当りの工数（各工程の工数÷各工程の成果物量）

◇ 中央値一覧（FP 規模、新規開発）

開発工程	要件定義	基本設計	詳細設計	製作	結合テスト	総合テスト (ベンダ確認)
データ数	86	127	93	184	192	229
開発規模当りの成果物量の中央値	FP当りの要件定義書ページ数 (ページ/FP)	FP当りの基本設計書ページ数 (ページ/FP)	FP当りの詳細設計書ページ数 (ページ/FP)	FP当りのKSLOC実績値 (KSLOC/FP)	FP当りの結合テストケース数 (ケース/FP)	FP当りの総合テストケース数 (ケース/FP)
	0.118	0.558	1.050	0.075	1.830	0.663
成果物量当りの工数の中央値	要件定義書ページ当りの要件定義工数 (人時/ページ)	基本設計書ページ当りの基本設計工数 (人時/ページ)	詳細設計書ページ当りの詳細設計工数 (人時/ページ)	KSLOC当りの製作工数 (人時/KSLOC)	結合テストケース当りの結合テスト工数 (人時/ケース)	総合テストケース当りの総合テスト工数 (人時/ケース)
	9.43	4.46	2.37	71.30	1.15	3.00

◇ 中央値一覧（SLOC 規模、新規開発）

開発工程	要件定義	基本設計	詳細設計	製作	結合テスト	総合テスト (ベンダ確認)
データ数	74	137	131	535	324	328
開発規模当りの成果物量の中央値	KSLOC当りの要件定義書ページ数 (ページ/KSLOC)	KSLOC当りの基本設計書ページ数 (ページ/KSLOC)	KSLOC当りの詳細設計書ページ数 (ページ/KSLOC)		KSLOC当りの結合テストケース数 (ケース/KSLOC)	KSLOC当りの総合テストケース数 (ケース/KSLOC)
	1.22	6.39	12.10		32.20	9.00
成果物量当りの工数の中央値	要件定義書ページ当りの要件定義工数 (人時/ページ)	基本設計書ページ当りの基本設計工数 (人時/ページ)	詳細設計書ページ当りの詳細設計工数 (人時/ページ)	KSLOC当りの製作工数 (人時/KSLOC)	結合テストケース当りの結合テスト工数 (人時/ケース)	総合テストケース当りの総合テスト工数 (人時/ケース)
	10.90	4.57	2.49	52.20	1.04	2.39

◇ 基本統計量一覧 (FP 規模、新規開発)

• FP 当りの要件定義書ページ数

							(ページ/FP)	
N	最小	P25	中央	P75	最大	平均	標準偏差	
86	0.001	0.057	0.118	0.187	1.104	0.162	0.173	

• FP 当りの基本設計書ページ数

							(ページ/FP)	
N	最小	P25	中央	P75	最大	平均	標準偏差	
127	0.011	0.274	0.558	1.008	3.391	0.713	0.601	

• FP 当りの詳細設計書ページ数

							(ページ/FP)	
N	最小	P25	中央	P75	最大	平均	標準偏差	
93	0.05	0.50	1.05	1.74	65.38	2.27	6.94	

• FP 当りのコード行数 (KSLOC 実績値)

							(KSLOC/FP)	
N	最小	P25	中央	P75	最大	平均	標準偏差	
184	0.001	0.046	0.075	0.137	0.990	0.112	0.120	

• FP 当りの結合テストケース数

							(ケース/FP)	
N	最小	P25	中央	P75	最大	平均	標準偏差	
192	0.02	0.71	1.83	3.10	54.48	2.85	4.95	

• FP 当りの総合テスト (ベンダ確認) ケース数

							(ケース/FP)	
N	最小	P25	中央	P75	最大	平均	標準偏差	
229	0.011	0.287	0.663	1.750	93.914	2.503	8.527	

• 要件定義書ページ当りの要件定義工数

							(人時/ページ)	
N	最小	P25	中央	P75	最大	平均	標準偏差	
86	1.44	5.01	9.43	23.77	588.00	25.31	66.13	

• 基本設計書ページ当りの基本設計工数

							(人時/ページ)	
N	最小	P25	中央	P75	最大	平均	標準偏差	
127	0.54	2.17	4.46	8.82	54.77	6.91	7.73	

• 詳細設計書ページ当りの詳細設計工数

							(人時/ページ)	
N	最小	P25	中央	P75	最大	平均	標準偏差	
93	0.12	1.33	2.37	3.73	20.36	3.40	3.62	

• KSLOC 当りの製作工数

							(人時/KSLOC)
N	最小	P25	中央	P75	最大	平均	標準偏差
184	0.7	39.3	71.3	110.9	854.1	92.8	96.8

• 結合テストケース当りの結合テスト工数

							(人時/ケース)
N	最小	P25	中央	P75	最大	平均	標準偏差
192	0.04	0.46	1.15	2.64	106.72	3.30	9.22

• 総合テスト（ベンダ確認）ケース当りの総合テスト（ベンダ確認）工数

							(人時/ケース)
N	最小	P25	中央	P75	最大	平均	標準偏差
229	0.00	0.80	3.00	8.01	130.94	7.87	13.89

◇ 基本統計量一覧（SLOC 規模、新規開発）

• KSLOC 当りの要件定義書ページ数

							(ページ/KSLOC)
N	最小	P25	中央	P75	最大	平均	標準偏差
74	0.02	0.57	1.22	2.78	21.41	2.37	3.18

• KSLOC 当りの基本設計書ページ数

							(ページ/KSLOC)
N	最小	P25	中央	P75	最大	平均	標準偏差
137	0.28	2.98	6.39	12.24	384.36	12.27	33.86

• KSLOC 当りの詳細設計書ページ数

							(ページ/KSLOC)
N	最小	P25	中央	P75	最大	平均	標準偏差
131	0.4	6.3	12.1	23.5	1143.6	32.5	112.9

• KSLOC 当りの結合テストケース数

							(ケース/KSLOC)
N	最小	P25	中央	P75	最大	平均	標準偏差
324	0.1	13.6	32.2	54.0	2115.3	60.7	162.9

• KSLOC 当りの総合テスト（ベンダ確認）ケース数

							(ケース/KSLOC)
N	最小	P25	中央	P75	最大	平均	標準偏差
328	0.02	2.81	9.00	22.26	986.62	30.86	82.60

• 要件定義書ページ当りの要件定義工数

							(人時/ページ)
N	最小	P25	中央	P75	最大	平均	標準偏差
74	1.1	4.9	10.9	21.3	588.0	29.5	76.1

- 基本設計書ページ当りの基本設計工数

							(人時/ページ)
N	最小	P25	中央	P75	最大	平均	標準偏差
137	0.37	2.09	4.57	10.33	110.15	9.26	13.97

- 詳細設計書ページ当りの詳細設計工数

							(人時/ページ)
N	最小	P25	中央	P75	最大	平均	標準偏差
131	0.12	1.37	2.49	4.17	82.08	4.40	8.01

- KSLOC 当りの製作工数

							(人時/KSLOC)
N	最小	P25	中央	P75	最大	平均	標準偏差
535	0.7	26.3	52.2	86.7	918.4	74.2	88.3

- 結合テストケース当りの結合テスト工数

							(人時/ケース)
N	最小	P25	中央	P75	最大	平均	標準偏差
324	0.01	0.47	1.04	2.48	106.72	3.11	8.98

- 総合テスト (ベンダ確認) ケース当りの総合テスト (ベンダ確認) 工数

							(人時/ケース)
N	最小	P25	中央	P75	最大	平均	標準偏差
328	0.00	0.76	2.39	7.15	286.32	8.24	20.91

3. マネジメント編

主に開発組織の管理層の方々に向けて、ステークホルダー・マネジメント、スコープ・マネジメント等のプロジェクト・マネジメントへの活用や、組織の成熟度向上（体制の整備、標準類の改良等）への活用に資することを目的として、いくつかの新規分析結果とメッセージを提示する。

3.1. プロジェクト・マネジメント関連

3.1.1. ユーザの関与（ステークホルダー・マネジメント関連）

(1) ユーザ担当者の要求仕様関与度合いの向上

【概要】

分析項目	<p>ユーザ担当者の要求仕様関与と SLOC 生産性との関係（新規開発）</p> <p>ユーザ担当者の要求仕様関与と発生不具合密度（信頼性）との関係（新規開発）</p>
メッセージ	<p>◇分析結果概要（読み取れる傾向）</p> <p>ユーザ担当者の要求仕様関与が多い方が、少ない方と比較して相対的に SLOC 生産性が高く、発生不具合密度がやや低い（相対的に信頼性がやや高い）傾向が見られる。具体的には、両者の SLOC 生産性の中央値には、約 1.7 倍の開きが見られる。また、両者の発生不具合密度の中央値には、約 7.9 倍の開きが見られる（両者の発生不具合密度の 75 パーセンタイル値には、約 2.4 倍の開きが見られる）。</p> <p>◇助言</p> <p>ユーザ担当者の要求仕様関与が多い方が、少ない方と比較して相対的に生産性及び信頼性が良い傾向が見られる。生産性向上及び信頼性向上に向けて、ユーザ担当者の要求仕様関与度合いを高めるようユーザと調整することが望ましい。あるいは、ユーザ担当者の要求仕様関与度合いによる変動を目安として参考にしながら、開発計画を策定することが望ましい。</p> <p>◇参考文献</p> <p>ステークホルダー・マネジメントに関連して、次の IPA/SEC の文献を併せて参考にして頂けると幸いである。</p> <ul style="list-style-type: none"> ○「経営者が参画する要求品質の確保」 http://www.ipa.go.jp/sec/publish/tn05-002.html 要件定義までの超上流工程でのステークホルダーが、共通の認識を持ち、それぞれが果たす役割と責任を明確にして問題解決にあたるためのヒント ○「機能要件の合意形成ガイド」 https://www.ipa.go.jp/sec/softwareengineering/reports/20100331.html 開発者が設計書を記述することのみではなく、発注者と開発者がシステム像を共有し、行き違いなく合意形成を行うための「コツ」
メトリクス	<p>◇ユーザ担当者の要求仕様関与が多いとは「十分に関与又は概ね関与」を意味する。ユーザ担当者の要求仕様関与が少ないとは「関与が不十分又は未関与」を意味する。</p> <p>◇発生不具合密度：稼働後（最長 6 ヶ月以内）に発生した不具合数÷開発規模（KSLOC）。</p>

<p>考察／備考</p>	<p>プロジェクト・マネジメントのステークホルダー・マネジメントにおいては、ステークホルダー（特にユーザ）を効果的に関与させることが、プロジェクトを成功に導く重要な要素となっている。また、スコープ・マネジメントにおいては、成果物スコープ（成果物とその要件）のベースラインを確立しステークホルダー間で合意形成するプロセス（要件定義に相当）が重要な要素となっている。</p> <p>これらのことから、ユーザの関与の中で「ユーザ担当者の要求仕様への関与」が、プロジェクトを成功に導くために特に重要な要素と考えられるので、「ユーザ担当者の要求仕様への関与度合いと生産性及び信頼性との関係」を分析した。</p> <p>ユーザ担当者の要求仕様への関与度合いが高いと、要求仕様が早期に明確になる、要求仕様の変更／紆余曲折が少なくなり手戻りが少なくなる等の効果が期待できるので、生産性、信頼性両面で有利と考えられる。</p> <p>文書化、レビュー、テスト等の品質保証プロセスの工夫・努力によって信頼性向上を図ることは主に開発者自身が工夫・努力すれば良いことであるが、本件のようにユーザの理解と協力を得ることによってより一層の信頼性向上を図ることも重要と考える。</p>
--------------	---

【検索条件】

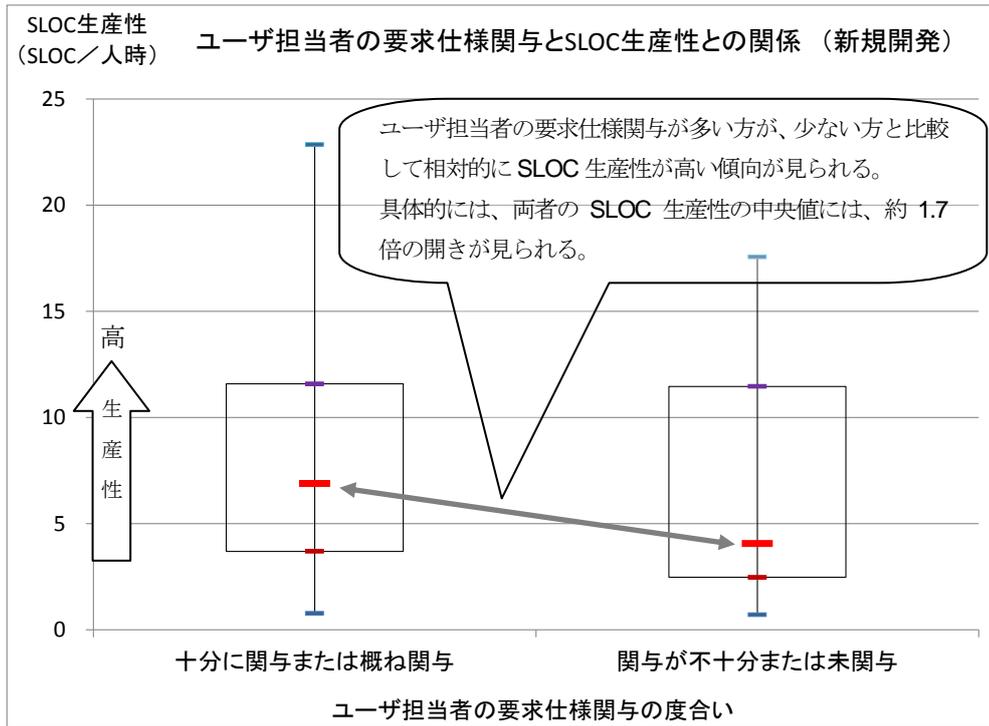
- ◇ 開発 5 工程が揃っているプロジェクト
- ◇ 開発プロジェクトの種別が新規開発
- ◇ 開発規模（SLOC 規模） $\geq 5\text{KSLOC}$
- ◇ 主開発言語が COBOL, C, VB, Java のいずれか
- ◇ 稼働後の発生不具合数 ≥ 0

【分析結果】

ユーザ担当者の要求仕様関与が多い方が、少ない方と比較して相対的に SLOC 生産性が高い傾向が見られる。(Welch の t 検定によれば、1%有意で差が認められる。) また、ユーザ担当者の要求仕様関与が多い方が、少ない方と比較して相対的に発生不具合密度がやや低い (信頼性がやや高い) 傾向が見られる。(Welch の t 検定によれば、ほぼ 20%有意で差が認められる。)

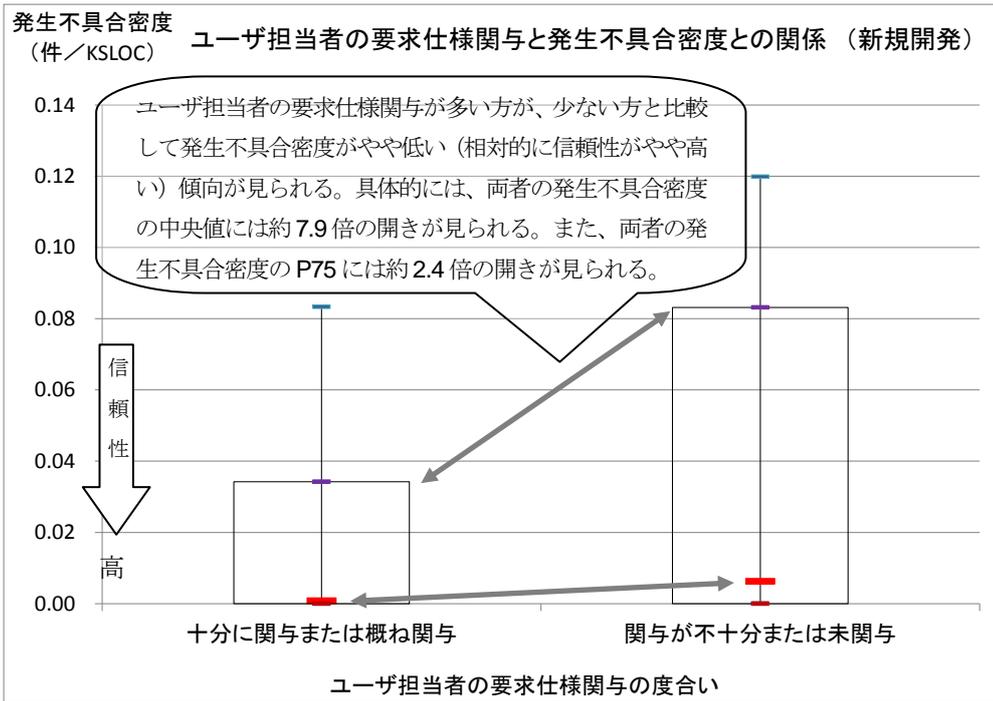
具体的には、両者の SLOC 生産性の中央値には、約 1.7 倍の開きが見られる。また、両者の発生不具合密度の中央値には、約 7.9 倍の開きが見られる。両者の発生不具合密度の 75 パーセンタイル値には、約 2.4 倍の開きが見られる。

箱ひげ図



ユーザ担当者の要求仕様関与	N	最小	P25	中央	P75	最大	SLOC生産性 (SLOC/人時)	
							平均	標準偏差
十分に関与または概ね関与	72	0.78	3.69	6.88	11.58	71.69	12.14	14.01
関与が不十分または未関与	16	0.71	2.47	4.05	11.46	17.56	6.45	5.59

箱ひげ図



ユーザ担当者の要求仕様関与	N	発生不具合密度 (件/KSLOC)						
		最小	P25	中央	P75	最大	平均	標準偏差
十分に関与または概ね関与	72	0.00000	0.00000	0.00079	0.03425	1.01339	0.05020	0.14872
関与が不十分または未関与	16	0.00000	0.00000	0.00623	0.08316	1.11732	0.14801	0.31442

