

ソフトウェア開発データの分析に基づく エンジニアリング研究の推進

～収集データの活用に向けたIPA/SECの取組み～

函館大学

FIT 2011

2011年9月8日

情報処理推進機構 (IPA)

技術本部ソフトウェア・エンジニアリング・センター (SEC)

山下博之・秋田君夫

IPA/SECでは、2004年の発足以来毎年、**ソフトウェア開発プロジェクトの基本要素(規模, 工数, 工期や信頼性等)の実績データ**をプロジェクト情報と合わせて収集すると共に、それらの分布や要素間の関係等の分析を行い、「**ソフトウェア開発データ白書**」として発行してきている。

この取組みは、**ベンチマーキングの枠組み国際標準 (IS29155-1)**に基づいており、現在**2500プロジェクト**を越えるデータが蓄積されている。この度、ソフトウェア品質向上につながる新たな分析手法の検討等を目的に、**大学等外部機関との共同によりこれら蓄積データの有効活用**を図る仕組みを確立した。

本講演では、これら取組みの概要とデータ活用事例を紹介する。大学等のみなさんには、是非、ソフトウェア開発定量データの利用をご検討頂きたい。



- ▶ 1. ソフトウェア開発データの収集・蓄積の取組み
- 2. ソフトウェア開発データの活用
- 3. データを活用した共同研究
- 4. まとめ

「ソフトウェア開発データ白書」には、**工数と工期、規模との関係、規模と発生不具合数との関係、生産性等のデータを収録している**。2010-2011年版は、**23社から提供を受けた2584件のプロジェクトデータを含む**。

プロジェクトマネージャは、これら各種のデータを組み合わせたり自プロジェクトのデータを重ねたりすることにより、生産性や工期、品質、工程等の情報をさまざまな切口で見ることができ、**プロジェクト計画立案時や評価時の参考にすることができる**。また、これらのデータの分析により、ソフトウェア開発の傾向を把握し、**各種施策に反映することができる**。

さらに、将来的には、対象のソフトウェア／システムの特性や開発組織・プロジェクトのプロファイル、開発方法の情報等と共に分析することにより、**システム障害の要因究明や再発予防方法の発見**が期待される。

ソフトウェア開発データの内容と用途

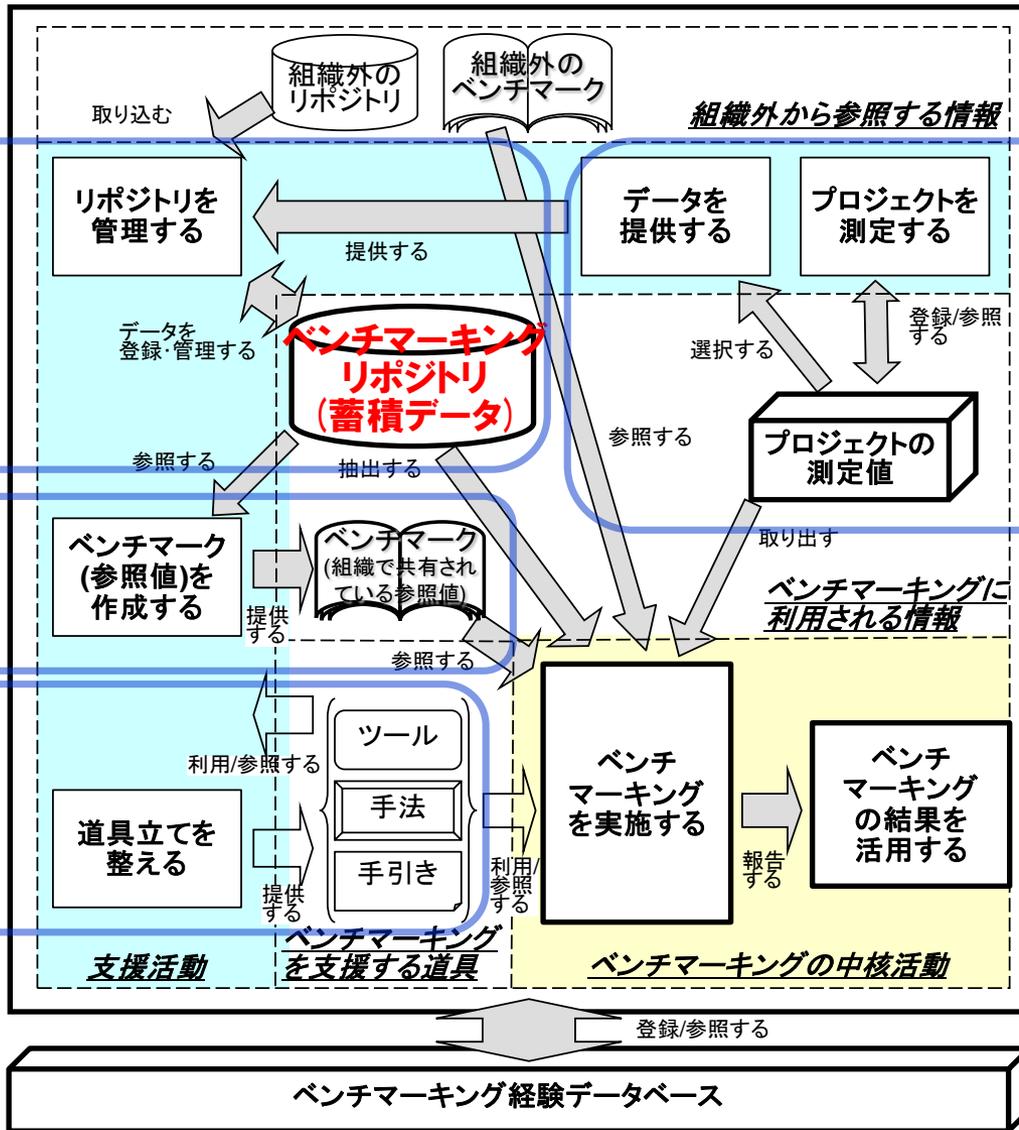
項番	区分	入力項目	狙い
1	生産性 水準	規模 (FP/SLOC)、 工数、業種、言語、 アーキテクチャ (システム構成)	工数当たりの開発規模の分布において、自社プロジェクトの位置を、規模の大きさ、業種、言語、アーキテクチャ等の層別に評価する。
2	工期 水準	工数、工期	「工期は工数の三乗根に比例する」傾向を活用し、自社プロジェクトの工数と工期が適正な状態にあるかどうかを評価する。
3	品質 水準	検出バグ数 (結合、 総合テスト分) 不具合数 (稼働後) 規模 (FP/SLOC)、言語	開発時の検出バグ数および稼働後の発生不具合密度の分布において、自社プロジェクトの位置を工程、言語等の層別に評価する。
4	工程 比率	工程別の工数と 工期	工程別の工数比の分布および工期比の分布より、標準的な工数比および工期比を参考に、自社プロジェクトの工程別工数、工期の配分をシミュレートする。

