

定量化の壁を破る

平成21年9月10日

東海大学

古山恒夫

私の立場

ソフトウェア工学の
知識 / ソフトウェア
開発経験

筆者

統計学的
知識

データ分析
が好き



定量化の壁はどこにあるか？

- ・定量化の意義がわからない(Why?)
 - ――開発で手いっぱいだ
 - ――測定や分析の時間がもったいない
開発の邪魔になる
- ・何をどのように測定すればよいかわからない(What & How?)
- ・測定データをどのように分析して利用すればよいかわからない(How?)
 - ――統計学って何？ ごまかしじゃないの？
 - ――分析結果なんて現場の感覚と違うよ



今日の話題

1. 定量化の目的と効果
 - ・・・定量化はなぜ重要か
2. 定量化の枠組み
 - ・・・何をどのように測定するか
3. データの分析方法
 - ・・・測定結果の分析と利用方法
4. データの分析事例



定量化の目的と効果^[3]



定量化の目的

現状を正確に把握

勘と経験の補強

定性モデルを数値による詳細化へ

- ・プロジェクトの計画作成と評価の助け
- ・計画に対する実績の客観的な追跡
- ・プロセス改善の判断と投資の指針
 - 生産性の向上／コスト削減
 - スケジュール短縮
 - 品質向上



定量化の(副次的)効果

- ・問題の早期特定と対応:客観的なリスクの特定
／問題の客観的な評価・優先順位付け
- ・重要なトレードオフの判断:コスト、スケジュール、機能、
品質、性能のトレードオフの判断を支援
- ・情報交換の効率化:組織に客観的な情報を提供
／受注側と発注側の意思疎通の促進
- ・判断の正当化:経営者・管理者に対する見積りや計画の
根拠及び最良の選択の有効な根拠を提示

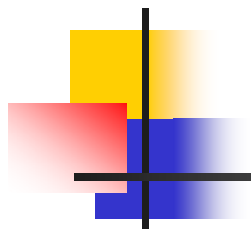
閑話休題

プロジェクトデータ税金論

データの収集は開発現場にとって負担

効率的なプロジェクト運営にはデータは必須
(税金徴収のようなもの)

- ・税金が軽く、結果が現場にフィードバックされるなら、「納税者」も納得
- ・目的の明確なデータを収集し、結果を現場にフィードバックすることが大切
- ・最初はSmall startでは始めるのがうまくいく秘訣



定量化の枠組み^{[1][2][3]}



コミットメント(確約)の確立と維持^[3]

測定成功の鍵

組織的なコミットメントと責任の明確化

- 経営者と管理者のコミットメント
- すべての関係者に適切な訓練を実施
- 測定に基づいてコミュニケーションを促進
- 関係者による有効性の理解



測定の計画作成^[3]

- ・測定プロセスは構造的で再現性のあるものに
- ・きちんとした測定計画を立案することが大切

測定計画の立案順序

- (1) 測定のニーズの特定と優先順位付け
 - 何のために測定するか
- (2) 測定量の選択と特定
 - 何を測定するか
- (3) プロジェクトプロセスへの統合



測定のニーズ

- まずは現状（ベースライン）分析を最初の目的に
- 次に影響要因の分析
- さらに改善効果の測定へ
 - 生産性の向上／コスト削減
 - スケジュール短縮
 - 品質向上
- そして計画作成の支援とプロジェクト実績の追跡へ



現状分析のためのテーマ例^[1]

- ・生産性
 - － すべての分野とタイプ毎の生産性
 - － 規模による生産性の比較
 - － 言語／ツール／手法による生産性の比較
- ・品質
 - － すべての分野とタイプ毎の品質
 - － 規模による品質の比較
- ・スケジュール
 - － すべての分野とタイプによる開発期間
 - － 規模による開発期間の比較
- ・規模
 - － 企業の保有するソフトウェア規模の合計
 - － 分野とタイプ毎のソフトウェア規模の合計
- ・顧客満足度
 - － ソフトウェアの使いやすさと顧客満足度の測定



現状分析の留意事項^[1]

- 現状調査で得られる結果は長所と短所
- 現状調査で得られない結果は
即効的な生産性の改善
- 対策を講じるのに1ヶ月から1年が必要



ベースライン分析の手順例^[1]