3.1.2. 成果物スコープの明確化（スコープ・マネジメント関連）

(1) 要求仕様の明確化

【概要】

分析項目	要求仕様の明確さと発生不具合密度（信頼性）との関係（新規開発）
メッセージ	<p>◇分析結果概要（読み取れる傾向）</p> <p>要求仕様が明確な方が、あいまいな方と比較して発生不具合密度がやや低い（相対的に信頼性がやや高い）傾向が見られる。</p> <p>具体的には、両者の発生不具合密度の中央値には、約 9.3 倍の開きが見られる（両者の発生不具合密度の 75 パーセンタイル値には、約 1.7 倍の開きが見られる）。</p> <p>◇助言</p> <p>実現性の高い開発計画策定に向けて、要求仕様を早期に明確化するようユーザと調整することが望ましい。あるいは、要求仕様の明確さによる変動を目安として参考にしながら、開発計画を策定することが望ましい。</p> <p>◇参考文献</p> <p>スコープ・マネジメントに関連して、次の IPA/SEC の文献を併せて参考にして頂けると幸いである。</p> <p>○「非機能要求グレード利用ガイド（2010年）」</p> <p>http://www.ipa.go.jp/sec/softwareengineering/reports/20100416.html</p> <p>「非機能要求」について、発注者と受注者との認識の行き違いや異なる理解を防止することを目的とし、重要な項目から段階的に詳細化しながら非機能要求の確認を行うツール群。</p>
メトリクス	<p>◇要求仕様が明確とは「非常に明確又はかなり明確」を意味する。要求仕様があいまいとは「ややあいまい又は非常にあいまい」を意味する。</p> <p>◇発生不具合密度は、稼働後（最長 6 ヶ月以内）に発生した不具合数 ÷ 開発規模（KSLOC）。</p>
考察／備考	<p>要求仕様を早期に明確化すると、要求仕様の変更／紆余曲折が少なくなり手戻りが少なくなる等の効果が期待できるので、生産性、信頼性両面で有利と考えられる。</p> <p>スコープ・マネジメントにおいては、成果物スコープ（成果物とその要件）のベースラインを確立しステークホルダー間で合意形成するプロセス（要件定義に相当）が重要な要素となっている。また、ステークホルダー・マネジメントにおいては、ステークホルダー（特にユーザ）を効果的に関与させることが、プロジェクトを成功に導く重要な要素となっている。</p> <p>要求仕様の早期明確化に向けて、要件定義に注力するとともにユーザの要求仕様への関与度合いを高めることが重要と考えられる。</p>

【検索条件】

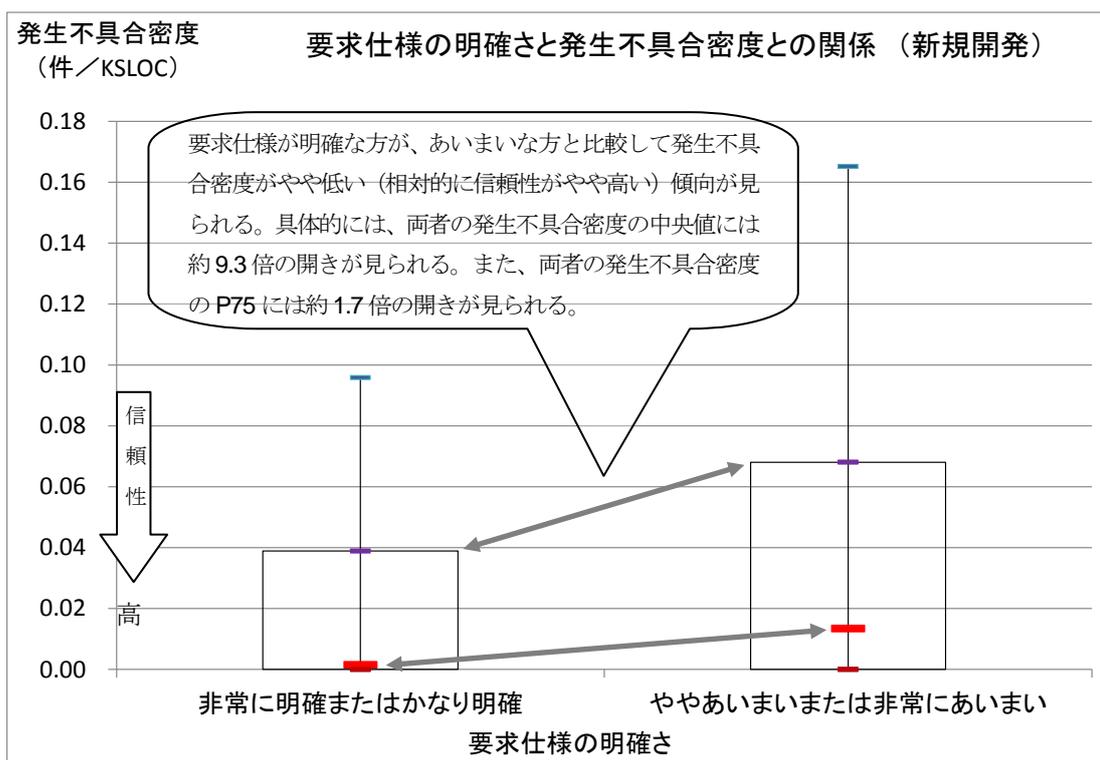
- ◇ 開発 5 工程が揃っているプロジェクト
- ◇ 開発プロジェクトの種別が新規開発
- ◇ 開発規模（SLOC 規模） $\geq 5\text{KSLOC}$
- ◇ 主開発言語が COBOL, C, VB, Java のいずれか
- ◇ 稼働後の発生不具合数 ≥ 0

【分析結果】

要求仕様が明確な方が、あいまいな方と比較して発生不具合密度がやや低い（相対的に信頼性がやや高い）傾向が見られる。（Welch の t 検定による有意性は確認できていないが、箱ひげ図で視覚的に差が見られることから掲載している。）

具体的には、両者の発生不具合密度の中央値には、約 9.3 倍の開きが見られる。また、両者の発生不具合密度の 75 パーセンタイル値には、約 1.7 倍の開きが見られる。

箱ひげ図



要求仕様の明確さ	N	発生不具合密度 (件/KSLOC)						
		最小	P25	中央	P75	最大	平均	標準偏差
非常に明確またはかなり明確	69	0.00000	0.00000	0.00143	0.03889	1.11732	0.06621	0.18886
ややあいまいまたは非常にあいまい	28	0.00000	0.00000	0.01336	0.06804	0.68138	0.08260	0.17582

3.2. 組織の成熟度向上に向けたマネジメント関連

3.2.1. 組織体制の整備

(1) 品質保証体制の強化

【概要】

分析項目	品質保証体制と発生不具合密度との関係（新規開発）
メッセージ	<p>◇分析結果概要（読み取れる傾向）</p> <p>品質保証体制として品質保証の専門スタッフが参加している集合の方が、参加していない集合と比較して発生不具合密度が低い（相対的に信頼性が高い）傾向が見られる。具体的には、両者の発生不具合密度の中央値には、約 15.0 倍の開きが見られる。両者の発生不具合密度の 75 パーセンタイル値には、約 3.5 倍の開きが見られる。</p> <p>◇助言</p> <p>信頼性向上に向けた組織体制整備の一環として、次のことをお勧めする。</p> <ul style="list-style-type: none"> ・より多くの開発プロジェクトに品質保証の専門スタッフが参加できるように、組織体制を強化する。 ・開発要員に対して、ソフトウェアの品質保証及びプロジェクト・マネジメントのスキル向上を図る。
メトリクス	◇発生不具合密度：稼働後（最長 6 ヶ月以内）に発生した不具合数 ÷ 開発規模 (KSLOC)。
考察／備考	<p>品質保証の専門スタッフがプロジェクトに参加することによって、専門スタッフのスキル／知識／経験が、主に次のような場面で、信頼性向上に向けて有効に働くと考えられる。</p> <p>◇リスクの少ない（実現可能性の高い）開発計画の策定支援</p> <p>種々の成功／失敗の事例を知っていることと、プロジェクト・マネジメントのスキルが高いことから、種々のプロジェクト・リスクを予見するとともにリスク対策について助言できる。</p> <p>◇レビュー及びテストに向けた効果的な助言</p> <p>種々の不具合事例に基づいた弱点（陥りやすい誤り、漏れやすい箇所）に関する知識が豊富なので、レビュー及びテストに対して効果的に助言できる。</p> <p>◇精度の高い信頼性予測・評価</p> <p>信頼性予測・評価スキルが高いため、開発時データ（文書化、レビュー、レビュー指摘、テスト、テスト時障害・誤り等のデータ）から、中間成果物及び成果物の信頼性／収束性を精度高く予測・評価できる。</p> <p>◇定量的管理の推進</p> <p>定量的管理のスキルと豊富な経験から、プロジェクトの定量的管理を指導／支援できる。特に値を読み解く力があるので、定量データの分析結果をより適切なマネジメント・アクションに繋げることができる。</p>

【検索条件】

- ◇ 開発 5 工程が揃っているプロジェクト

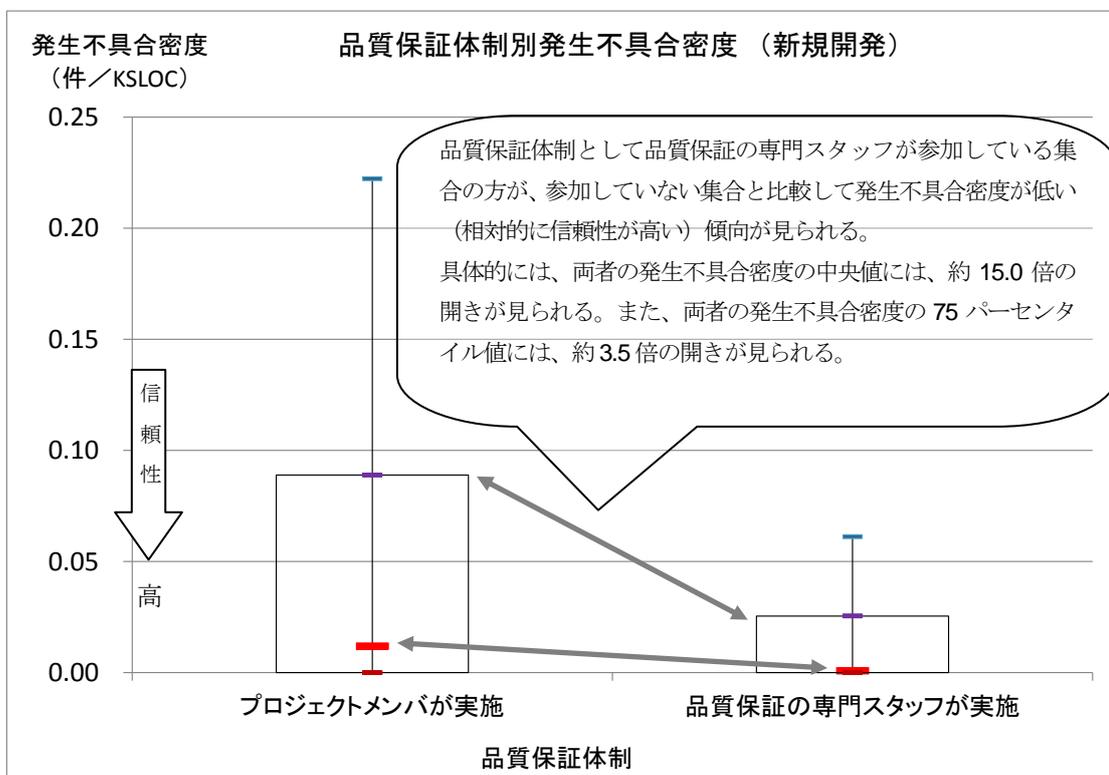
- ◇ 開発プロジェクトの種別が新規開発
- ◇ 開発規模（SLOC 規模） $\geq 5\text{KSLOC}$
- ◇ 主開発言語が COBOL, C, VB, Java のいずれか
- ◇ 稼働後の発生不具合数 ≥ 0

【分析結果】

品質保証体制として品質保証の専門スタッフが参加している集合の方が、参加していない集合と比較して発生不具合密度が低い（相対的に信頼性が高い）傾向が見られる。（Welch の t 検定によれば、5%有意で差が認められる。）

具体的には、両者の発生不具合密度の中央値には、約 15.0 倍の開きが見られる。両者の発生不具合密度の 75 パーセンタイル値には、約 3.5 倍の開きが見られる。

箱ひげ図



品質保証体制	N	発生不具合密度（件/KSLOC）						
		最小	P25	中央	P75	最大	平均	標準偏差
プロジェクトメンバが実施	77	0.0000	0.0000	0.0118	0.0889	0.5816	0.0711	0.1217
品質保証の専門スタッフが実施	72	0.0000	0.0000	0.0008	0.0255	0.4704	0.0336	0.0781

3.2.2. ドメイン別マネジメント

(1) 業種ドメイン別マネジメント

【概要】

分析項目	発生不具合密度（信頼性）の業種間比較 KSLOC 当り工数（生産性）の業種間比較 生産性／信頼性の変動要因候補の業種間比較
メッセージ	◇分析結果概要（読み取れる傾向） 発生不具合密度（信頼性）及び KSLOC 当り工数（生産性）は業種間で差異が見られる。また、信頼性要求レベルや開発プロセス関連の生産性／信頼性の変動要因候補においても業種間の差異が見られる。 特に金融・保険業では、他の業種より信頼性が高く生産性が低い傾向が見られる。また、その要因面では、他の業種より品質保証（設計文書化、レビュー、テスト）に多くの工数をかけている傾向が見られる。 ◇助言 マネジメントや分析を進める上で業種の影響を無視できないと判断できるので、業種によってドメインを分けてマネジメントすることをお勧めする。
メトリクス	◇発生不具合密度：稼働後（最長 6 ヶ月以内）に発生した不具合数÷開発規模（KSLOC） ◇開発プロセス関連の生産性／信頼性の変動要因候補として分析したものは次の通り。 ・設計文書化密度 （基本設計書のページ数+詳細設計書のページ数）÷開発規模（KSLOC） ・設計レビュー工数密度 基本設計から製作までの設計レビュー工数÷開発規模（KSLOC） ・設計レビュー指摘密度 基本設計から製作までのレビュー指摘数÷開発規模（KSLOC） ・テスト密度 （結合テストケース数+総合テスト（ベンダ確認）ケース数）÷開発規模（KSLOC） ・テスト検出不具合密度 （結合テスト検出不具合数+総合テスト（ベンダ確認）検出不具合数）÷開発規模（KSLOC） ・上流工程での不具合摘出比率 基本設計から製作までのレビュー指摘数÷（基本設計から製作までのレビュー指摘数+結合テストから総合テスト（ベンダ確認）までの検出不具合数）
考察／備考	◇特に金融・保険業では、他の業種より信頼性が高く生産性が低い傾向が見られる。また、その要因面では、他の業種より品質保証（設計文書化、レビュー、テスト）に多くの工数をかけている傾向が見られることから、システムリスクが高い（信頼性、公共性及び社会性の要求が高い）ソフトウェアの開発には、それ相応の品質保証（設計文書化、レビュー、テスト等）の工数が必要と考えられる。 ◇信頼性及び生産性は業種間で差異が見られるとともに、生産性／信頼性の変動

要因候補においても業種間の差異が見られることから、変動要因候補の多くが、生産性／信頼性の変動要因になっている可能性が高いと考えられる。

【検索条件（共通分）】

- ◇ 開発 5 工程が揃っているプロジェクト
- ◇ 開発プロジェクトの種別が新規開発又は改良開発（改修・保守又は拡張）
- ◇ 開発規模（SLOC 規模） > 0
- ◇ 主開発言語が COBOL, C, VB, Java のいずれか
- ◇ 実績工数（開発 5 工程） > 0
- ◇ 稼働後の発生不具合数 ≥ 0
- ◇ 業種（大業種）：製造業、情報通信業、卸売・小売業、金融・保険業、公務（他に分類されないもの）

【分析方法】

各業種の信頼性を他の業種と比較するために、ある業種とその業種以外の業種全体との二群に分けて比較する。

【分析結果】

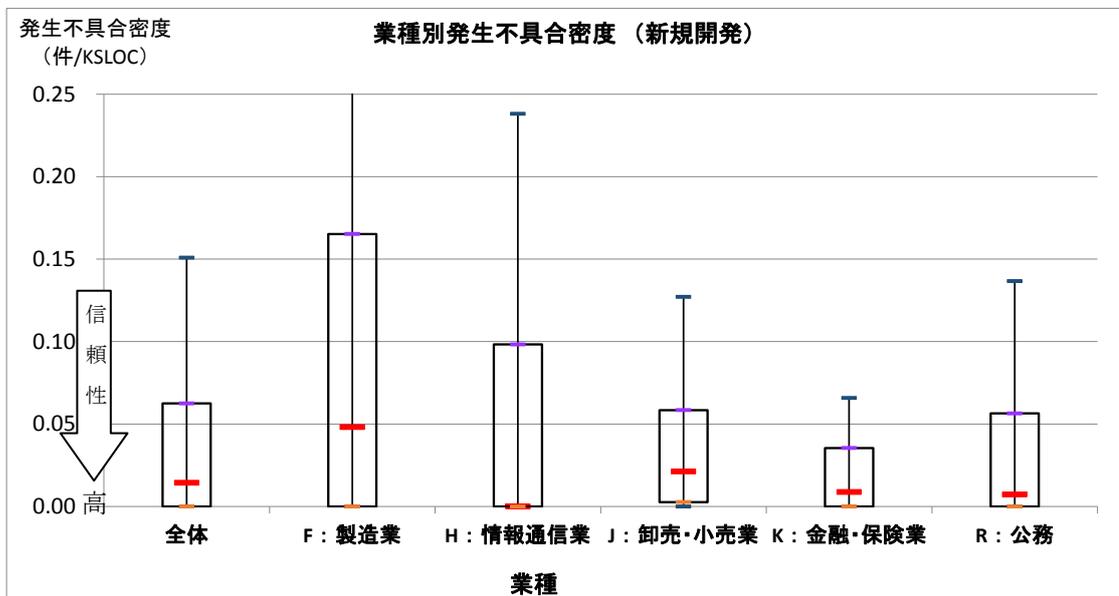
(1) 発生不具合密度（信頼性）の業種間比較

<新規開発の場合>

製造業は他より発生不具合密度が高い傾向が見られる（Welch の t 検定で 10% 有意）。

金融・保険業は他より発生不具合密度が低い傾向が見られる（Welch の t 検定で 1% 有意）。

箱ひげ図

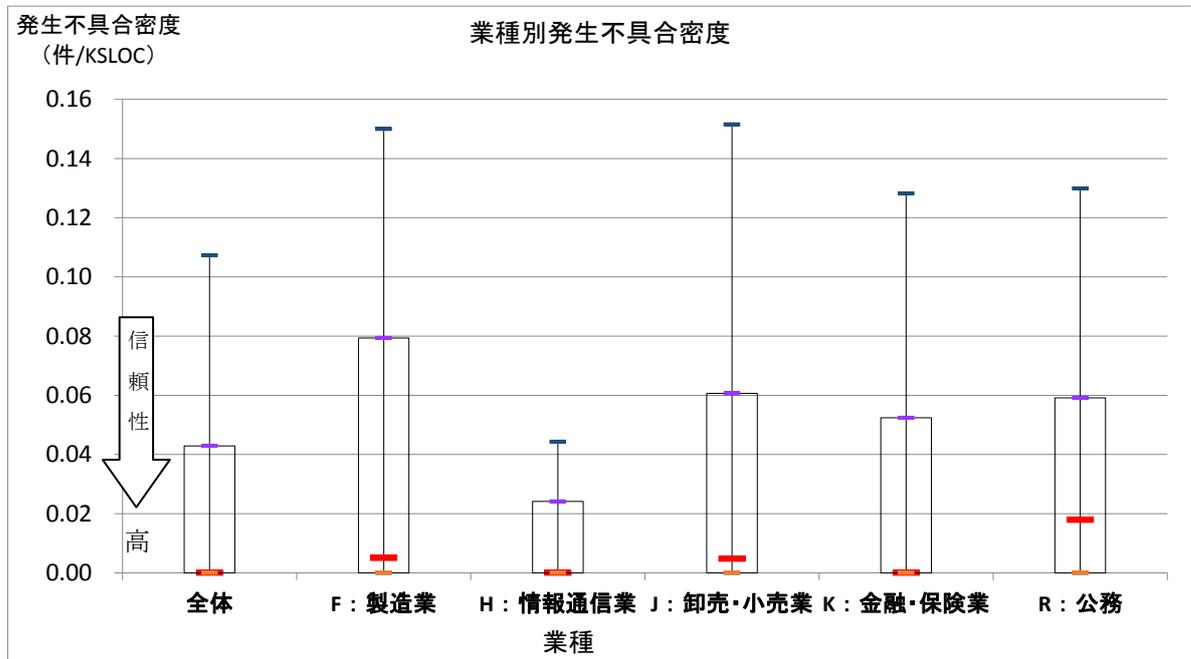


大業種	N	発生不具合密度 (件/KSLOC)						
		最小	P25	中央	P75	最大	平均	標準偏差
全体	373	0.000	0.000	0.014	0.063	2.413	0.081	0.204
F: 製造業	45	0.000	0.000	0.048	0.165	1.013	0.123	0.200
H: 情報通信業	43	0.000	0.000	0.000	0.098	0.483	0.066	0.112
J: 卸売・小売業	34	0.000	0.003	0.021	0.058	2.413	0.127	0.416
K: 金融・保険業	128	0.000	0.000	0.009	0.035	1.117	0.044	0.134
R: 公務 (他に分類されないもの)	51	0.000	0.000	0.007	0.056	1.269	0.114	0.265