ベンチマーキング枠組みモデル

国際標準(ISO/IEC 29155-1)に基づく

＜IPA/SECの取組みとの対応＞

- ① データの収集・蓄積と機密室でのセキュアな管理
- ② データ白書の出版, 大学との共同研究, 等
- ③ プロジェクト診断ツール, データ活用ガイド, 等
- ④ データ測定・提供支援ツール



1. ソフトウェア開発データの収集・蓄積の取組み

➡ 2. ソフトウェア開発データの活用

3. データを活用した共同研究

4. まとめ

活用事例(1)

入力した利用者のプロジェクトデータを、画面上に表示したデータ
白書掲載のグラフ上に重ねてプロットできる「**プロジェクト診断支
援ツール**」を公開。

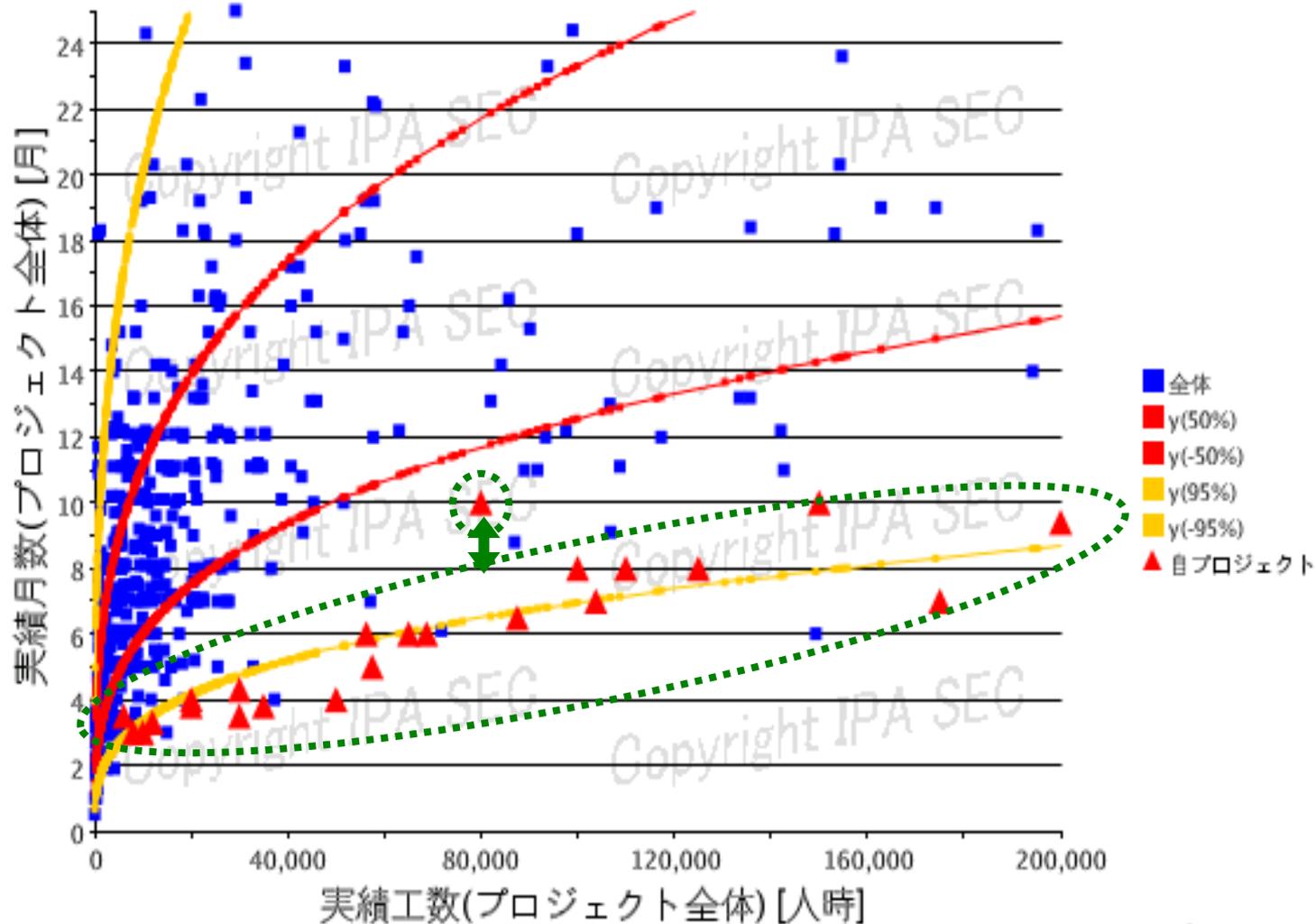
→他社データとのベンチマーキングが可能。

→ユーザ・ベンダ間での目標設定、プロジェクト状況の把握、プロ
ジェクト特性や立ち位置の確認を、定量的かつ客観的に可能。

活用事例(2)

「ソフトウェア開発データ白書」の具体的な利用方法・事例と、この
ような定量データ一般の具体的な活用法・事例を説明した、「**データ
白書の見方と定量データ活用ポイント**」を発行。