<改良開発の場合>

情報通信業は他より発生不具合密度が低い傾向が見られる (Welch の t 検定で 1%有意)。

箱ひげ図



大業種	N	発生不具合密度 (件/KSLOC)						
		最小	P25	中央	P75	最大	平均	標準偏差
全体	361	0.000	0.000	0.000	0.043	5.155	0.093	0.361
F: 製造業	31	0.000	0.000	0.005	0.079	0.738	0.080	0.161
H: 情報通信業	40	0.000	0.000	0.000	0.024	0.171	0.022	0.045
J: 卸売・小売業	18	0.000	0.000	0.005	0.061	1.007	0.128	0.268
K: 金融・保険業	136	0.000	0.000	0.000	0.052	2.331	0.096	0.312
R: 公務 (他に分類されないもの)	53	0.000	0.000	0.018	0.059	0.663	0.078	0.160

(2) KSLOC 当り工数（生産性）の業種間比較

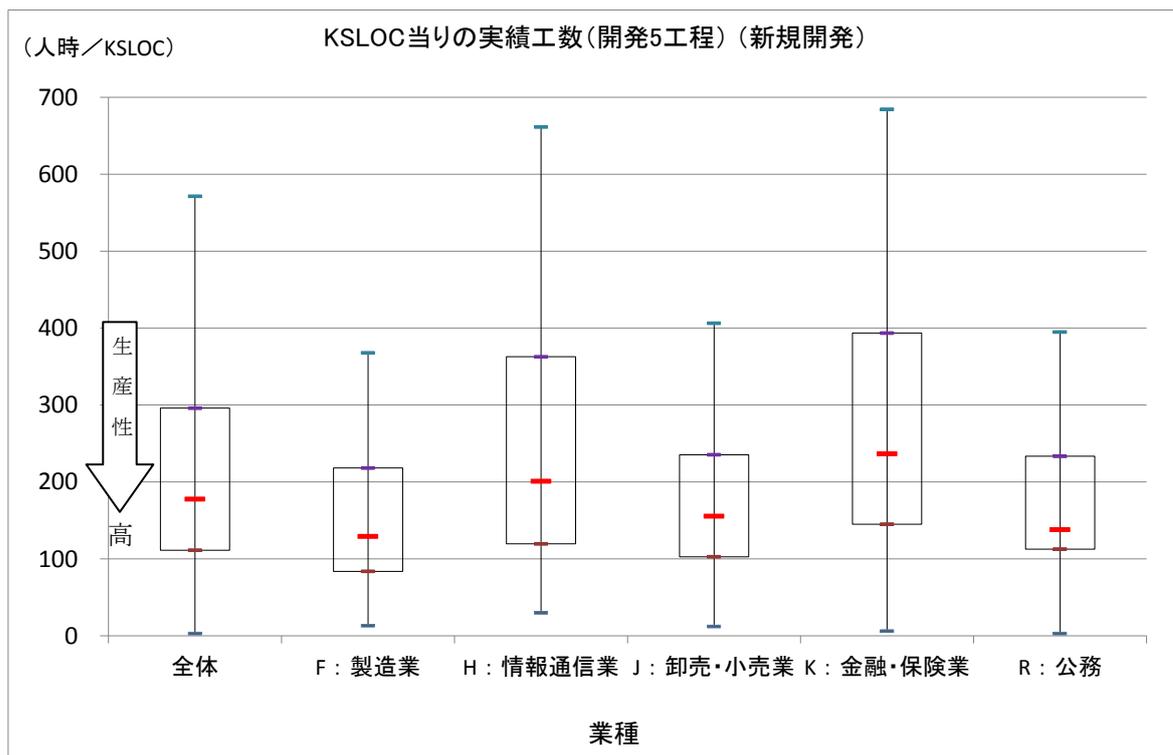
<新規開発の場合>

製造業は他より KSLOC 当り工数が少ない傾向が見られる（Welch の t 検定で 1%有意）。

情報通信業は他より KSLOC 当り工数が多い傾向が見られる（Welch の t 検定で 10%有意）。

金融・保険業は他より KSLOC 当り工数が多い傾向が見られる（Welch の t 検定で 1%有意）。

箱ひげ図



大業種	N	最小	P25	中央	P75	KSLOC当り工数 (人時/KSLOC)		
						最大	平均	標準偏差
全体	632	3	111	178	296	2788	256	278
F: 製造業	91	13	84	129	218	1185	189	203
H: 情報通信業	99	30	120	201	363	2788	287	315
J: 卸売・小売業	55	12	103	155	235	1391	226	247
K: 金融・保険業	201	6	145	236	394	1876	331	320
R: 公務(他に分類されないもの)	65	3	113	138	233	2441	223	319

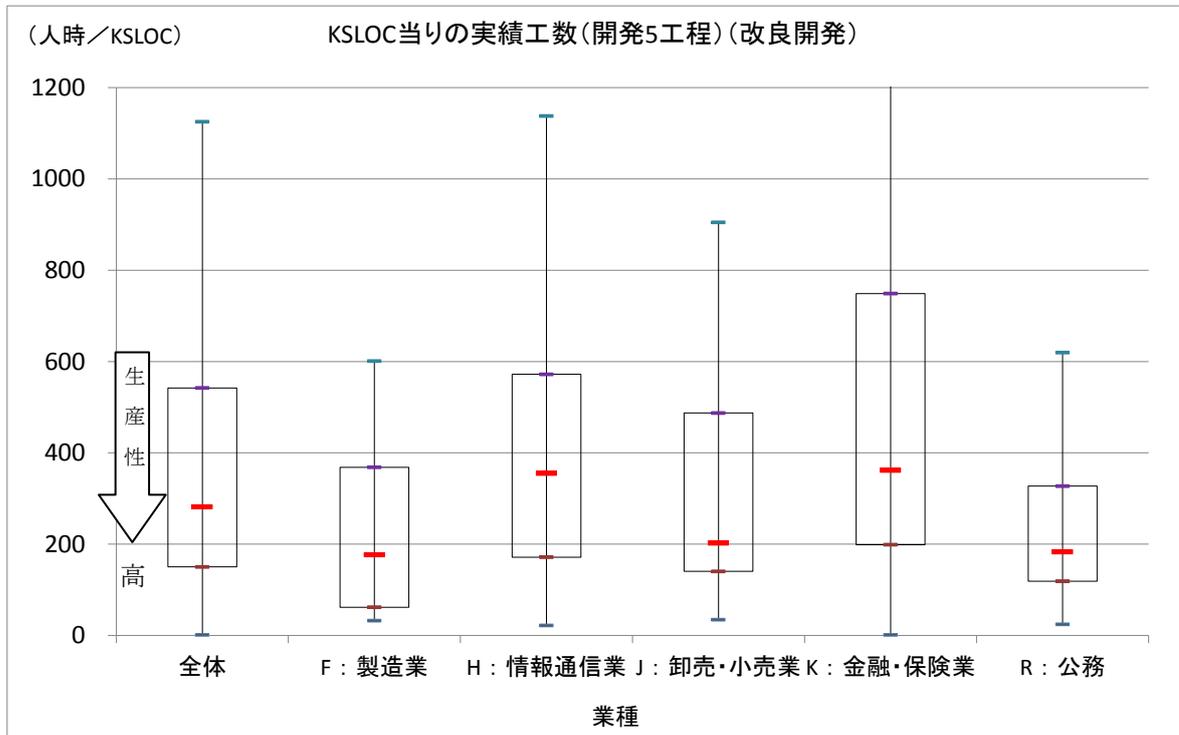
<改良開発の場合>

製造業は他より KSLOC 当り工数が少ない傾向が見られる (Welch の t 検定で 5%有意)。

情報通信業は他より KSLOC 当り工数が多い傾向が見られる (Welch の t 検定で 5%有意)。

金融・保険業は他より KSLOC 当り工数が多い傾向が見られる (Welch の t 検定で 1%有意)。

箱ひげ図



大業種	N	最小	P25	中央	P75	KSLOC当り工数 (人時/KSLOC)		
						最大	平均	標準偏差
全体	600	1	150	282	542	26800	719	1982
F: 製造業	43	32	62	176	368	3949	427	723
H: 情報通信業	115	22	172	355	572	8083	695	1211
J: 卸売・小売業	34	35	140	202	487	3872	555	833
K: 金融・保険業	216	1	199	361	749	26800	1103	2985
R: 公務 (他に分類されないもの)	70	25	119	183	327	1506	256	256

(3) 生産性／信頼性の変動要因候補の業種間比較

上記(1)及び(2)で示したように、信頼性、生産性共に業種間で差異が見られる。

本項では、次のことを確認すべく、生産性／信頼性の変動要因候補の値について業種間で差異が見られるかどうかを分析した結果を示す。

◇生産性および信頼性の業種間差異の中身として、(業種固有の性質以外の) どのような要素が作用しているのかを知りたい。調べられることの一例として、変動要因候補のいくつかが生産性および信頼性の業種間差異の要素になっていることを確認したい。

◇生産性および信頼性に関する種々の分析を進める上で、業種の影響をどのように勘案すると良いかを検討したい。変動要因候補のうち、生産性および信頼性の業種間差異の要素になっているものが少なくなければ、生産性および信頼性の業種間差異が確実に存在することが明らかになり、種々の分析を進める上で業種の影響を無視できないと判断できる。

<業種間比較結果概要>

生産性／信頼性の変動要因候補の業種間比較の分析の結果、次の変動要因候補に業種間の差が見られた。

[凡例]

◎：1%有意 (Welch の t 検定の P 値が 1%以下)

○：5%有意 (Welch の t 検定の P 値が 1%より大きくて 5%以下)

□：10%有意 (Welch の t 検定の P 値が 5%より大きくて 10%以下)

△：20%有意 (Welch の t 検定の P 値が 10%より大きくて 20%以下)

通番	変動要因候補	新規開発	改良開発	傾向
1	信頼性要求レベル	◎	◎	金融・保険業は、他より信頼性要求レベルが高いものの割合が多い。
2	設計文書化密度	○	◎	卸売・小売業は他より設計文書化密度が高い (改良◎) 金融・保険業は他より設計文書化密度が高い (新規○)
3	設計レビュー工数密度	□	○	情報通信業は他より設計レビュー工数密度がやや高い (改良△) 金融・保険業は他より設計レビュー工数密度が高い (新規□、改良○)
4	設計レビュー指摘密度	□	◎	製造業は他より設計レビュー指摘密度が低い。(改良◎) 卸売・小売業は他より設計レビュー指摘密度が低い (新規□) 金融・保険業は他より設計レビュー指摘密度が高い (改良◎)
5	テスト密度	○	○	製造業は他よりテスト密度が低い (新規□、改良○) 情報通信業は他よりテスト密度が高い (新規○、改良○) 金融・保険業は他よりテスト密度がやや高い (新規△、改良△)
6	テスト検出不具合密度	□	□	製造業は他よりテスト検出不具合密度が高い (新規□、改良△) 卸売・小売業は他よりテスト検出不具合密度が高い (改良□) 金融・保険業は他よりテスト検出不具合密度が低い (新規□、改良□)
7	上流工程での不具合摘出比率	○	◎	製造業は他より上流工程での不具合摘出比率がやや低い (改良△) 情報通信業は他より上流工程での不具合摘出比率がやや高い (改良△) 卸売・小売業は他より上流工程での不具合摘出比率が低い (新規○) 金融・保険業は他より上流工程での不具合摘出比率が高い (新規○、改良◎)

この結果から、次のように判断できる。

◇これらの変動要因候補が生産性および信頼性の業種間差異の要素になっている可能性があり、業種間の変動要因候補の値の差が業種間の生産性および信頼性の差に、ある程度寄与しているものと考えられる。

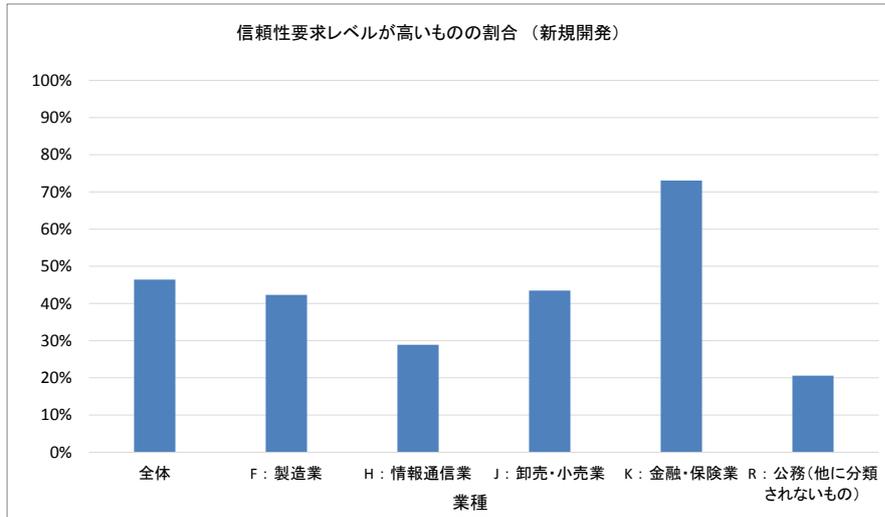
◇マネジメントや種々の分析を進める上で、業種の影響を無視できない。

以下に、これらの変動要因候補の業種間差異を具体的に示す。

◇信頼性要求レベル

<新規開発の場合>

金融・保険業は、他より信頼性要求レベルが高いものの割合が多い傾向が見られる。

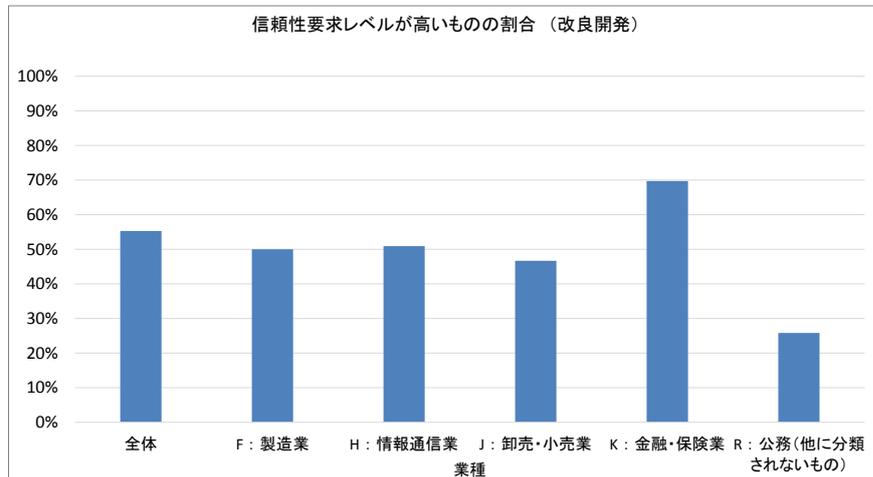


信頼性要求レベルが記入されているもののうち、信頼性要求レベルが高いもの（a又はb）の割合

業種	N	信頼性要求レベルが a又はbの件数	信頼性要求レベルが 高いものの割合
全体	237	110	46.4%
F: 製造業	26	11	42.3%
H: 情報通信業	45	13	28.9%
J: 卸売・小売業	23	10	43.5%
K: 金融・保険業	78	57	73.1%
R: 公務（他に分類されないもの）	34	7	20.6%

<改良開発の場合>

金融・保険業は、他より信頼性要求レベルが高いものの割合が多い傾向が見られる。



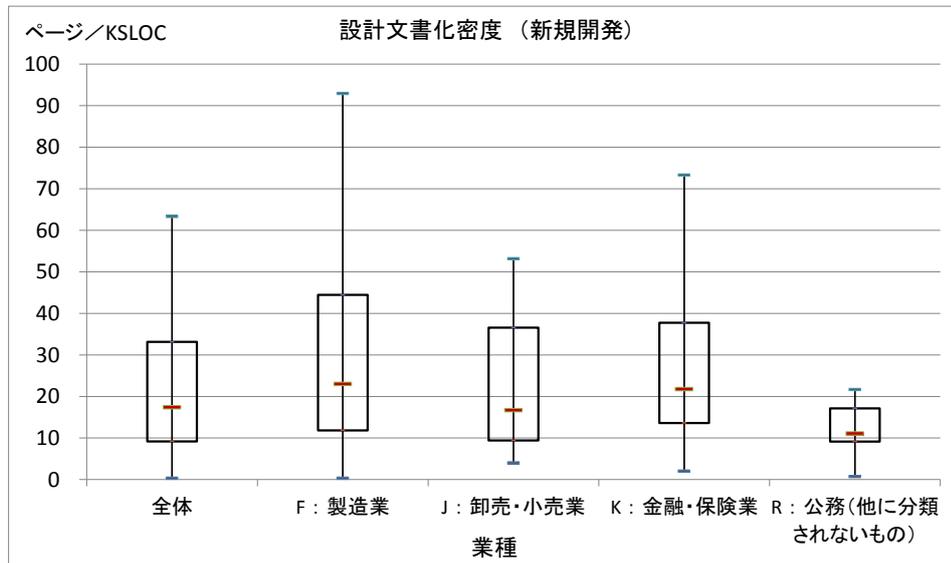
信頼性要求レベルが記入されているもののうち、信頼性要求レベルが高いもの（a又はb）の割合

業種	N	信頼性要求レベルが a又はbの件数	信頼性要求レベルが 高いものの割合
全体	275	152	55.3%
F: 製造業	26	13	50.0%
H: 情報通信業	55	28	50.9%
J: 卸売・小売業	15	7	46.7%
K: 金融・保険業	109	76	69.7%
R: 公務（他に分類されないもの）	31	8	25.8%

◇設計文書化密度

<新規開発の場合>

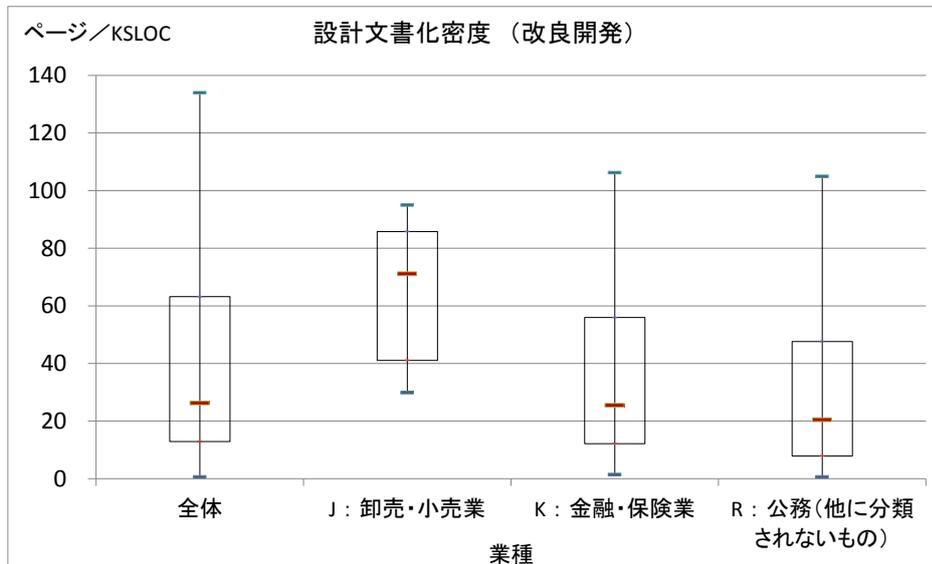
金融・保険業は、他と比べて設計文書化密度が高い傾向が見られる（5%有意）。



業種	N	最小	P25	中央	P75	設計文書化密度（ページ/KSLOC）		
						最大	平均	標準偏差
全体	160	0.3	9.2	17.4	33.1	1527.9	39.2	130.4
F: 製造業	24	0.3	11.8	23.0	44.4	109.9	33.6	30.7
J: 卸売・小売業	14	4.0	9.4	16.7	36.6	97.3	30.5	31.5
K: 金融・保険業	68	2.0	13.6	21.7	37.7	1527.9	61.3	196.8
R: 公務（他に分類されないもの）	24	0.7	9.1	11.1	17.1	47.1	13.5	8.9
H: 情報通信業	3	4.5	6.2	7.9	17.7	27.4	13.3	12.4

<改良開発の場合>

卸売・小売業は、他と比べて設計文書化密度が高い傾向が見られる（1%有意）。



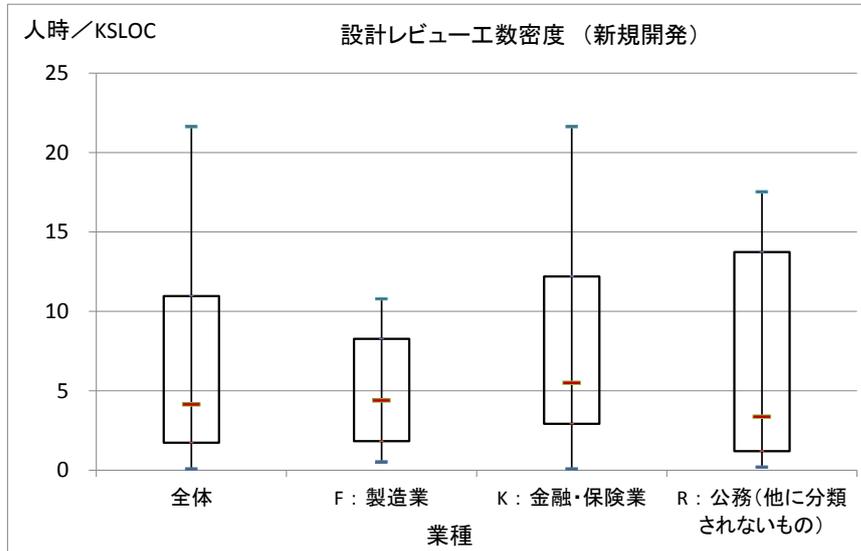
業種	N	最小	P25	中央	P75	設計文書化密度（ページ/KSLOC）		
						最大	平均	標準偏差
全体	114	0.6	12.9	26.2	63.2	2195.9	95.7	295.2
J: 卸売・小売業	10	29.9	41.1	71.1	85.8	531.2	108.5	150.4
K: 金融・保険業	57	1.5	12.1	25.4	55.9	2195.9	132.5	406.8
R: 公務（他に分類されないもの）	13	0.6	7.9	20.4	47.6	278.5	61.9	91.5
F: 製造業	9							
H: 情報通信業	8							

データ不足
データ不足

◇設計レビュー工数密度

<新規開発の場合>

金融・保険業は、他より設計レビュー工数密度が高い傾向が見られる（10%有意）。



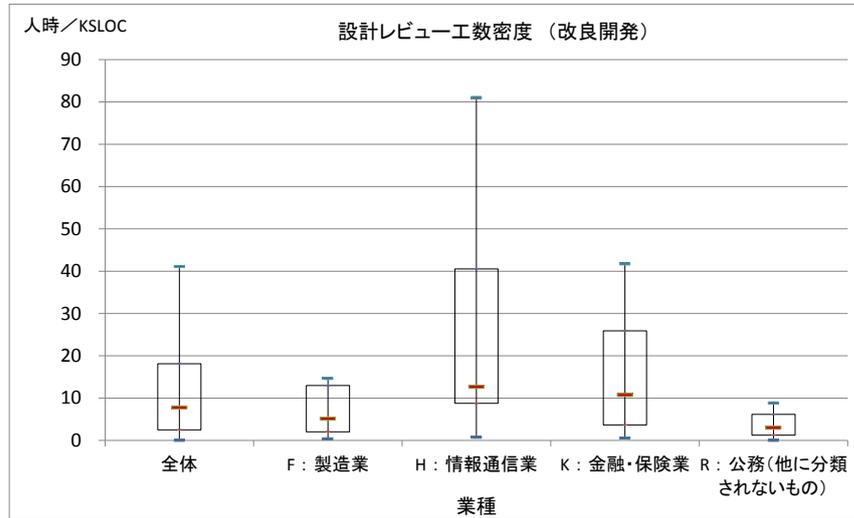
業種	N	設計レビュー工数密度(人時/KSLOC)						
		最小	P25	中央	P75	最大	平均	標準偏差
全体	104	0.08	1.73	4.13	10.96	1559.36	34.11	173.07
F: 製造業	16	0.53	1.83	4.39	8.27	55.63	10.10	14.87
K: 金融・保険業	39	0.08	2.92	5.49	12.20	1559.36	70.85	274.66
R: 公務(他に分類されないもの)	23	0.20	1.20	3.36	13.74	17.54	6.74	6.39
H: 情報通信業	5							
J: 卸売・小売業	5							

データ不足
データ不足

<改良開発の場合>

情報通信業は、他より設計レビュー工数密度がやや高い傾向が見られる（20%有意）。

金融・保険業は、他より設計レビュー工数密度が高い傾向が見られる（5%有意）。



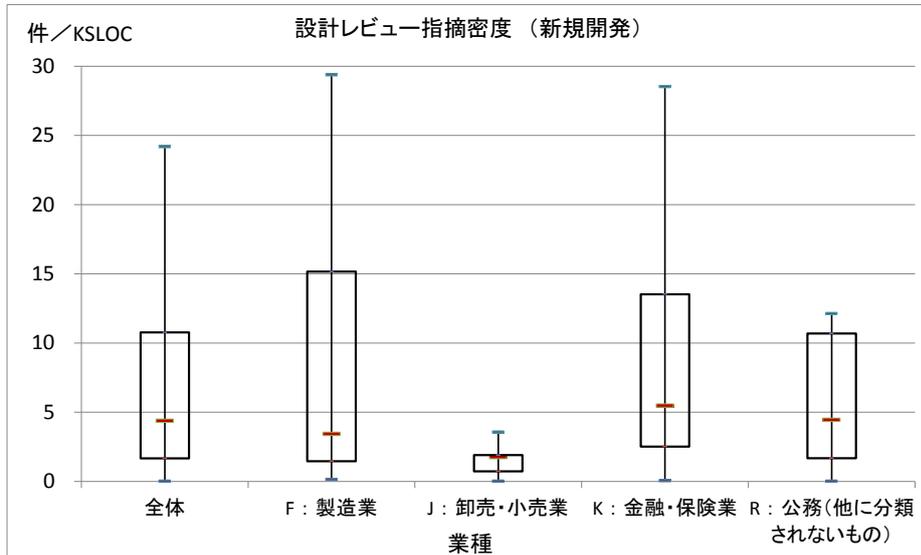
業種	N	設計レビュー工数密度(人時/KSLOC)						
		最小	P25	中央	P75	最大	平均	標準偏差
全体	120	0.03	2.49	7.77	18.12	33402.06	801.07	4894.42
F: 製造業	10	0.40	2.03	5.09	12.97	266.67	43.00	87.33
H: 情報通信業	18	0.73	8.78	12.63	40.53	81.08	22.65	21.33
K: 金融・保険業	44	0.57	3.67	10.72	25.90	33402.06	2154.14	7958.30
R: 公務(他に分類されないもの)	16	0.03	1.27	2.96	6.16	31.00	5.77	8.22
J: 卸売・小売業	9							

データ不足

◇設計レビュー指摘密度

<新規開発の場合>

卸売・小売業は、他より設計レビュー指摘密度が低い傾向が見られる（10%有意）。



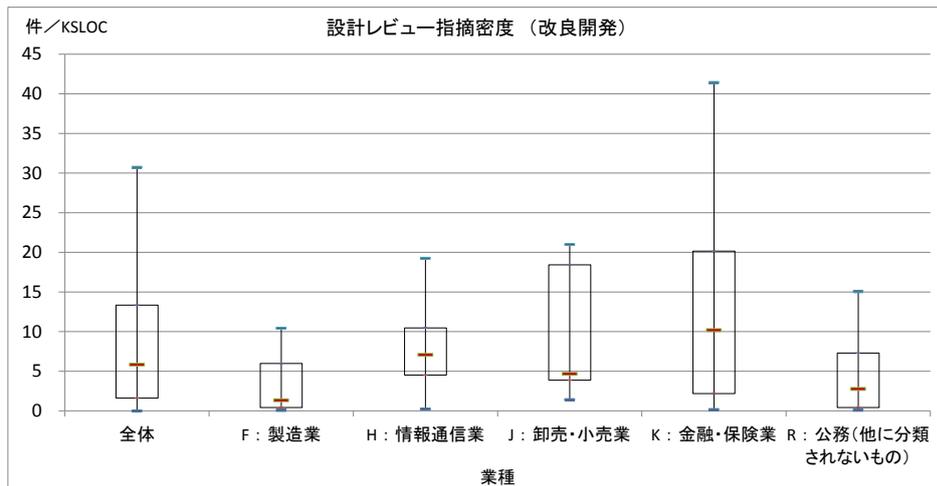
業種	N	設計レビュー指摘密度 (件/KSLOC)						
		最小	P25	中央	P75	最大	平均	標準偏差
全体	138	0.02	1.67	4.38	10.76	75.43	8.17	11.95
F: 製造業	25	0.15	1.46	3.43	15.16	75.43	11.27	16.99
J: 卸売・小売業	10	0.02	0.73	1.75	1.90	26.92	3.93	8.15
K: 金融・保険業	43	0.06	2.52	5.47	13.52	73.19	10.81	14.88
R: 公務 (他に分類されないもの)	28	0.03	1.68	4.44	10.69	12.14	5.32	4.46
H: 情報通信業	6							

データ不足

<改良開発の場合>

製造業は、他より設計レビュー指摘密度が低い傾向が見られる（1%有意）。

金融・保険業は、他より設計レビュー指摘密度が高い傾向が見られる（1%有意）。



業種	N	設計レビュー指摘密度 (件/KSLOC)						
		最小	P25	中央	P75	最大	平均	標準偏差
全体	171	0.00	1.62	5.83	13.35	2100.00	28.37	167.16
F: 製造業	16	0.06	0.42	1.32	5.98	14.56	3.49	4.29
H: 情報通信業	27	0.23	4.52	7.06	10.47	33.69	8.81	7.01
J: 卸売・小売業	10	1.42	3.88	4.67	18.43	78.46	15.25	23.36
K: 金融・保険業	70	0.16	2.19	10.18	20.15	2100.00	57.52	259.23
R: 公務 (他に分類されないもの)	19	0.07	0.43	2.76	7.30	15.04	4.29	4.86

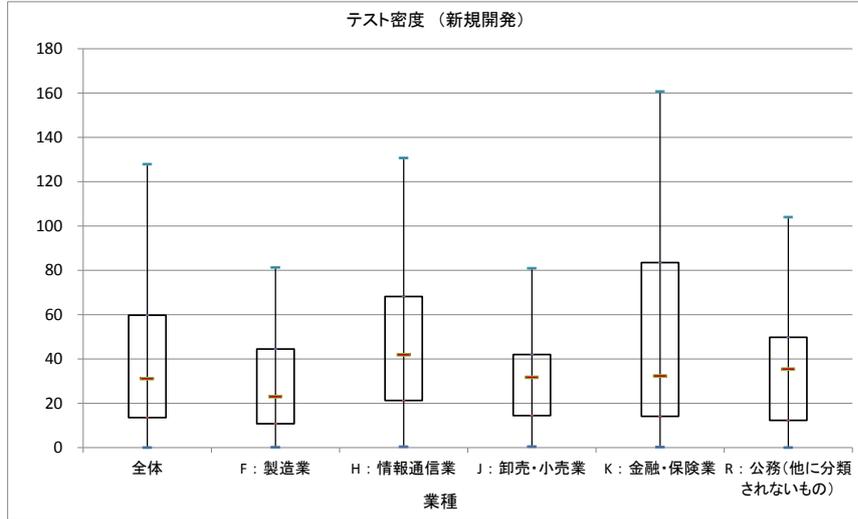
◇テスト密度

<新規開発の場合>

製造業は、他よりテスト密度が低い傾向が見られる (10%有意)。

情報通信業は、他よりテスト密度が高い傾向が見られる (5%有意)。

金融・保険業は、他よりテスト密度がやや高い傾向が見られる (20%有意)。



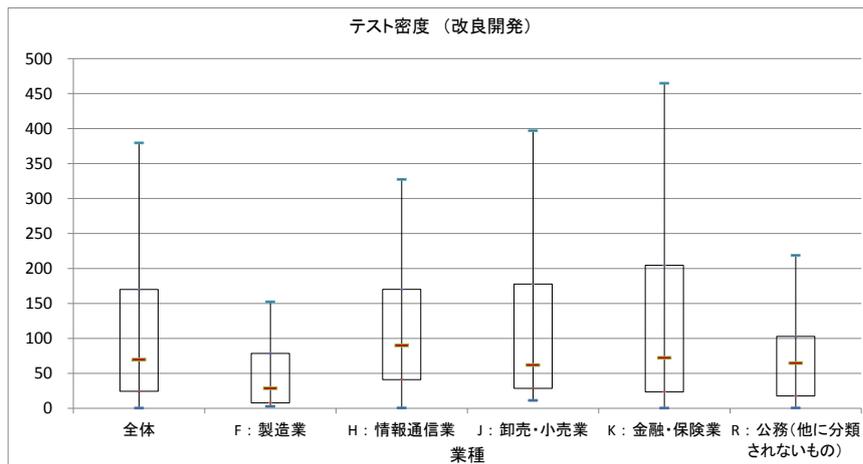
業種	N	最小	P25	中央	P75	最大	テスト密度 (ヶ-λ/KSLOC)	
							平均	標準偏差
全体	509	0.0	13.5	31.1	59.8	1703.3	65.4	146.1
F: 製造業	66	0.2	10.8	22.9	44.5	302.8	43.9	59.9
H: 情報通信業	87	0.4	21.2	41.9	68.2	970.7	64.5	111.5
J: 卸売・小売業	42	0.5	14.4	31.7	42.0	587.1	49.1	95.0
K: 金融・保険業	162	0.2	14.1	32.3	83.5	1703.3	94.1	225.2
R: 公務 (他に分類されないもの)	58	0.0	12.3	35.4	49.8	698.5	59.4	101.2