事例(1) 工数・工期配分の妥当性の確認

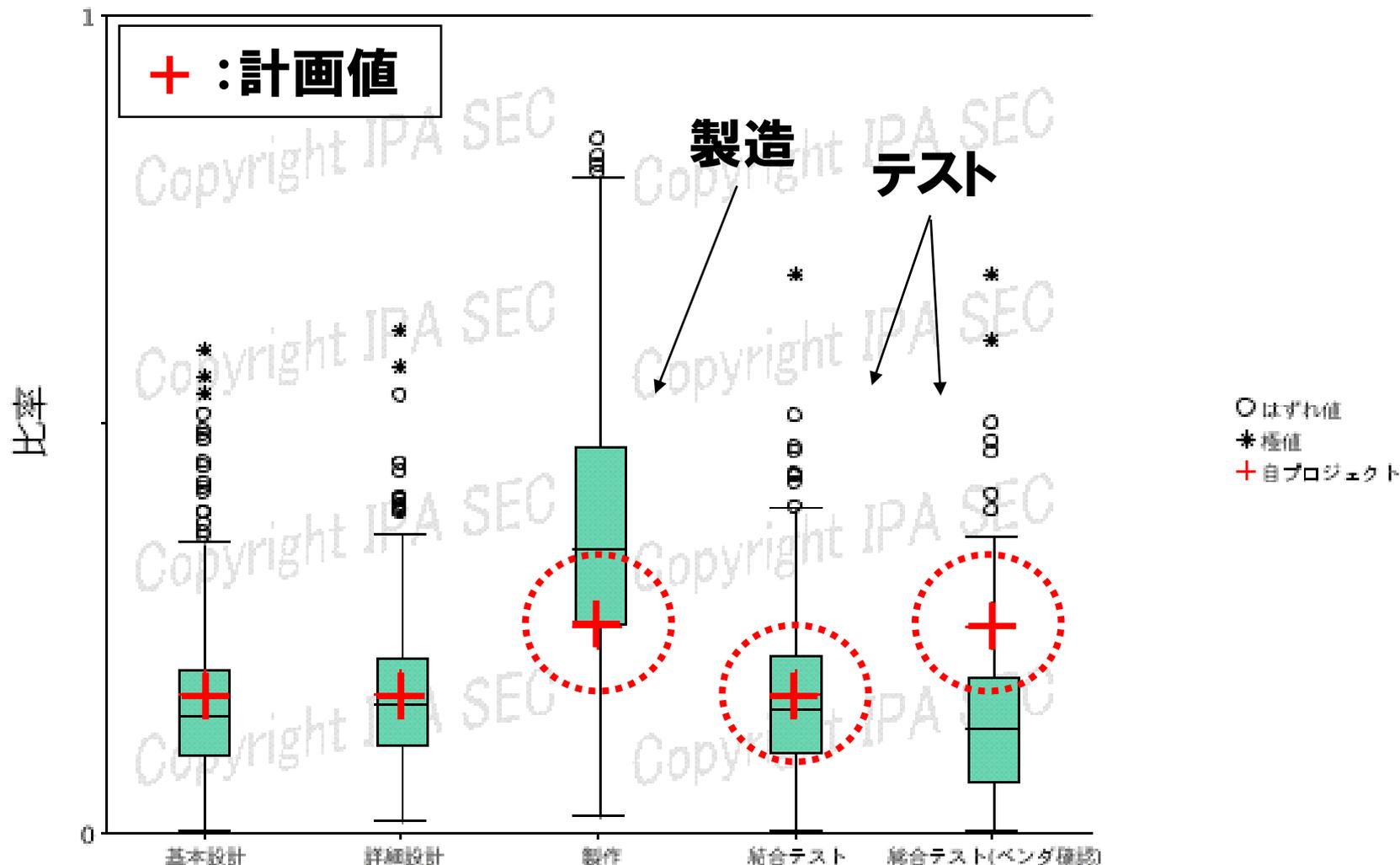


N=506

Copyright IPA SEC

「プロジェクト診断支援ツール」の利用例

事例(2) 工数・工期配分の妥当性の確認



Copyright IPA SEC

出典:「データ白書の見方と定量データ活用ポイント」

高度な分析手法の研究等，一層の有効活用を図るため，
大学等の公的研究機関に対して
一定の守秘条件の下で開示する仕組みを整備。
この仕組みに基づく研究が進行中。

事例

- ◇「ヒューマン・ファクタが開発にどう影響するか」というテーマでの定量データの分析
- ◇規模と工数等の測定量間の関係分析

1. ソフトウェア開発データの収集・蓄積の取組み

2. ソフトウェア開発データの活用

 3. データを活用した共同研究

4. まとめ

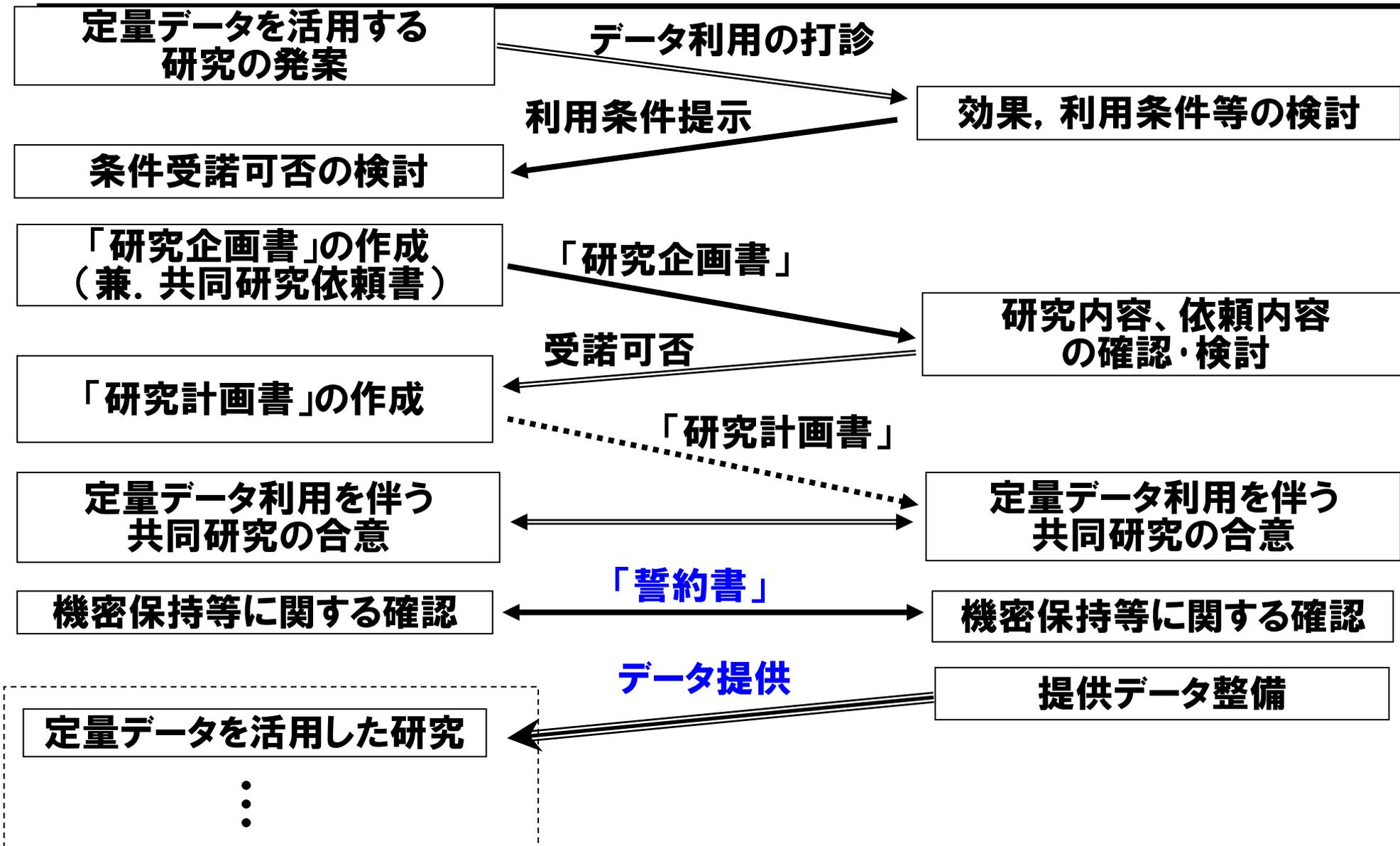