<改良開発の場合>

製造業は、他よりテスト密度が低い傾向が見られる (5%有意)。

情報通信業は、他よりテスト密度が高い傾向が見られる (5%有意)。

金融・保険業は、他よりテスト密度がやや高い傾向が見られる (20%有意)。



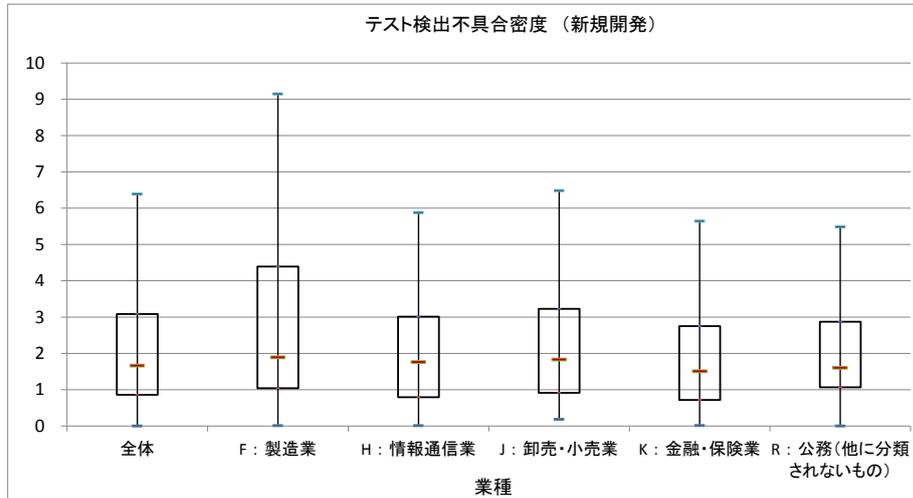
業種	N	最小	P25	中央	P75	最大	テスト密度 (ヶ-λ/KSLOC)	
							平均	標準偏差
全体	477	0.4	24.4	69.6	170.0	12800.0	288.3	951.4
F: 製造業	31	2.9	7.9	28.5	78.6	2756.9	158.5	494.0
H: 情報通信業	94	0.5	41.0	89.8	170.1	2568.0	218.7	416.4
J: 卸売・小売業	29	11.4	28.6	61.9	177.7	1569.2	185.5	321.2
K: 金融・保険業	185	0.4	23.7	72.2	204.5	12800.0	399.1	1266.9
R: 公務 (他に分類されないもの)	55	0.6	17.8	64.6	102.9	349.6	83.3	83.0

◇テスト検出不具合密度

<新規開発の場合>

製造業は、他よりテスト検出不具合密度が高い傾向が見られる（10%有意）。

金融・保険業は、他よりテスト検出不具合密度が低い傾向が見られる（10%有意）。



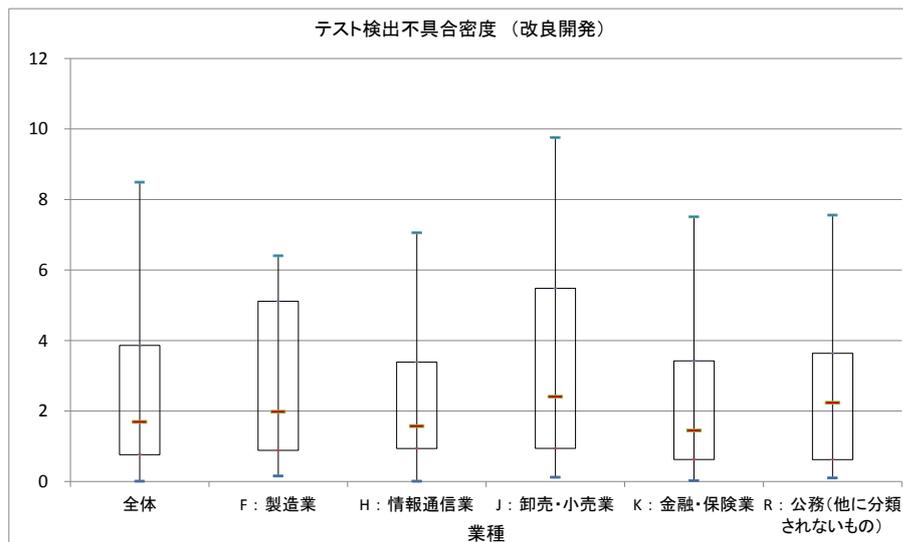
業種	N	最小	P25	中央	P75	テスト検出不具合密度 (件/KSLOC)		
						最大	平均	標準偏差
全体	511	0.00	0.86	1.66	3.08	43.93	2.56	3.50
F: 製造業	67	0.01	1.04	1.89	4.39	24.13	3.35	4.18
H: 情報通信業	84	0.01	0.79	1.76	3.01	27.77	2.70	3.66
J: 卸売・小売業	46	0.18	0.92	1.83	3.23	13.46	2.60	2.76
K: 金融・保険業	160	0.01	0.72	1.51	2.75	43.93	2.51	4.30
R: 公務(他に分類されないもの)	59	0.00	1.07	1.60	2.87	11.98	2.33	2.27

<改良開発の場合>

製造業は、他よりテスト検出不具合密度がやや高い傾向が見られる（20%有意）。

卸売・小売業は、他よりテスト検出不具合密度が高い傾向が見られる（10%有意）。

金融・保険業は、他よりテスト検出不具合密度が低い傾向が見られる（10%有意）。



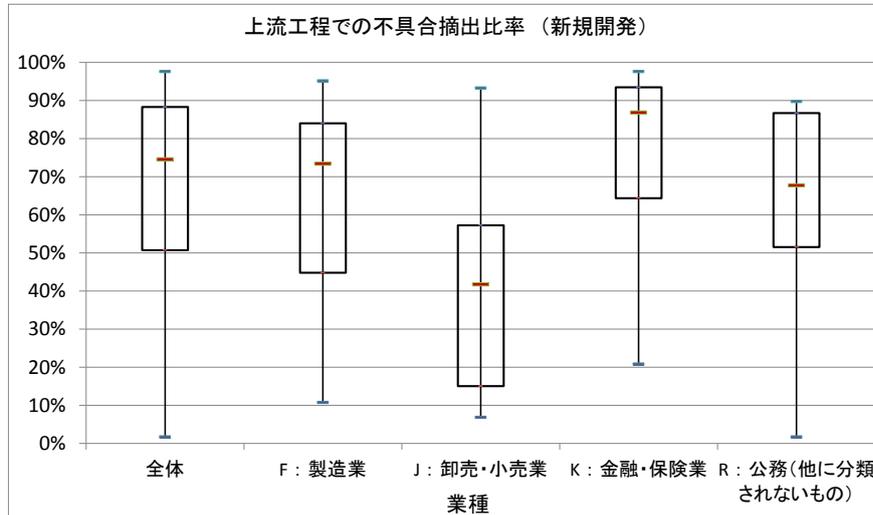
業種	N	最小	P25	中央	P75	テスト検出不具合密度 (件/KSLOC)		
						最大	平均	標準偏差
全体	474	0.01	0.76	1.69	3.86	542.83	4.98	26.02
F: 製造業	34	0.16	0.88	1.98	5.11	53.33	5.72	10.23
H: 情報通信業	91	0.01	0.94	1.57	3.39	542.83	8.94	56.80
J: 卸売・小売業	30	0.12	0.94	2.40	5.48	92.31	9.24	20.09
K: 金融・保険業	170	0.02	0.62	1.45	3.42	42.55	3.14	5.26
R: 公務(他に分類されないもの)	57	0.10	0.62	2.23	3.64	21.93	3.36	4.31

◇上流工程での不具合摘出比率

<新規開発の場合>

卸売・小売業は、他より上流工程での不具合摘出比率が低い傾向が見られる（5%有意）。

金融・保険業は、他より上流工程での不具合摘出比率が高い傾向が見られる（5%有意）。



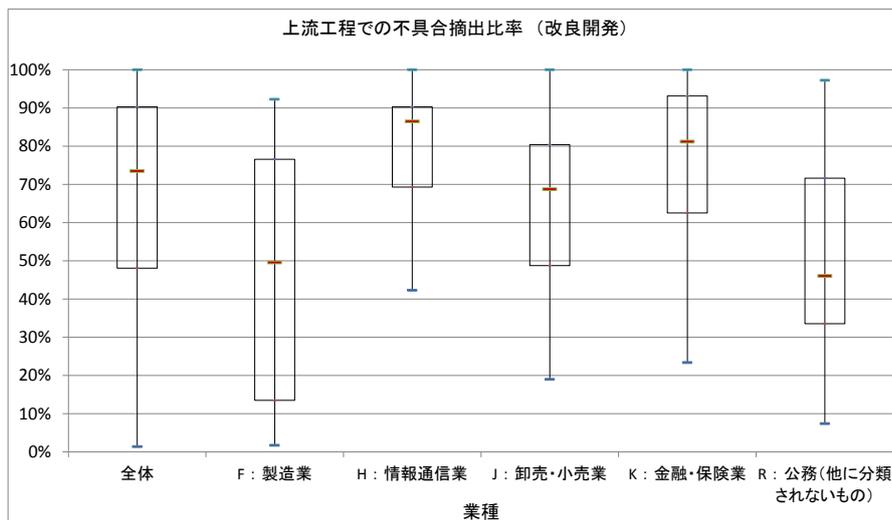
業種	N	上流工程での不具合摘出比率 (%)							
		最小	P25	中央	P75	最大	平均	標準偏差	
全体	133	1.7%	50.7%	74.6%	88.3%	97.6%	65.9%	27.5%	
F: 製造業	24	10.8%	44.8%	73.4%	84.0%	95.2%	63.2%	26.6%	
J: 卸売・小売業	10	6.9%	15.0%	41.7%	57.2%	93.3%	40.9%	28.6%	
K: 金融・保険業	41	1.8%	64.4%	86.8%	93.5%	97.6%	74.0%	27.7%	
R: 公務(他に分類されないもの)	27	1.7%	51.5%	67.7%	86.7%	89.7%	62.1%	28.1%	
H: 情報通信業	6	66.3%	71.3%	75.0%	83.1%	89.5%	76.9%	8.9%	データ不足

<改良開発の場合>

製造業は、他より上流工程での不具合摘出比率がやや低い傾向が見られる（20%有意）。

情報通信業は、他より上流工程での不具合摘出比率がやや高い傾向が見られる（20%有意）。

金融・保険業は、他より上流工程での不具合摘出比率が高い傾向が見られる（1%有意）。



業種	N	上流工程での不具合摘出比率 (%)						
		最小	P25	中央	P75	最大	平均	標準偏差
全体	165	1.4%	48.1%	73.5%	90.3%	100.0%	65.8%	28.2%
F: 製造業	16	1.7%	13.5%	49.5%	76.6%	92.3%	44.9%	31.7%
H: 情報通信業	25	15.4%	69.3%	86.5%	90.3%	100.0%	76.9%	21.3%
J: 卸売・小売業	10	19.0%	48.8%	68.7%	80.4%	100.0%	64.5%	23.5%
K: 金融・保険業	66	8.4%	62.5%	81.2%	93.1%	100.0%	75.1%	22.9%
R: 公務(他に分類されないもの)	19	7.4%	33.5%	46.0%	71.6%	97.3%	50.9%	30.2%

3.2.3. 生産性変動要因の分析と活用

【概要】

分析項目	生産性 (SLOC 規模当たり実績工数) の変動要因分析 (開発プロセス関連)
メッセージ	<p>◇分析結果概要 (読み取れる傾向)</p> <p>開発プロセス関連の生産性変動要因候補 (メトリクス欄参照) によって、生産性 (SLOC 規模当たり実績工数 (開発 5 工程)) が変動する傾向が見られる。</p> <p>◇助言</p> <ul style="list-style-type: none"> ・生産性向上のための組織の重点強化領域の特定に向けて 生産性向上を進めて行くためには、個々のプロジェクトのマネジメントよりも、組織の改善に向けたマネジメント・サイクルを回すことが一層重要であろう。重点的に強化すると効果的な領域を特定し、適切な方策を立てることが望まれる。そのために、組織の生産性変動要因群を把握することが効果的であり、お勧めする。 ・開発計画の実現可能性検討、工数見積りの妥当性評価等に向けて 生産性の目標値とそれを達成するための開発プロセスの目標値あるいは見積り工数が、一定の妥当な範囲 (例えば管理指標の 25 パーセンタイル値～P75 パーセンタイル値の範囲) に収まっているか否かで評価するのが基本的な妥当性評価方法ではあるが、その範囲に収まっていないというだけで NG (妥当でない) と評価するのは早計である。どういうプロジェクトなのかによっては、生産性変動要因による変動を始めとして一定の範囲に収まらなくなる合理的な理由が存在する可能性がある。評価対象プロジェクトに該当する生産性変動要因によって生じる変動幅を勘案して、一定の妥当な範囲を上方修正/下方修正しながら妥当性評価することをお勧めする。その結果においても妥当な範囲外となり、かつ生産性変動要因以外の合理的な理由がない場合には、計画や見積りを見直すことが望ましい。
メトリクス	<p>開発プロセス関連の生産性変動要因候補として分析したものは次の通り。</p> <p>◇設計文書化密度：(基本設計書のページ数+詳細設計書のページ数) ÷ 開発規模 (KSLOC)</p> <p>◇設計レビュー工数密度：基本設計から製作までの設計レビュー工数 ÷ 開発規模 (KSLOC)</p> <p>◇設計レビュー指摘密度：基本設計から製作までのレビュー指摘数 ÷ 開発規模 (KSLOC)</p> <p>◇テスト密度：(結合テストケース数+総合テスト (ベンダ確認) ケース数) ÷ 開発規模 (KSLOC)</p> <p>◇テスト検出不具合密度：(結合テスト検出不具合数+総合テスト (ベンダ確認) 検出不具合数) ÷ 開発規模 (KSLOC)</p> <p>◇上流工程での不具合摘出比率：基本設計から製作までのレビュー指摘数 ÷ (基本設計から製作までのレビュー指摘数+結合テストから総合テスト (ベンダ確認) までの検出不具合数)</p>
考察/備考	<p>IPA/SEC データベースにおいて業種を始め様々なプロファイルのプロジェクトデータが混在した中では、開発プロセス関連の生産性変動要因候補 (メトリクス欄参照) によって生産性が変動する傾向が見られる。各組織の同種のソフトウェ</p>

	ア開発を行うドメインの中では、同様な傾向が見られるとは限らないが、より強い傾向を示す変動要因が見つかることと期待される。自組織の生産性変動要因群を分析・把握することが重要かつ効果的である。
--	--

【検索条件】

- ◇ 開発 5 工程が揃っているプロジェクト
- ◇ 開発プロジェクトの種別が新規開発又は改良開発（改修・保守又は拡張）
- ◇ 開発規模（SLOC 規模） > 0
- ◇ 主開発言語が COBOL, C, VB, Java のいずれか
- ◇ 実績工数（開発 5 工程） > 0

【分析方法】

- ◇ 変動要因候補のデータの中央値によって分析対象プロジェクト群を値が大きいものと小さいものとの二群に大別して、両者の SLOC 規模当たりの実績工数（開発 5 工程）を比較する。
- ◇ 傾向が見られるかどうか（二群に大別した両者に差があるかどうか）に関しては、原則として常用対数化した開発規模当たりの実績工数に対して Welch の t 検定（分散が等しくないと仮定した 2 標本による平均値の差の t 検定）を行った上で記述する。ただし、データ分布の非対称性が弱いケース（歪度の絶対値が 2 未満のケース）が少数あるが、そのケースに対しては対数化しないデータを用いて検定する。

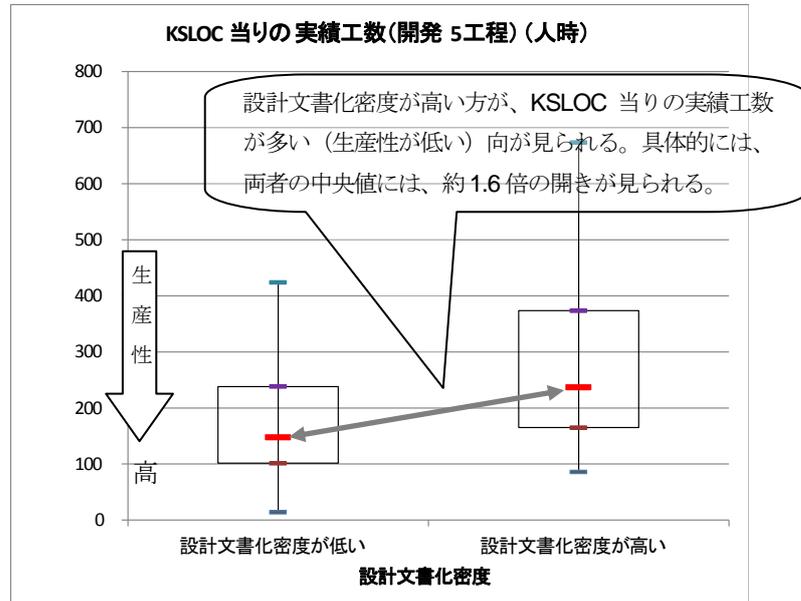
【分析結果】

マトリクス欄の生産性変動要因候補によって、生産性（SLOC 規模当たり実績工数（開発 5 工程））が変動する傾向が見られる。

- [凡例] ◎ : 1%有意 (Welch の t 検定の P 値が 1%以下)
 ○ : 5%有意 (Welch の t 検定の P 値が 1%より大きくて 5%以下)
 □ : 10%有意 (Welch の t 検定の P 値が 5%より大きくて 10%以下)
 △ : 20%有意 (Welch の t 検定の P 値が 10%より大きくて 20%以下)
 空白 : 有意でない (Welch の t 検定の P 値が 20%より大きい)

通番	変動要因候補	新規開発	改良開発	傾向
1	設計文書化密度	◎	◎	設計文書化密度が高い方が KSLOC 当り工数が多い
2	設計レビュー工数密度	◎	◎	設計レビュー工数密度が高い方が KSLOC 当り工数が多い
3	設計レビュー指摘密度	◎	◎	設計レビュー指摘密度が高い方が KSLOC 当り工数が多い
4	テスト密度	◎	◎	テスト密度が高い方が KSLOC 当り工数が多い
5	テスト検出不具合密度	◎	◎	テスト検出不具合密度が高い方が KSLOC 当り工数が多い
6	上流工程での不具合摘出比率	△	○	上流での不具合摘出比率が高い方が KSLOC 当り工数が多い (改良)

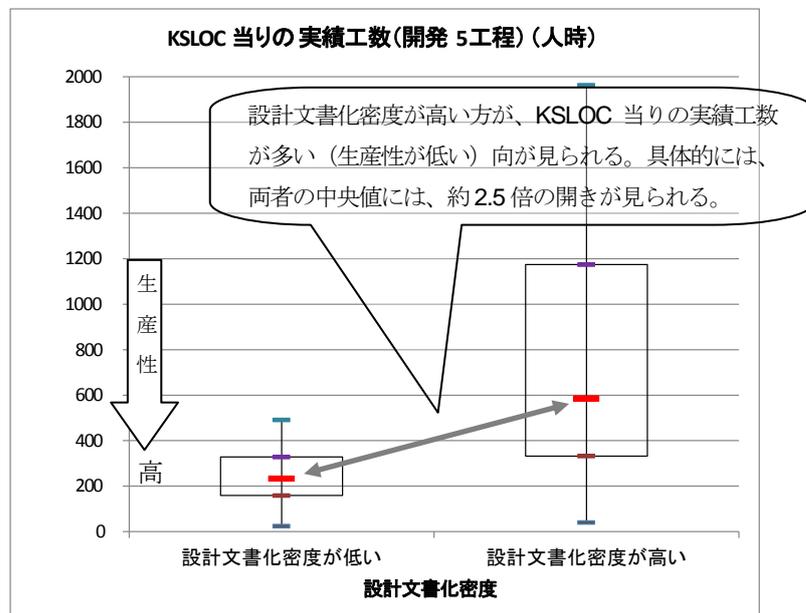
(1) 設計文書化密度による生産性の変動
 <新規開発の場合>



設計文書化密度	N	最小	P25	中央	P75	最大	平均	標準偏差	(人時/KSLOC)
設計文書化密度が低い	80	14	101	147	238	664	187	131	
設計文書化密度が高い	79	86	165	236	373	1412	321	252	

設計文書化密度が低い: 設計文書化密度が中央値17.31以下
 設計文書化密度が高い: 設計文書化密度が中央値17.31より大

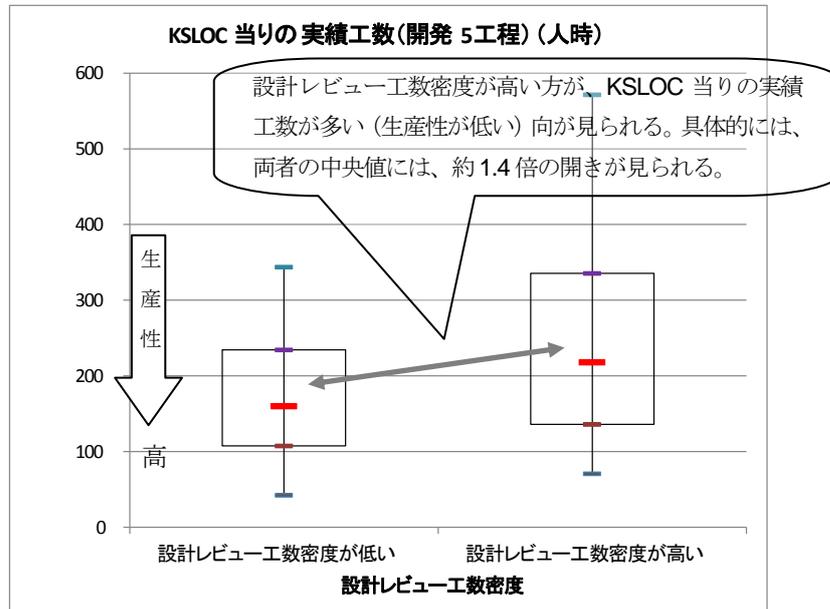
<改良開発の場合>



設計文書化密度	N	最小	P25	中央	P75	最大	平均	標準偏差	(人時/KSLOC)
設計文書化密度が低い	57	25	159	232	328	1564	289	257	
設計文書化密度が高い	57	40	332	585	1174	12550	1294	2385	

設計文書化密度が低い: 設計文書化密度が中央値26.2以下
 設計文書化密度が高い: 設計文書化密度が中央値26.2より大

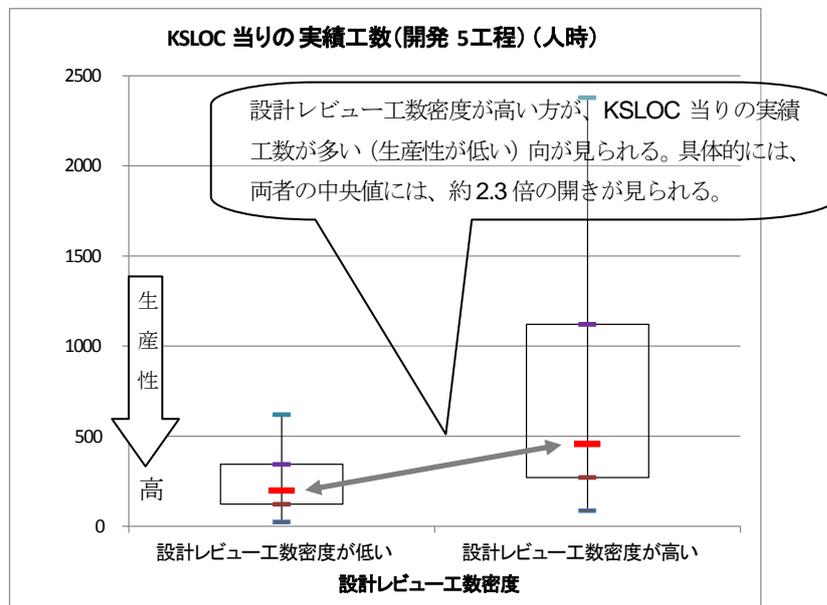
(2) 設計レビュー工数密度による生産性の変動
 <新規開発の場合>



KSLOC当りの実績工数(開発5工程)の基本統計量								(人時/KSLOC)	
設計レビュー工数密度	N	最小	P25	中央	P75	最大	平均	標準偏差	
設計レビュー工数密度が低い	50	42	108	160	234	874	202	167	
設計レビュー工数密度が高い	50	71	136	218	335	1412	304	256	

設計レビュー工数密度が低い: 設計レビュー工数密度が中央値3.68以下
 設計レビュー工数密度が高い: 設計レビュー工数密度が中央値3.68より大

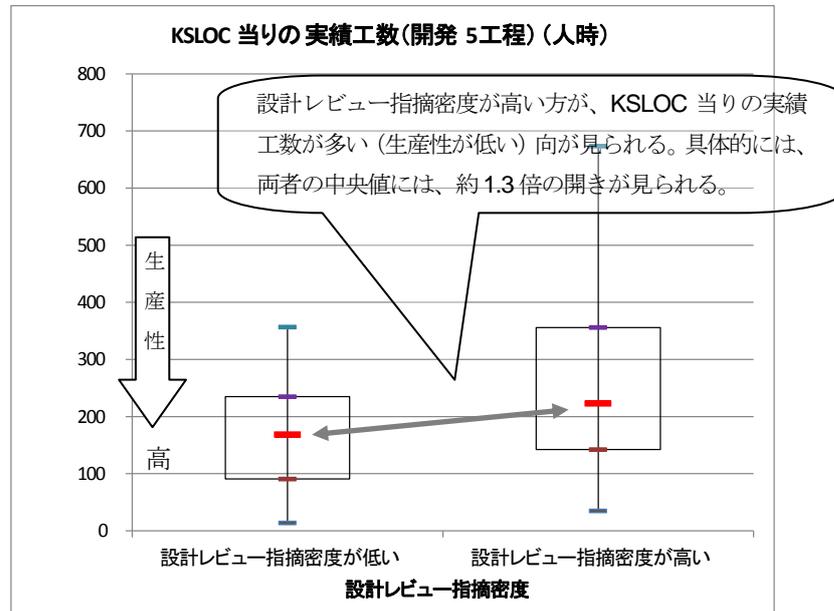
<改良開発の場合>



KSLOC当りの実績工数(開発5工程)の基本統計量								(人時/KSLOC)	
設計レビュー工数密度	N	最小	P25	中央	P75	最大	平均	標準偏差	
設計レビュー工数密度が低い	58	25	122	198	344	1689	326	356	
設計レビュー工数密度が高い	58	86	271	456	1120	7447	912	1217	

設計レビュー工数密度が低い: 設計レビュー工数密度が中央値7.31以下
 設計レビュー工数密度が高い: 設計レビュー工数密度が中央値7.31より大

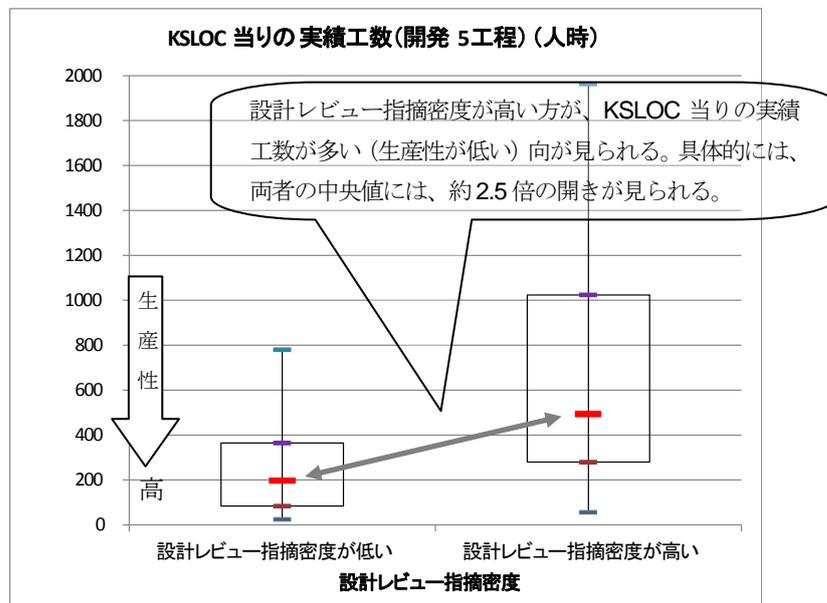
(3) 設計レビュー指摘密度による生産性の変動
 <新規開発の場合>



KSLOC当りの実績工数(開発5工程)の基本統計量								(人時/KSLOC)	
設計レビュー指摘密度	N	最小	P25	中央	P75	最大	平均	標準偏差	
設計レビュー指摘密度が低い	70	14	91	168	235	1100	213	198	
設計レビュー指摘密度が高い	70	35	142	222	356	2441	327	348	

設計レビュー指摘密度が低い: 設計レビュー指摘密度が中央値4.357以下
 設計レビュー指摘密度が高い: 設計レビュー指摘密度が中央値4.357より大

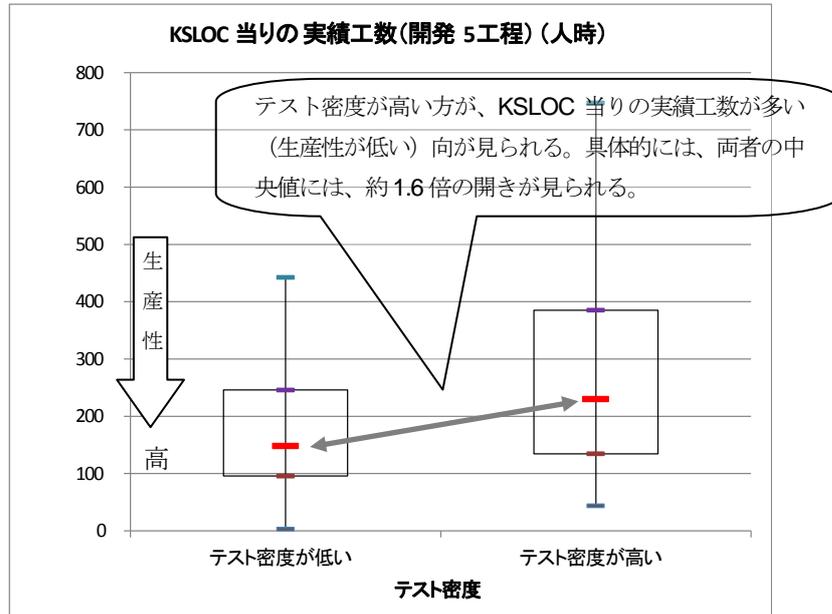
<改良開発の場合>



KSLOC当りの実績工数(開発5工程)の基本統計量								(人時/KSLOC)	
設計レビュー指摘密度	N	最小	P25	中央	P75	最大	平均	標準偏差	
設計レビュー指摘密度が低い	87	25	84	196	364	3941	376	580	
設計レビュー指摘密度が高い	87	56	280	493	1024	26800	1394	3470	

設計レビュー指摘密度が低い: 設計レビュー指摘密度が中央値5.7以下
 設計レビュー指摘密度が高い: 設計レビュー指摘密度が中央値5.7より大

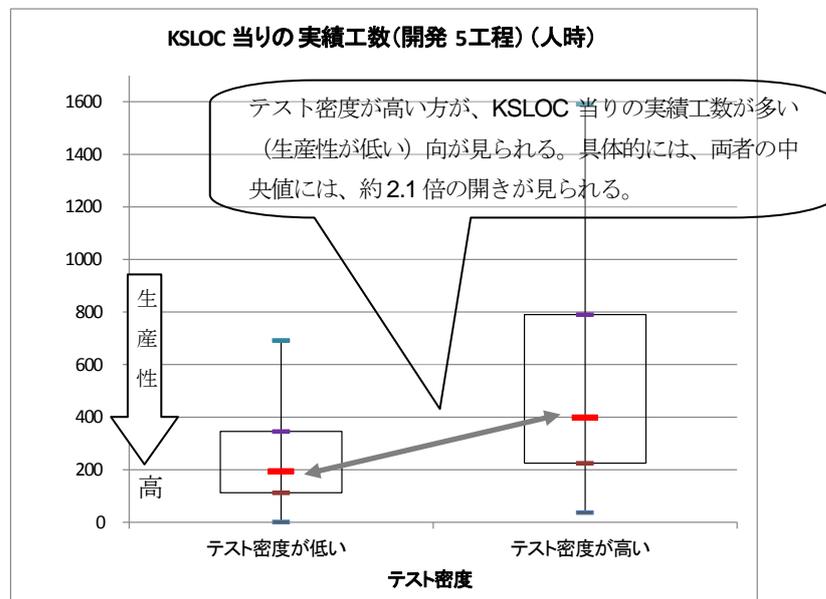
(4) テスト密度による生産性の変動
 <新規開発の場合>



KSLOC当りの実績工数(開発5工程)の基本統計量								(人時/KSLOC)	
テスト密度	N	最小	P25	中央	P75	最大	平均	標準偏差	
テスト密度が低い	252	3	96	148	246	2441	199	212	
テスト密度が高い	252	44	135	230	385	2788	327	328	

テスト密度が低い: テスト密度が中央値31.57以下
 テスト密度が高い: テスト密度が中央値31.57より大

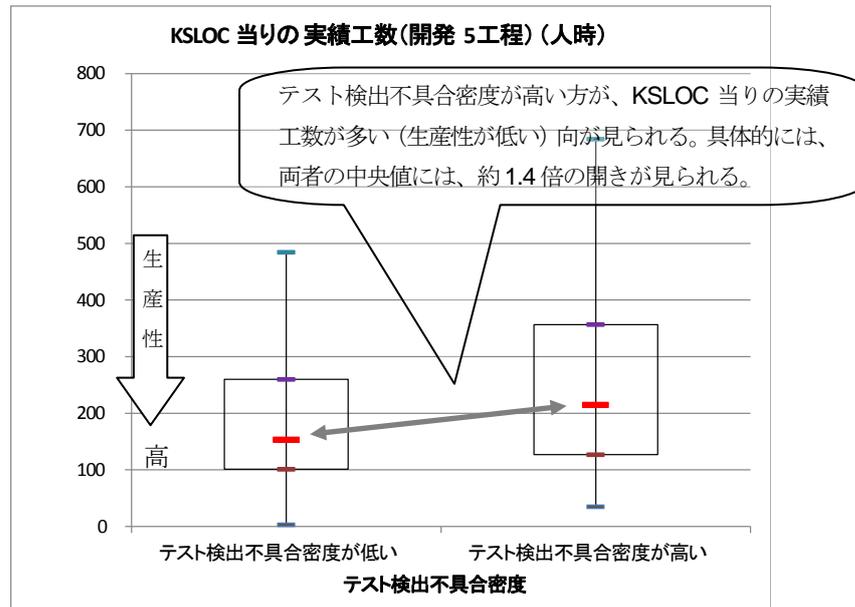
<改良開発の場合>



KSLOC当りの実績工数(開発5工程)の基本統計量								(人時/KSLOC)	
テスト密度	N	最小	P25	中央	P75	最大	平均	標準偏差	
テスト密度が低い	236	1	112	193	346	3788	311	432	
テスト密度が高い	235	37	225	398	790	26800	1111	2823	

テスト密度が低い: テスト密度が中央値71.77以下
 テスト密度が高い: テスト密度が中央値71.77より大

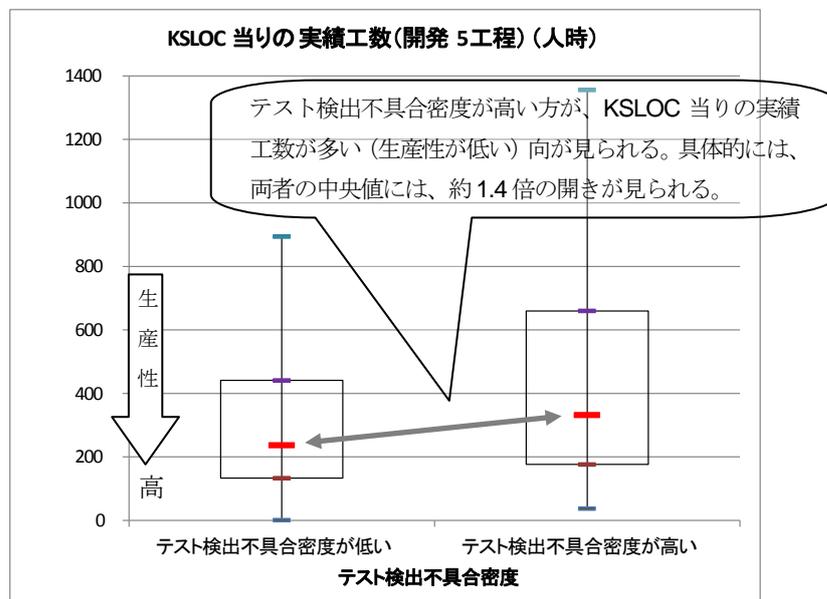
(5) テスト検出不具合密度による生産性の変動
 <新規開発の場合>



テスト検出不具合密度	N	最小	P25	中央	P75	最大	平均	標準偏差	(人時/KSLOC)
テスト検出不具合密度が低い	261	3	101	153	260	1412	208	183	
テスト検出不具合密度が高い	261	35	127	214	357	2788	311	344	