データの開示は、

データ提供元とIPA/SECとの機密保持契約の範囲内で、
研究機関とIPA/SECとの「共同研究」の枠組みで実施。

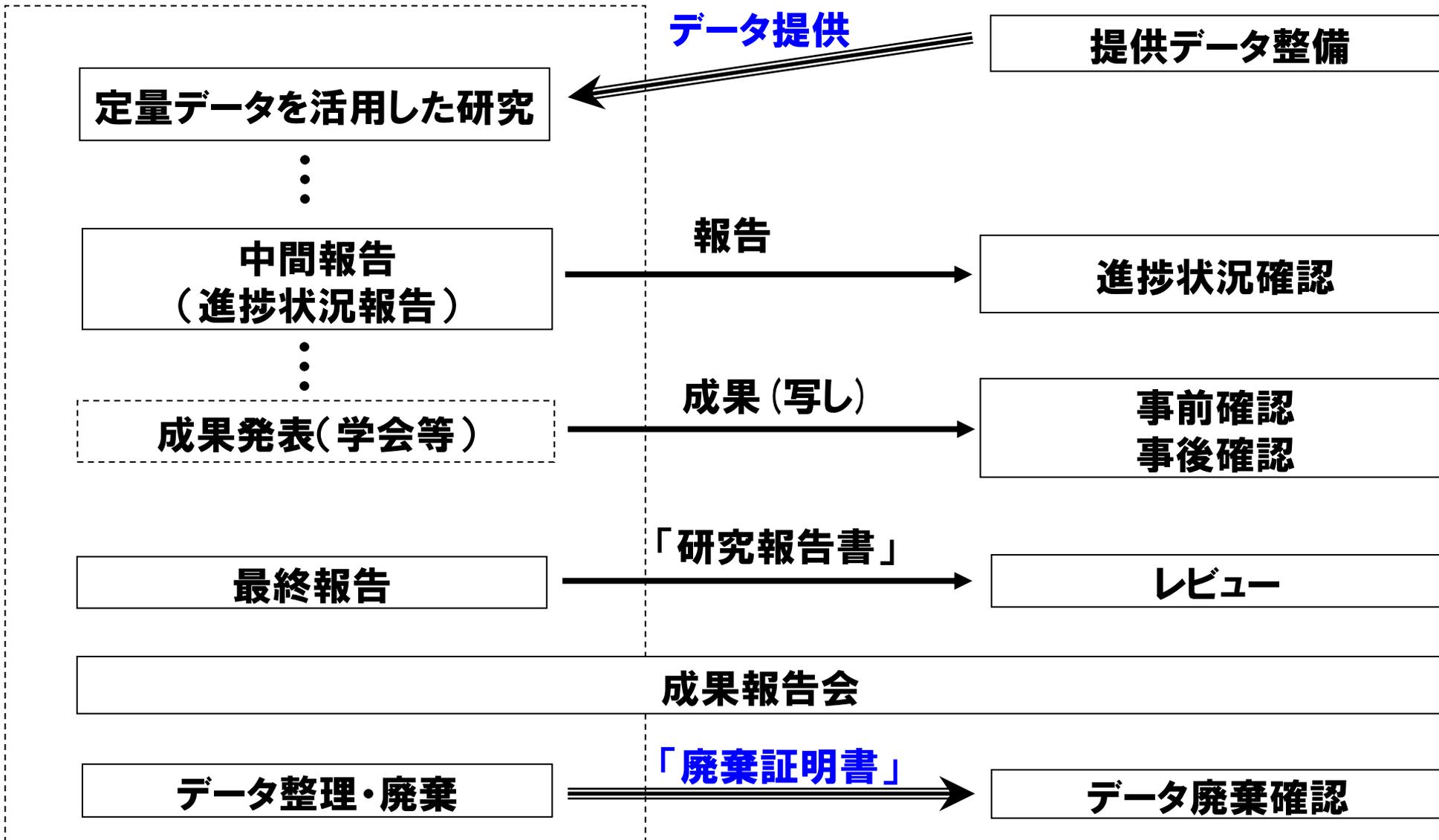
主要な利用条件：

- (1) IPA/SEC事業への展開
- (2) 成果の公表
- (3) データの匿名性確保
- (4) データのセキュリティ

「データ活用」共同研究のフロー(1/2)



「データ活用」共同研究のフロー(2/2)





対数正規分布に従う測定量の回帰分析例(東海大学)



p.20以降のデータ類参照

1. ソフトウェア開発データの収集・蓄積の取組み
2. ソフトウェア開発データの活用
3. データを活用した共同研究
-  4. まとめ

IPA/SECでは、2004年の発足以来毎年、**ソフトウェア開発プロジェクトの基本要素(規模, 工数, 工期や信頼性等)の実績データ**をプロジェクト情報と合わせて収集・蓄積している。