テスト検出不具合密度が低い: テスト検出不具合密度が中央値1.603以下
 テスト検出不具合密度が高い: テスト検出不具合密度が中央値1.603より大

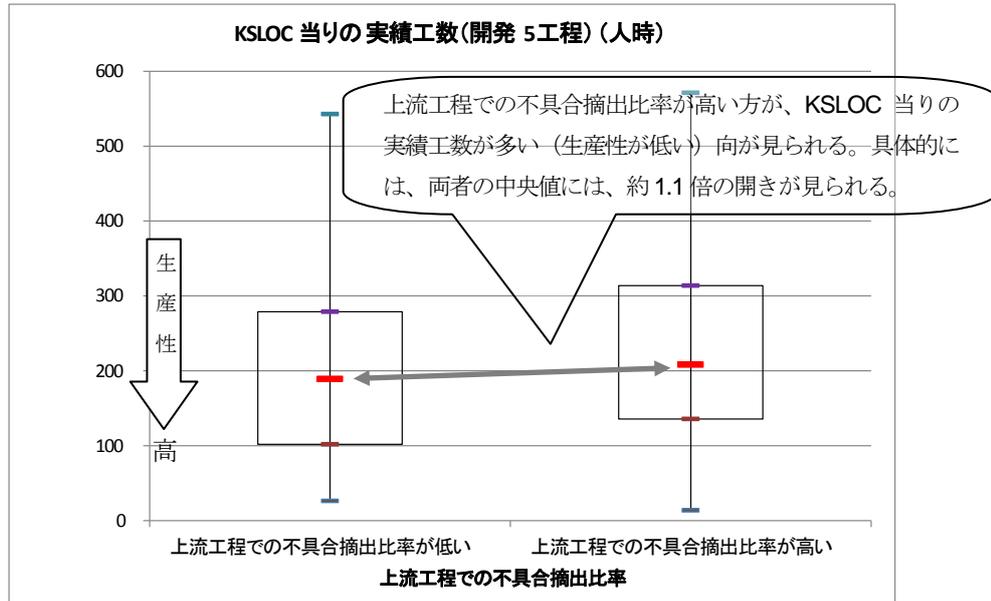
<改良開発の場合>



テスト検出不具合密度	N	最小	P25	中央	P75	最大	平均	標準偏差	(人時/KSLOC)
テスト検出不具合密度が低い	252	1	133	236	441	26800	683	2454	
テスト検出不具合密度が高い	249	37	176	331	660	12550	726	1400	

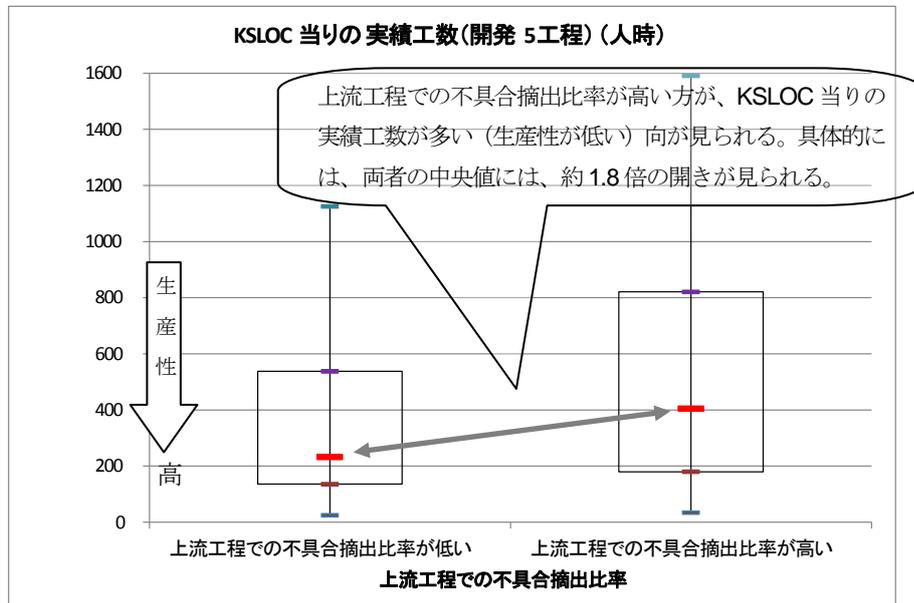
テスト検出不具合密度が低い: テスト検出不具合密度が中央値1.6以下
 テスト検出不具合密度が高い: テスト検出不具合密度が中央値1.6より大

(6) 上流工程での不具合抽出比率による生産性の変動
 <新規開発の場合>



上流工程での不具合抽出比率	N	最小	P25	中央	P75	最大	平均	標準偏差	(人時/KSLOC)
上流工程での不具合抽出比率が低い	67	26	102	189	279	2441	257	331	
上流工程での不具合抽出比率が高い	68	14	136	208	314	1412	274	228	
上流工程での不具合抽出比率が低い: 抽出比率が中央値73.7%以下									
上流工程での不具合抽出比率が高い: 抽出比率が中央値73.7%より大									

<改良開発の場合>



上流工程での不具合抽出比率	N	最小	P25	中央	P75	最大	平均	標準偏差	(人時/KSLOC)
上流工程での不具合抽出比率が低い	80	25	136	231	538	10938	545	1266	
上流工程での不具合抽出比率が高い	80	34	180	405	821	12550	904	1903	
上流工程での不具合抽出比率が低い: 抽出比率が中央値72.8%以下									
上流工程での不具合抽出比率が高い: 抽出比率が中央値72.8%より大									

3.2.4. 信頼性変動要因の分析と活用

【概要】

分析項目	信頼性（発生不具合密度）の変動要因分析（開発プロセス関連）
メッセージ	<p>◇分析結果概要（読み取れる傾向）</p> <p>発生不具合密度が 0.02 件/KLSLOC 未満のプロジェクトを良群、0.02/KLSLOC 以上のプロジェクトを否群に分けて、両者に開発プロセス関連の信頼性変動要因候補（メトリクス欄参照）の値にどの程度の差異があるかを分析した。その結果、信頼性変動要因候補のいくつかによって、信頼性（発生不具合密度）が変動する傾向が見られる。</p> <p>具体的には、良群に次の傾向が見られる。</p> <ul style="list-style-type: none"> ・レビュー工数密度が高い ・上流工程での不具合摘出比率が高い ・テスト検出不具合密度が低い ・テスト検出能率（テスト項目当りの検出不具合数）が低い <p>◇助言</p> <ul style="list-style-type: none"> ・自組織の各ドメインにおいて自組織の定量データを用いて、信頼性変動要因を分析することをお勧めする。組織の信頼性向上方策の検討にあたっては、信頼性変動要因に着目して重点的に強化すると効果的な領域を特定し、その領域の信頼性向上方策を検討することをお勧めする。 ・IPA/SEC データベースに基づいた当分析結果からは、「レビュー工数密度を高めると信頼性が向上する」、「上流工程での不具合摘出比率を高めると信頼性が向上する」等から、設計及びレビューのプロセス改善が効果的と考えられる。 (例) レビュー工数密度を高める。 レビュー効率を高める（設計文書化密度を上げる、設計書作成標準やチェックリストを改良する等）
メトリクス	<p>◇発生不具合密度：稼働後（最長 6 ヶ月以内）に発生した不具合数 ÷ 開発規模 (KSLOC)。</p> <p>開発プロセス関連の信頼性変動要因候補として分析したものは次の通り。</p> <p>◇設計レビュー工数密度：基本設計から製作までの設計レビュー工数 ÷ 開発規模 (KSLOC)</p> <p>◇設計レビュー指摘密度：基本設計から製作までのレビュー指摘数 ÷ 開発規模 (KSLOC)</p> <p>◇上流工程での不具合摘出比率：基本設計から製作までのレビュー指摘数 ÷ (基本設計から製作までのレビュー指摘数 + 結合テストから総合テスト (ベンダ確認) までの検出不具合数)</p> <p>◇テスト密度：(結合テストケース数 + 総合テスト (ベンダ確認) ケース数) ÷ 開発規模 (KSLOC)</p> <p>◇テスト検出不具合密度：(結合テスト検出不具合数 + 総合テスト (ベンダ確認) 検出不具合数) ÷ 開発規模 (KSLOC)</p> <p>◇テスト検出能率：(結合テスト検出不具合数 + 総合テスト (ベンダ確認) 検出不具合数) ÷ (結合テストケース数 + 総合テスト (ベンダ確認) ケース数)</p>

考察／備考	IPA/SEC データベースにおいて業種を始め様々なプロファイルのプロジェクトデータが混在した中では、開発プロセス関連の信頼性変動要因候補（メトリクス欄参照）のいくつかによって信頼性が変動する傾向が見られる。各組織の同種のソフトウェア開発を行うドメインの中では、同様な傾向が見られるとは限らないが、より強い傾向を示す変動要因が見つかることと期待される。自組織の信頼性変動要因群を分析・把握することが重要かつ効果的である。
-------	--

【検索条件（共通分）】

- ◇ 開発 5 工程が揃っているプロジェクト
- ◇ 開発プロジェクトの種別が新規開発又は改良開発（改修・保守又は拡張）
- ◇ 開発規模（SLOC 規模） > 0
- ◇ 主開発言語が COBOL, C, VB, Java のいずれか
- ◇ 実績工数（開発 5 工程） > 0
- ◇ 稼働後の発生不具合数 >= 0

【分析方法】

- ◇ 発生不具合密度が 0.02 件/KLSLOC 未満のプロジェクトを良群、0.02/KLSLOC 以上のプロジェクトを否群に分けて、両者にどのような開発プロセスの差異があるかを分析する。
具体的には、良群と否群とで、信頼性変動要因候補（メトリクス欄参照）の値にどの程度の差異があるかを分析する。
(注) 0.02 件/KLSLOC は、発生不具合密度のほぼ中央値に相当する。
- ◇ 傾向が見られるかどうか（良群と否群に差があるかどうか）に関しては、原則として常用対数化した信頼性変動要因候補の数値に対して Welch の t 検定（分散が等しくないと仮定した 2 標本による平均値の差の t 検定）を行った上で記述する。ただし、データ分布の非対称性が弱いケース（歪度の絶対値が 2 未満のケース）が少数あるが、そのケースに対しては対数化しないデータを用いて検定する。

【分析結果】

良群に次の傾向が見られる。

- ・レビュー工数密度が高い（1%有意）
- ・上流工程での不具合摘出比率が高い（5%有意）
- ・テスト検出不具合密度が低い（5%有意）
- ・テスト検出能率（テスト項目当りの検出不具合数）が低い（5%有意）

<分析結果のサマリ表>

ドメイン等		開発規模	設計レビュー指摘密度	設計レビュー工数密度	上流工程での不具合摘出比率	テスト検出不具合密度	テスト密度	テスト検出能率
IPA/SEC 白書 新規開発	傾向	↓	→	↑	↑	↓	—	↓
	良群件数	250	58	48	56	203	196	188
	否群件数	232	70	61	65	197	192	186
IPA/SEC 白書 改良開発	傾向	↓	—	—	—	→	↑	↓
	良群件数	235	54	46	53	187	178	167
	否群件数	150	59	47	58	136	127	125

(凡例)

上記の表では、各メトリクスが示す良群の傾向を次のように矢印で表示している。

➤ 矢印の色 (または向き)

良群の平均値が否群の平均値より大きい場合には赤色 (上向き矢印)、小さい場合には緑色 (上向き矢印) で表示している。

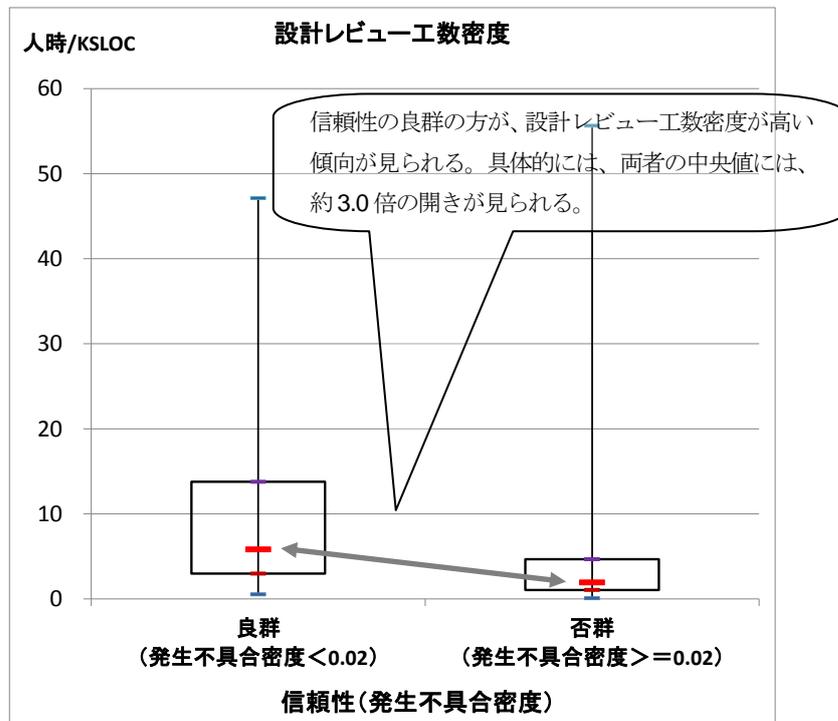
➤ 矢印の向き

良群、否群の平均値の差を Welch の t 検定を用いて検定した結果のうち、両側検定 P 値によって、次のように表示している。

↑ 及び ↓ は、1~10%有意であり、良群、否群の平均値の差が顕著であることを示している。

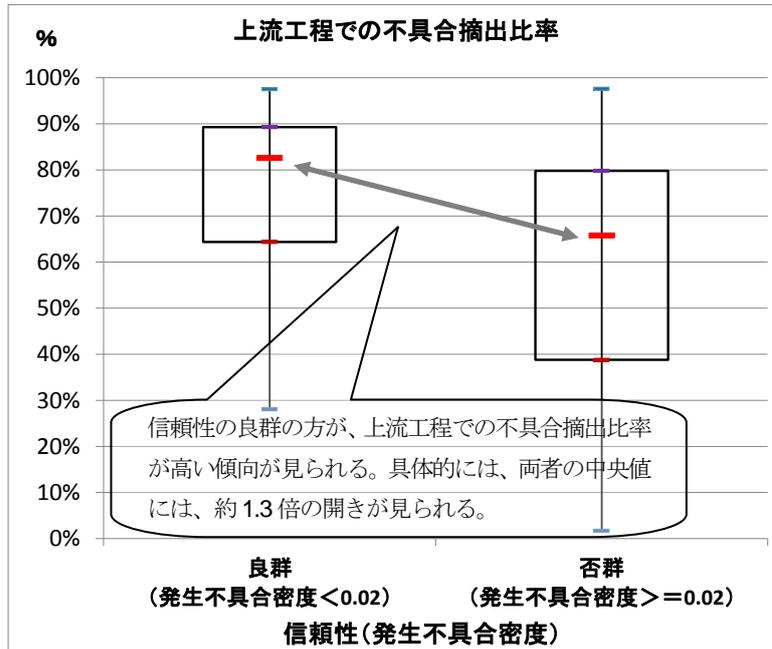
- ◇良群の方が指標値が大きい ↑ P<0.1 ↗ P<0.2 → P<0.5 — P>=0.5
- ◇良群の方が指標値が小さい ↓ P<0.1 ↘ P<0.2 ← P<0.5 — P>=0.5

<設計レビュー工数密度>



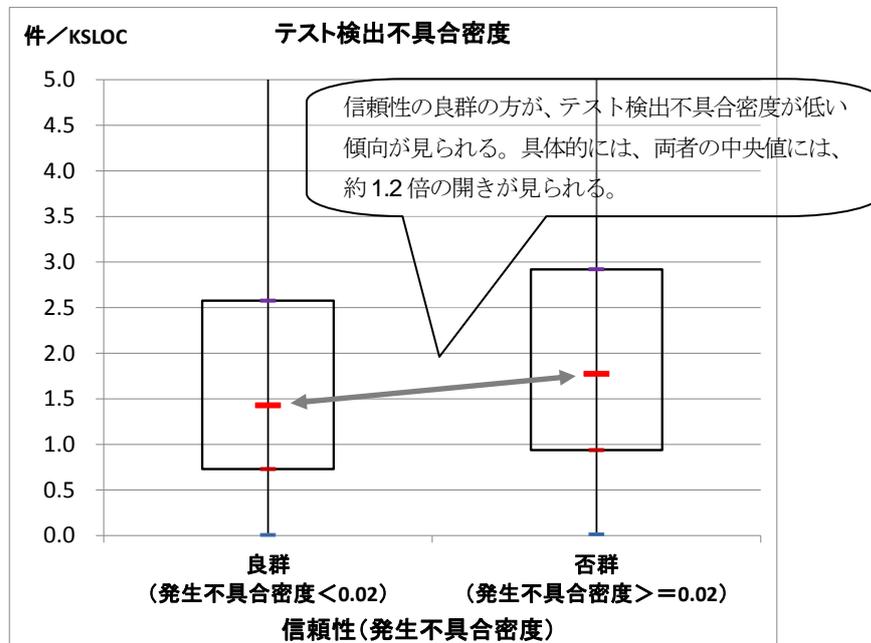
設計レビュー工数密度							(人時/KSLOC)	
信頼性 (発生不具合密度)	N	最小	P25	中央	P75	最大	平均	標準偏差
良群 (発生不具合密度 < 0.02)	36	0.54	2.97	5.81	13.77	47.12	9.44	9.39
否群 (発生不具合密度 >= 0.02)	41	0.08	1.05	1.93	4.67	55.63	5.20	9.80