この度、ソフトウェア品質向上につながる新たな分析手法の検討等を目的に、大学等外部機関との**共同によりこれら蓄積データの有効活用**を図る仕組みを確立した。

IPA/SECは、ソフトウェア開発データの分析に基づく研究が一層促進され、産業界にとって有益な成果が生み出されることを期待している。その結果として、企業から開発データを提供して頂くことが困難になってきている状況がいくらかでも打開されることを願っている。

大学等のみなさんには、是非、利用をご検討頂きたい。

ご清聴，ありがとうございました

「ソフトウェア開発データ白書2010-2011」公開中

<http://sec.ipa.go.jp/publish/tn10-002.html>

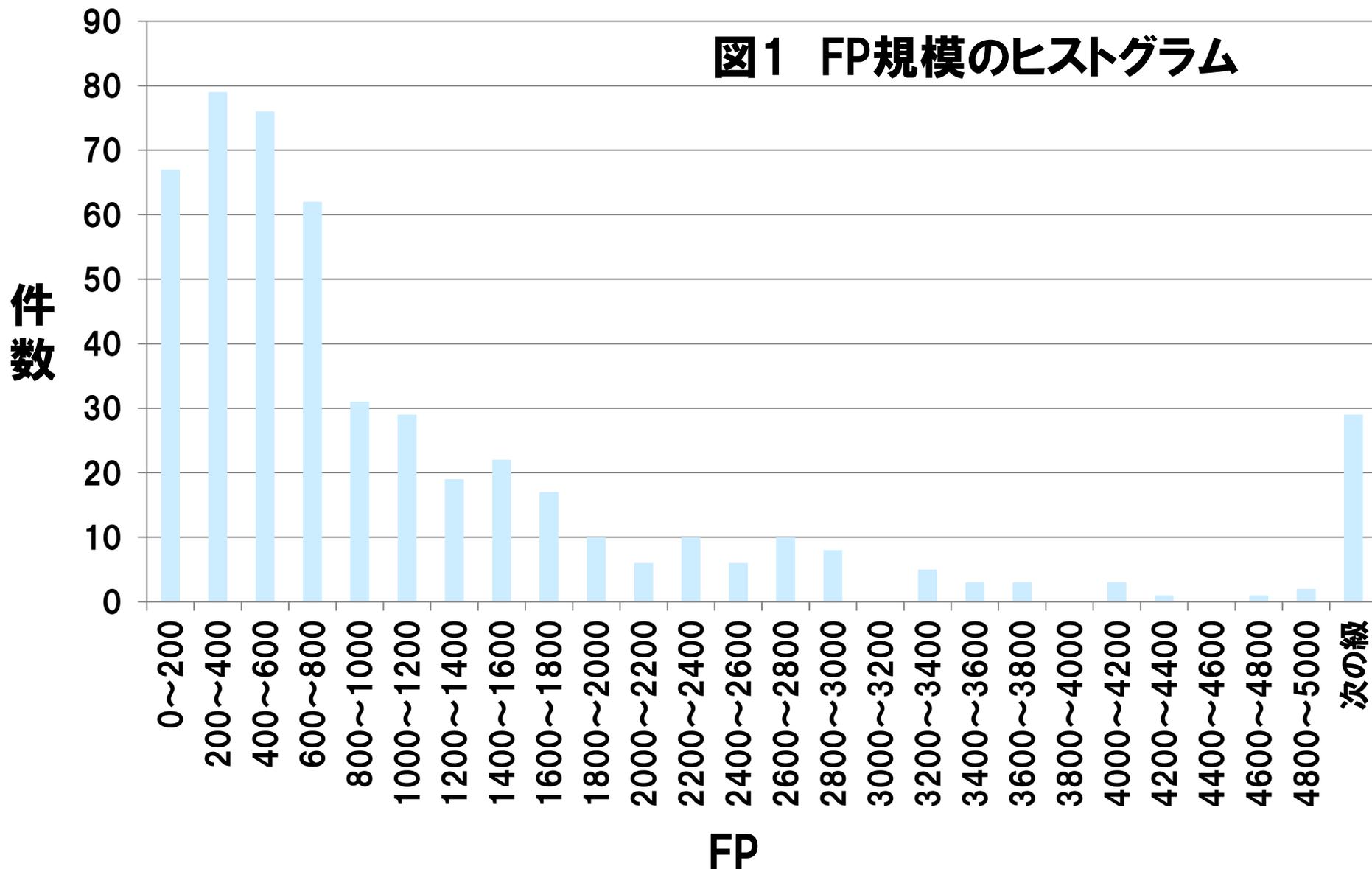
<http://sec.ipa.go.jp/publish/index.html#ent>

IPA/SECホームページ:

<http://sec.ipa.go.jp/index.html>

以降, データ類

図1 FP規模のヒストグラム



対数変換後のFP規模のヒストグラム

図2 対数変換後のFP規模のヒストグラム

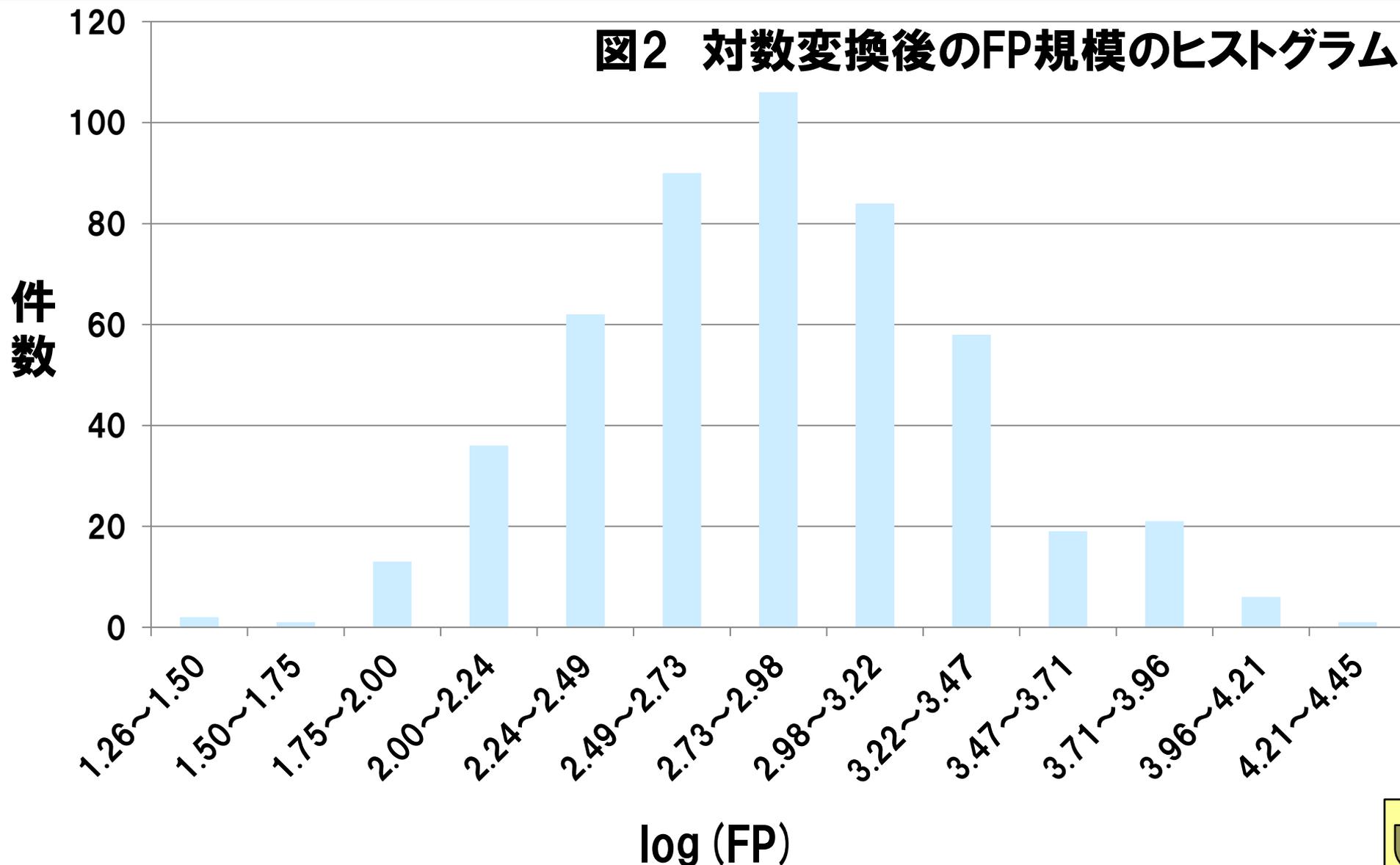