<上流工程での不具合摘出比率>



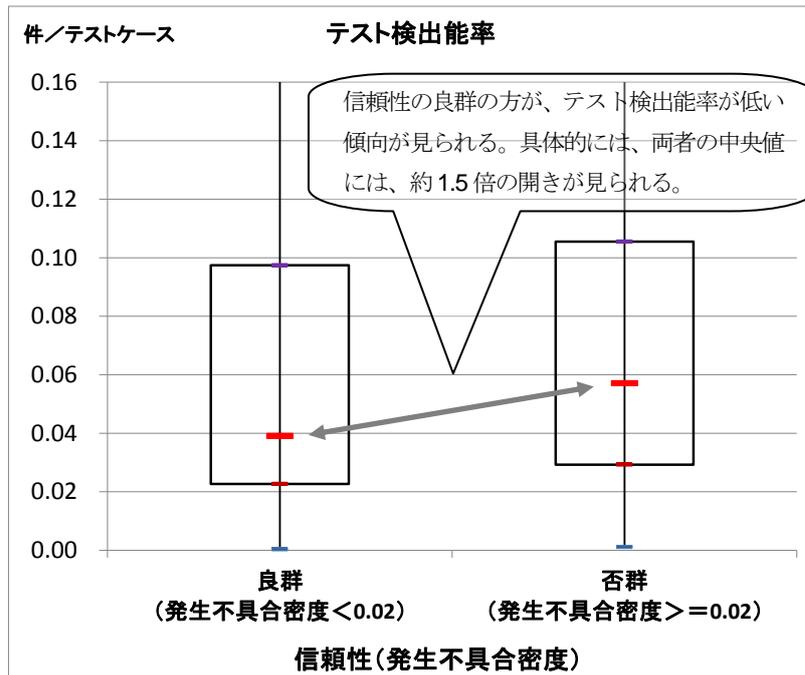
上流工程での不具合摘出比率								(%)
信頼性(発生不具合密度)	N	最小	P25	中央	P75	最大	平均	標準偏差
良群 (発生不具合密度 < 0.02)	45	0.0%	64.4%	82.6%	89.3%	97.5%	71.3%	27.9%
否群 (発生不具合密度 >= 0.02)	48	1.7%	38.8%	65.8%	79.8%	97.6%	58.3%	28.4%

<テスト検出不具合密度>



テスト検出不具合密度								(件/KSLOC)
信頼性(発生不具合密度)	N	最小	P25	中央	P75	最大	平均	標準偏差
良群 (発生不具合密度 < 0.02)	203	0.01	0.73	1.43	2.58	24.13	2.12	2.59
否群 (発生不具合密度 >= 0.02)	197	0.01	0.94	1.77	2.92	13.85	2.21	1.92

<テスト検出能率>



テスト検出能率							(件/テストケース)	
信頼性(発生不具合密度)	N	最小	P25	中央	P75	最大	平均	標準偏差
良群 (発生不具合密度<0.02)	188	0.000	0.023	0.039	0.097	2.801	0.104	0.258
否群 (発生不具合密度>=0.02)	186	0.001	0.029	0.057	0.106	0.776	0.091	0.105

4.トレンド編

主に IT 業界や国等の政策立案担当者の方々に向けて、施策検討の際の参考情報として利用して頂くことを期待して、信頼性及び生産性のトレンドに関する新規分析結果とメッセージを提示する。

4.1. 信頼性のトレンド

信頼性向上に向けたマネジメント活動や諸施策の総合的な効果は、信頼性の結果指標の経年変化（推移）を監視することによって確認できる。また、その推移を踏まえて、品質マネジメントの年度計画や中長期計画における信頼性目標を設定することができる。

本節では、ご参考のために、IPA/SEC のソフトウェア開発プロジェクトデータベースを用いた 2004 年から 2012 年までの新規開発における信頼性の経年変化を示す。ここで、信頼性とは発生不具合密度（開発規模（KSLOC）当りの稼働後に発生した不具合数）を言う。

【概要】

分析項目	発生不具合密度（信頼性）の経年変化（新規開発）
メッセージ	<p>◇分析結果概要（読み取れる傾向）</p> <p>IPA/SEC のソフトウェア開発プロジェクトデータベースにおいては、プロジェクトデータを収集開始してから最近まで（プロジェクトの実績終了年が 2004 年から 2012 年まで）の発生不具合密度の中央値の経年変化に低下傾向（信頼性向上の傾向）が見られる。</p> <p>具体的には、新規開発の場合、発生不具合密度の中央値が 0.02 件/KSLOC から 0 件/KSLOC に低下している。2011 年及び 2012 年の中央値は 0 件/KSLOC。ただし、発生不具合密度の 75 パーセンタイル値については、顕著な低下傾向は見られない。いまだに、ばらつきは小さくないと言える。</p> <p>◇助言</p> <p>上記の IPA/SEC データベースに基づく発生不具合密度（信頼性）の経年変化の分析結果を、次のようなシーンで参考にして頂けると幸いである。</p> <ul style="list-style-type: none"> ・信頼性向上に向けたマネジメント活動や諸施策の総合的な効果の確認 ・品質マネジメントの年度計画や中長期計画における目標と施策の検討等
メトリクス	<p>◇発生不具合密度：稼働後（最長 6 ヶ月以内）に発生した不具合数÷開発規模（KSLOC）。</p>
考察/備考	<p>信頼性変動要因に関連する IPA/SEC データベースの今までの分析結果から、次の方策も信頼性向上に寄与すると考えられる。ヒントとして参考にされたい。</p> <p>◇レビュー工数を増強する。</p> <p>目安としては、設計レビュー工数密度（設計レビュー工数÷開発規模）が 10 人時/KSLOC 程度となるよう増強する。</p> <p>◇上流（設計及び製作）工程での不具合摘出比率（上流工程での不具合摘出数÷開発完了までの不具合摘出総数）を高める。</p> <p>目安としては、上流工程での不具合摘出比率を 74%以上とする。</p> <p>設計レビュー効果を上げるために、設計文書化密度（設計書ページ数÷開発規模）を約 16 ページ/KSLOC 以上とする（新規開発の場合）。</p> <p>◇ユーザ担当者の要求仕様関与度合いを高めるよう、ユーザと調整する。</p> <p>◇要求仕様を早期に明確化するよう、ユーザと調整する。</p>

【検索条件】

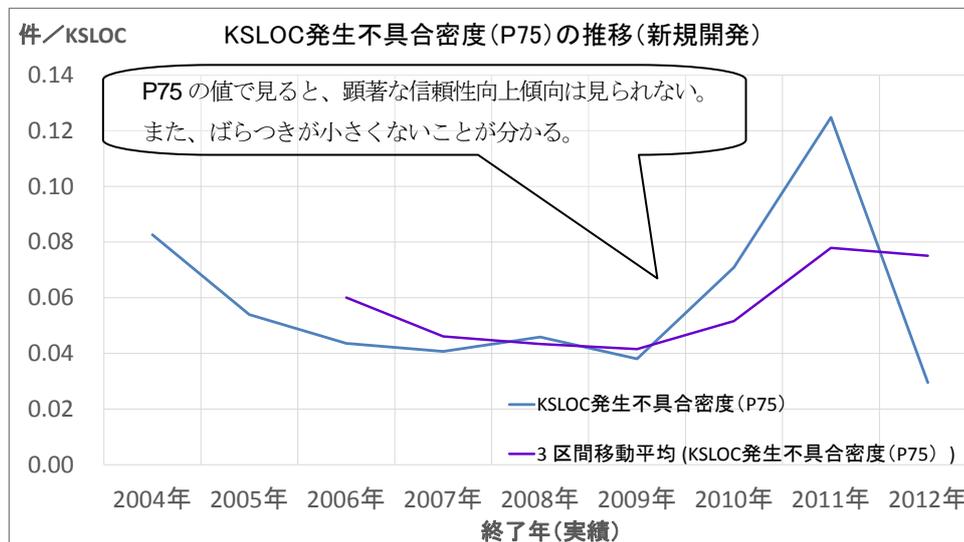
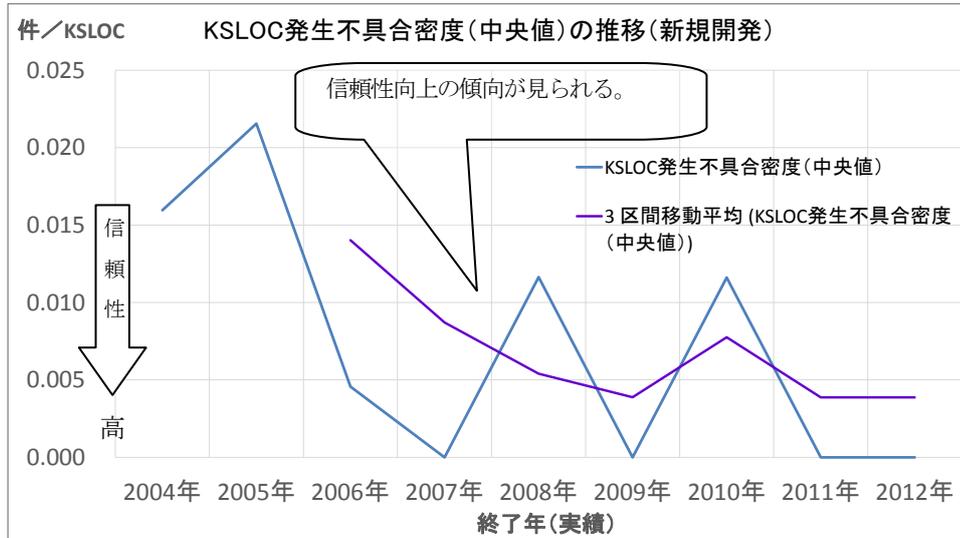
- ◇ 開発5工程が揃っているプロジェクト
- ◇ 開発プロジェクトの種別が新規開発
- ◇ 開発規模（SLOC 規模） >0
- ◇ 主開発言語が COBOL, C, VB, Java のいずれか
- ◇ 稼働後の発生不具合数 >= 0

【留意事項】

- ◇ 当分析結果は、種々の組織や業種等が混在したサンプルデータ集合から得られたものであることに留意されたい。また、稼働後の発生不具合数が記入されているという検索条件によって、データ数が少なくなったサンプルデータ集合から得られたものであることにも留意されたい。
- ◇ プロジェクトの終了年（実績）が 2013 年のデータについては、他の年のデータ数の半数程度のデータしか収集できていないので、今回の分析対象から除外している。
なお、次回のデータ白書用のデータを収集できた時点からは、2013 年に終了したプロジェクトのデータ数が他の年と同程度になるものと期待される。

【分析結果】

プロジェクトデータを収集開始してから最近まで（プロジェクトの実績終了年が 2004 年から 2012 年まで）の発生不具合密度の経年変化において、発生不具合密度の低下（信頼性の向上）の傾向が見られる。具体的には、発生不具合密度の中央値が 0.02 件/KSLOC から 0 件/KSLOC に低下している。2011 年及び 2012 年の中央値は 0 件/KSLOC。
ただし、発生不具合密度の 75 パーセンタイル値については、顕著な低下傾向は見られない。



KSLOC発生不具合密度								(件/KSLOC)	
終了年(実績)	N	最小	P25	中央値	P75	最大	平均	標準偏差	
2004年	28	0.0000	0.0000	0.0160	0.0826	0.3323	0.0656	0.0952	
2005年	53	0.0000	0.0000	0.0216	0.0539	0.8503	0.0695	0.1526	
2006年	54	0.0000	0.0000	0.0046	0.0436	0.7407	0.0469	0.1204	
2007年	47	0.0000	0.0000	0.0000	0.0407	2.4613	0.1222	0.3962	
2008年	42	0.0000	0.0000	0.0116	0.0459	0.6710	0.0668	0.1531	
2009年	42	0.0000	0.0000	0.0000	0.0380	0.4275	0.0407	0.0923	
2010年	29	0.0000	0.0000	0.0116	0.0709	1.0134	0.0870	0.1959	
2011年	31	0.0000	0.0000	0.0000	0.1248	0.3475	0.0594	0.0924	
2012年	39	0.0000	0.0000	0.0000	0.0295	0.4828	0.0348	0.0855	

4.2. 生産性のトレンド

生産性向上に向けたマネジメント活動や諸施策の総合的な効果は、生産性の結果指標の経年変化（推移）を監視することによって確認できる。また、その推移を踏まえて、生産性マネジメントの年度計画や中長期計画における生産性目標を設定することができる。

本節では、ご参考のために、IPA/SECのソフトウェア開発プロジェクトデータベースを用いた2004年から2012年までの新規開発における生産性の経年変化を示す。ここで、生産性とは SLOC 生産性（人時当りの開発規模（SLOC））を言う。

【概要】

分析項目	SLOC 生産性の経年変化
メッセージ	<p>◇分析結果概要（読み取れる傾向）</p> <p>IPA/SEC のソフトウェア開発プロジェクトデータベースにおいては、プロジェクトデータを収集開始してから最近まで（プロジェクトの実績終了年が2004年から2012年まで）の SLOC 生産性の中央値の経年変化に、生産性向上（または生産性低下）の傾向は見られない。</p> <p>ソフトウェア開発プロジェクトは、低価格と短納期の強いプレッシャーに晒されることが多々ある。また、予算管理や価格交渉等の場面で、年率5%の開発コスト削減や前年度比10%のライン単価低減等が要求されるケースが散見される。しかし、データによる裏付け等の根拠が希薄であると、このような状況は開発プロジェクトのリスク増大や品質低下を招く恐れがある。</p> <p>◇助言</p> <p>上記の IPA/SEC データベースに基づく SLOC 生産性の経年変化の分析結果を、次のようなシーンで参考にして頂けると幸いである。</p> <ul style="list-style-type: none"> ・生産性向上に向けたマネジメント活動や諸施策の総合的な効果の確認 ・生産性マネジメントの年度計画や中長期計画における目標と施策の検討 ・予算管理、単価の改定/交渉
メトリクス	◇SLOC 生産性：SLOC 規模÷開発工数（開発5工程）
考察/備考	<p>生産性に向上傾向が見られない要因の一つとして、次のことが考えられる。</p> <p>「複雑さ・難しさが増している一方で、開発要員のスキルレベルが追いついていない。開発プロジェクトにとっての相対的な難易度は高まっていることが考えられる。」</p> <p>少数精鋭部隊で開発していた時代は去っている。次のような状況下でも、如何にして要員のスキルレベルを確保/向上させるか、また効率的な開発方法を実現するかが課題の一つと考える。</p> <p>◇開発すべきソフトウェアの総量が増大するのに伴って、ますます大勢の開発要員が必要になっている（要員の資質の平均レベルが低下する）。</p> <p>◇大半の開発プロジェクトが改良開発（改修・保守や拡張）になって来ており、新規開発（特にスクラッチ開発）によってスキルアップするチャンスが少なくなっている。</p>

【検索条件】

- ◇ 開発5工程が揃っているプロジェクト
- ◇ 開発プロジェクトの種別が新規開発

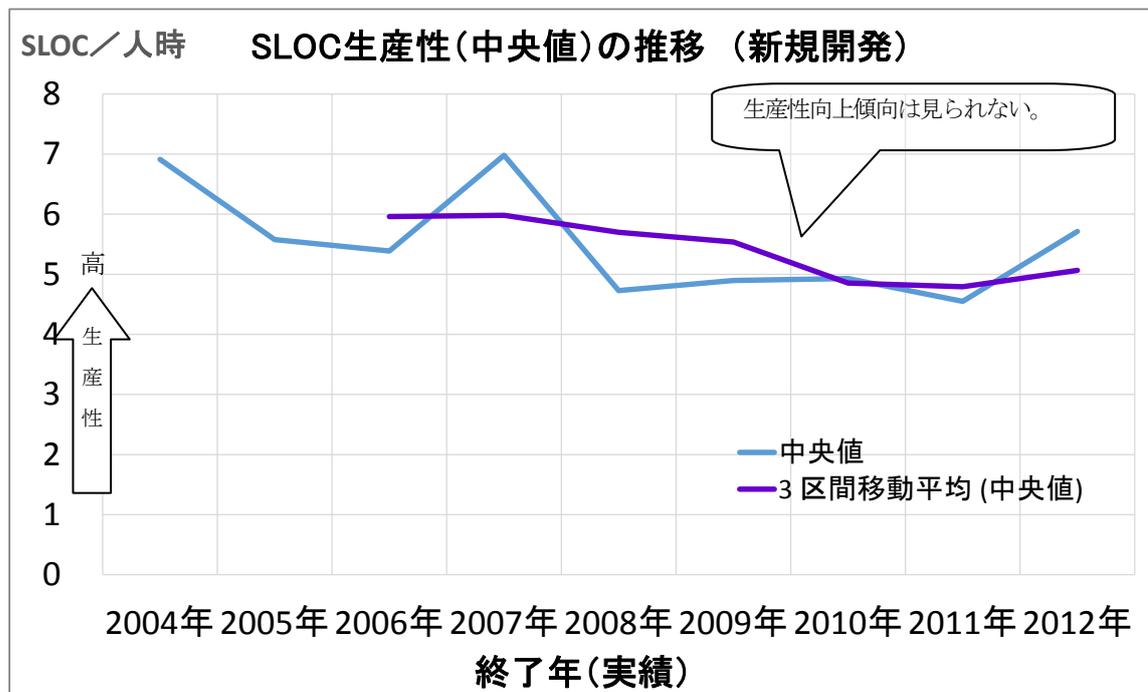
- ◇ 開発規模 (SLOC 規模) >0
- ◇ 主開発言語が COBOL, C, VB, Java のいずれか
- ◇ 実績工数 (開発 5 工程) > 0

【留意事項】

- ◇ 当分析結果は、種々の組織や業種等が混在したサンプルデータ集合から得られたものであることに留意されたい。また、SLOC 規模が記入されているという検索条件によって、データ数が少なくなったサンプルデータ集合から得られたものであることにも留意されたい。
- ◇ プロジェクトの終了年 (実績) が 2013 年のデータについては、他の年のデータ数の半数程度のデータしか収集できていないので、今回の分析対象から除外している。
 なお、次回のデータ白書用のデータを収集できた時点からは、2013 年に終了したプロジェクトのデータ数が他の年と同程度になるものと期待される。

【分析結果】

SLOC 生産性の中央値には、顕著な生産性向上 (または生産性低下) の傾向は見られない。



SLOC生産性								(SLOC/人時)
終了年(実績)	N	最小	P25	中央値	P75	最大	平均	標準偏差
2004年	76	0.73	4.46	6.91	10.55	325.24	14.37	40.62
2005年	71	0.53	3.59	5.58	8.88	29.49	7.11	6.06
2006年	67	0.90	2.92	5.39	9.78	33.49	7.69	6.57
2007年	58	0.71	3.03	6.98	11.74	82.58	8.99	11.28
2008年	53	0.36	2.91	4.73	8.71	35.27	6.92	6.61
2009年	52	0.65	3.67	4.90	8.10	42.43	6.74	6.55
2010年	49	0.84	3.21	4.93	8.97	52.74	9.70	12.09
2011年	32	1.51	3.43	4.55	7.04	76.12	11.68	19.32
2012年	40	2.17	3.40	5.71	9.95	23.70	7.05	4.91