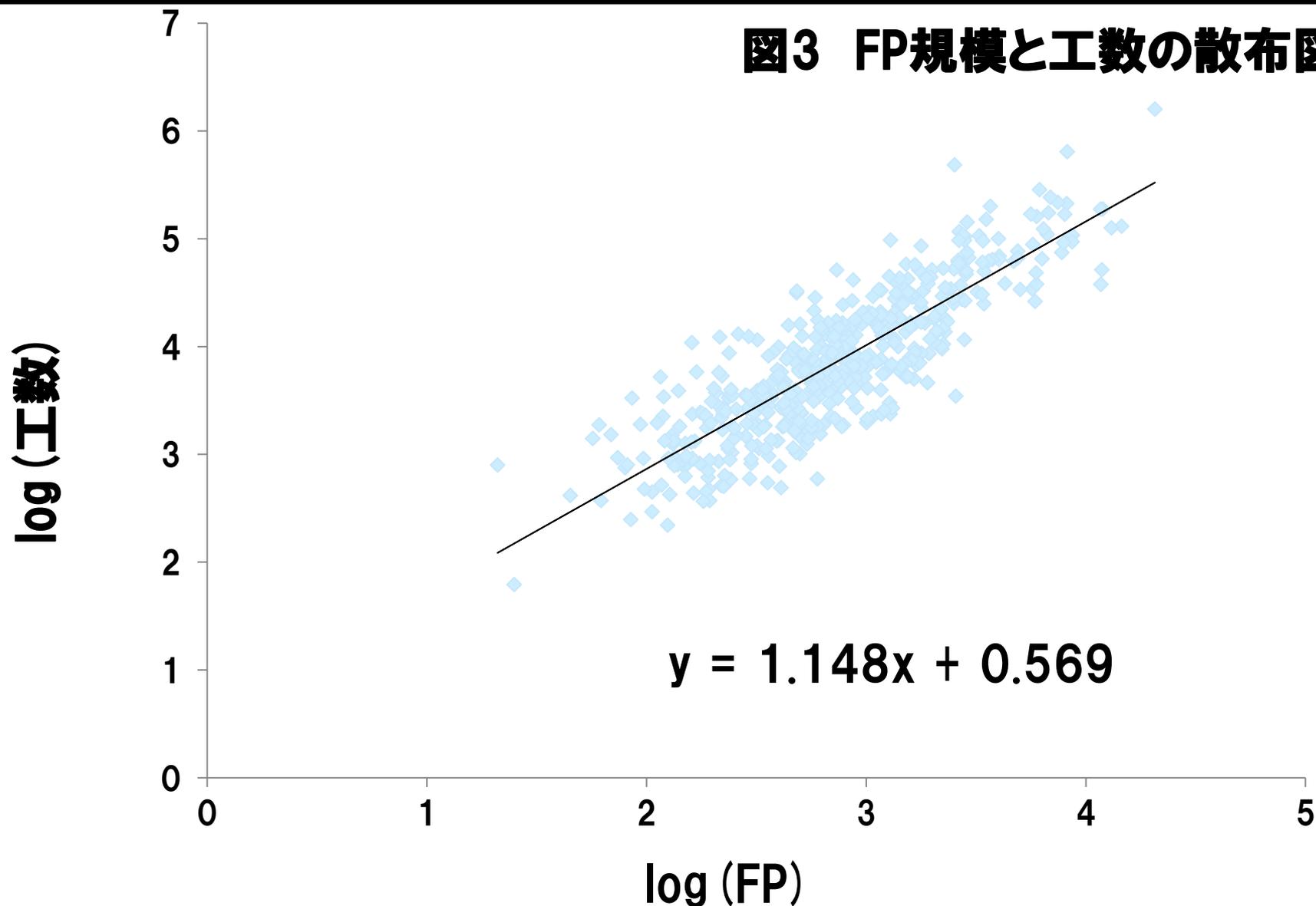
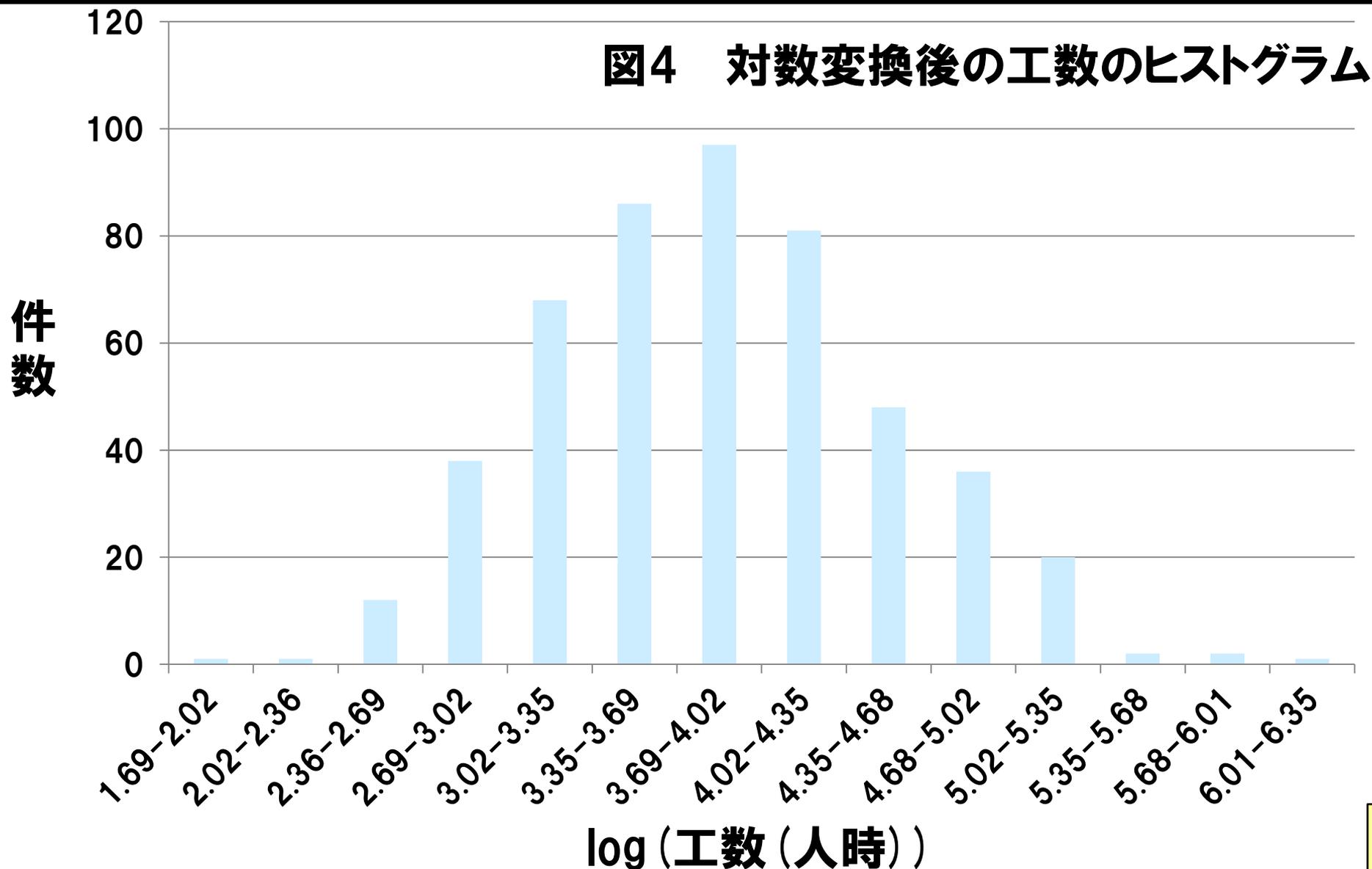
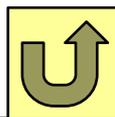
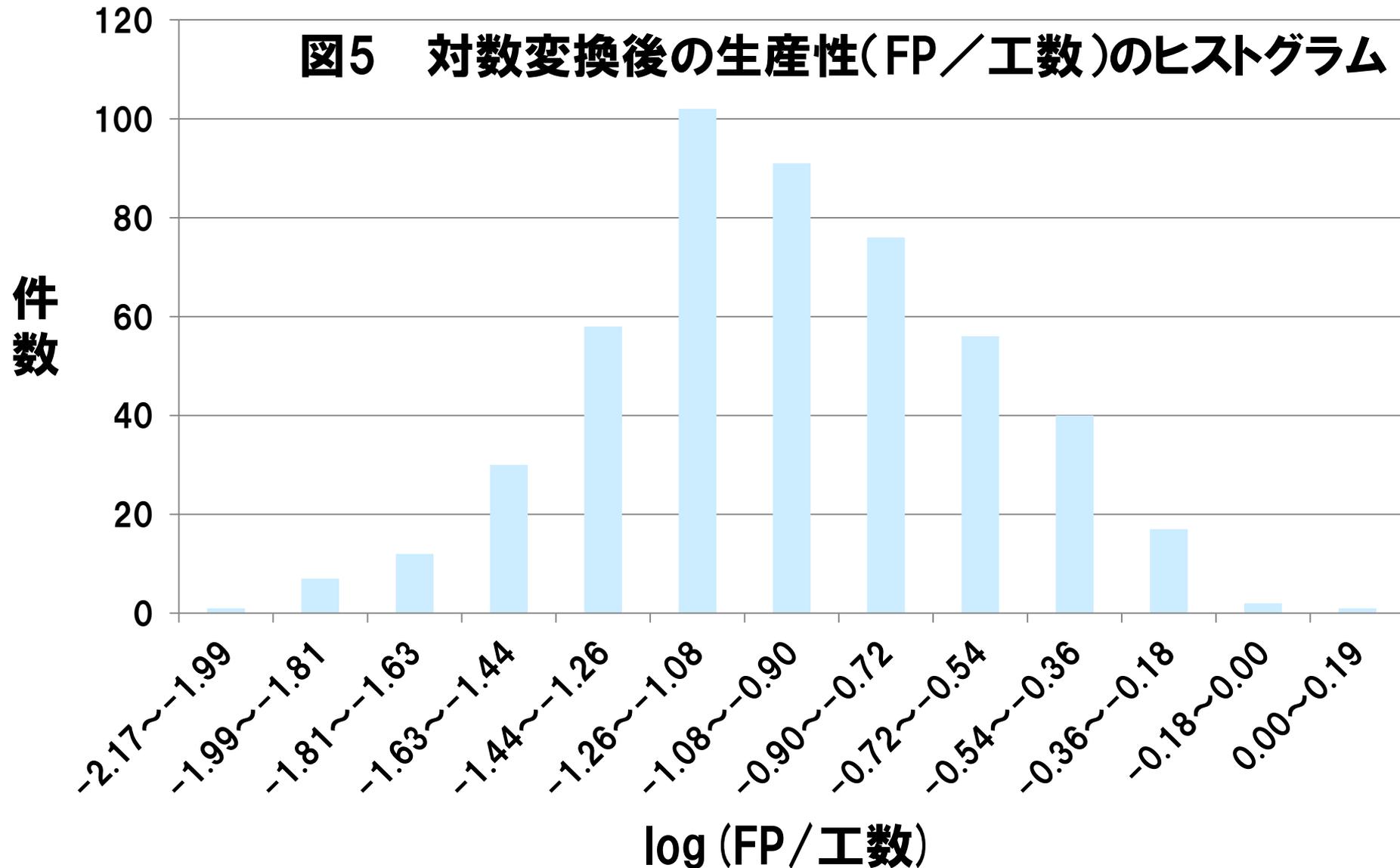


図3 FP規模と工数の散布図



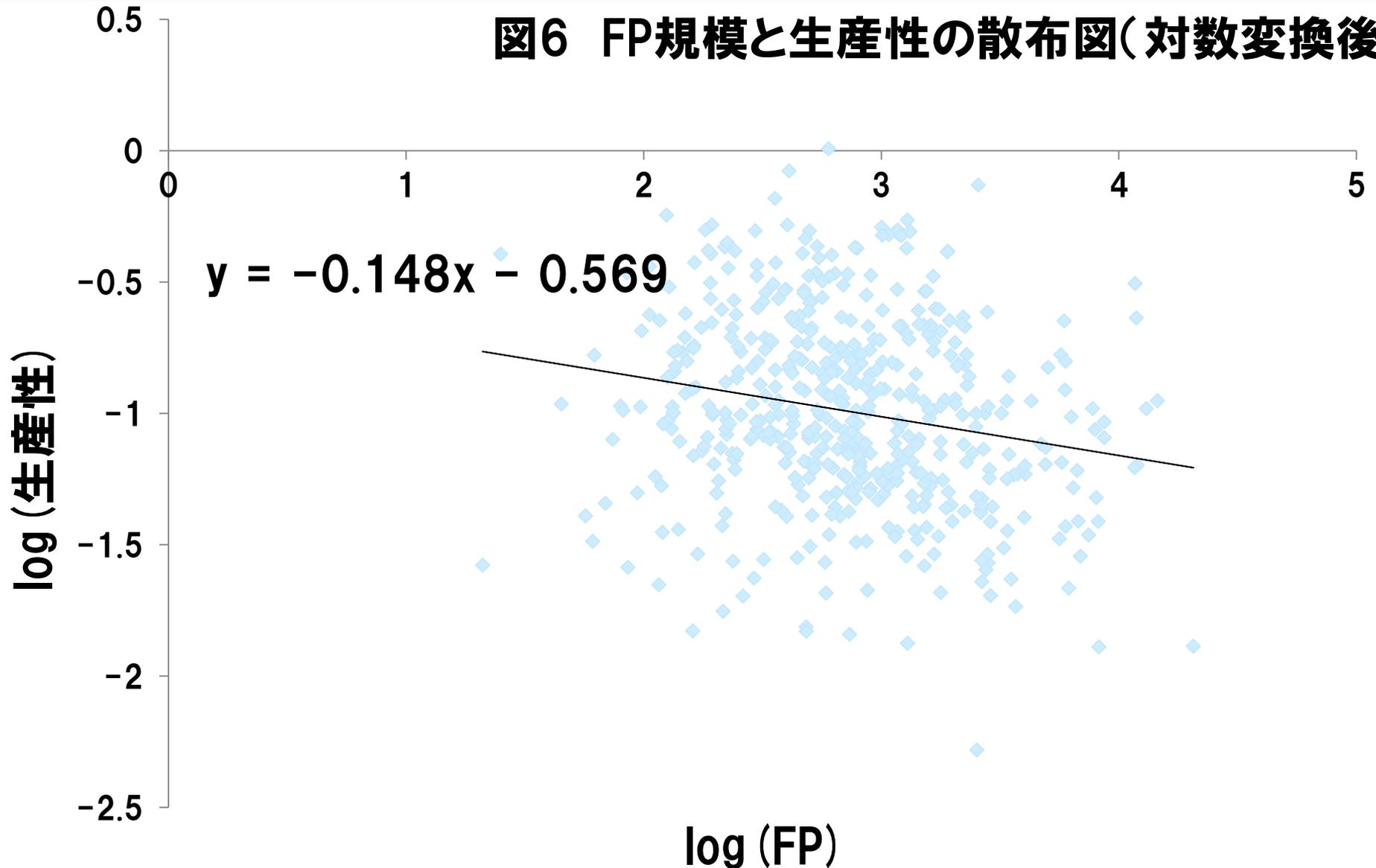


対数変換後の生産性(FP/工数)のヒストグラム



FP規模と生産性の散布図(対数変換後)

図6 FP規模と生産性の散布図(対数変換後)



Q1. 共同研究の対象として、企業は含まれないのか？[東芝]

A1. 提供頂いているデータは企業の機微なデータであるため、現時点では、制約を設けている。

**Q2. たとえば、要件定義の内容等、深いデータは収集していないのか？
研究者としては、そちらの方に興味がある。[座長(東京女子大)]**

A2. 現時点では、数値データと簡単なプロフィール情報のみを収集している。これをトリガとして、個別の企業との共同研究等に進めばよいと考える。

Q3. データを提供しているのは余裕のある企業と思われ、データに偏りがあるのではないか？[仙台高専]

A3. 確かに、現状では、大手ベンダからの提供が中心となっている。別の事業として、データ収集を自動化するツールの開発を進めている。