- ① 経営陣の後押し
- ② 定量化チームの選定
- ③ データ収集手段の設計
- ④ 予備検討とプロジェクト選択
- ⑤ プロジェクトインタビューのスケジューリング
- ⑥ 個々のプロジェクト分析
- ⑦ データの集計と統計分析
- ⑧ ベースラインプレゼンテーションの作成
- ⑨ 年次ベースライン報告書の作成



用語について

- 測定量 (measurement)
- 測定方法 (measurement method)
- 測定法 (metrics) = 測定量 + 測定方法
- 尺度 (scale)
 - 比尺度: 規模、工数、工期など
 - 間隔尺度: (プロジェクトデータにない)
 - 順位尺度: PMスキルレベルなど
 - 名義尺度: 開発方法
(OO分析法・構造化分析法など)



基礎データ項目の収集と留意事項

量的データ

- ・規模データ(FPとLOC)
- ・工数データ(WBSに沿った工数把握を)
- ・工期データ(始めと終わり)
→組織内で統一されていればよい

質的データ

- ・影響要因データ(SECのデータ白書を参考)



提言に含むべき情報^[3]

- ・プロジェクトの全体的な評価
- ・具体的な問題、リスク、不足している情報の特定
- ・特定した根本的なリスクと、問題を解決するための代替アクションについての提言
- ・利点と欠点を併記して記述
- ・想定できる新しい課題



測定終了後に行う 測定の評価

- 測定成果物、測定プロセスそのものの評価
- 経験データベースの更新
- 改善事項の決定と実施



データの分析方法^[4]



プロジェクトデータの問題点と特徴

- (1) データが収集しにくい
- (2) 欠損データが多い(ところどころにデータがある)
- (3) データの精度が低い
- (4) 人間の主観に基づくデータが多い
- (5) データは非負の非対称な分布をなすことが多い
- (6) 規模が主要な説明変数である
- (7) 影響要因が規模と相関を持つことが多い

閑話休題

中央値のススメ

- 測定データの「中心」を特徴付ける統計値
 - (算術) 平均値
 - 中央値 (大きさの順で真ん中に位置する値)
- 非対称な分布をするデータでは「平均値」≠「中央値」
- 平均値は外れ値に影響を受けやすい
- 中央値は外れ値の影響をほとんど受けない
- 中央値は最も典型的なプロジェクトのデータ



プロジェクトデータでは「平均値」よりも「中央値」を



統計学の使い方

- 仮説や直感を科学的に裏付ける手段
- 時には仮説や直感の誤りを指摘することも
 - … 経験的には概ね仮説の7割が合致
- データの精度に見合ったレベルの統計分析手法
 - … ソフトウェアプロジェクトデータの分析には大学の教養レベルの統計学で十分
- 最も事実を適切に説明する分析手法を選択
 - … 経験と良心が必要



データ分析における留意事項

- (1) データの収集と分析の目的を明確に
- (2) 分析の前にデータ群の特性の把握を
- (3) 適切な層別を
- (4) 必要なら変数変換を
- (5) まず、2変数間の分析を
- (6) なるべく基本測定量を分析対象に



データ分析における留意事項(続き)

- (7) 相関係数の信頼度に留意
- (8) 検定／推定的前提条件を忘れずに
- (9) 疑わしきは経過観察
- (10) 統計学的基準に振り回されない
- (11) 分析結果は利用者の立場で
- (12) 分析結果はひとつの指針
- (13) 誤った分析事例



(1) データの収集と分析の目的を明確に

- ・思いつくままに多くの種別のデータを収集・分析

→ 非効率

- ・何のために／何を知りたいか明確にしてから
データを収集・分析

- ・しかるべき書物(例えば参考文献[1], [3]など)や各自の
経験をもとに、データ収集・分析モデル(各データ項目
の相互関係に関する仮説)を明らかにしてから、データを
収集・分析



(2) 分析の前にデータ群の特性把握を

- ・ 相関分析・回帰分析の前に散布図を
 - 複数グループの混在発見
 - はずれ値の発見
- ・ 各変数のヒストグラムを
 - 分布の確認(正規分布か?)



(3) 適切な層別を

- ・異なった属性を持つものを混在して分析
 - 分析結果が曖昧
- ・異なった属性をもつデータは別々に分類(層別)
 - 例) システムソフトとアプリケーションソフトの開発条件の差
 - 生産性に差
 - 生産性の分析では層別によって分析精度が向上



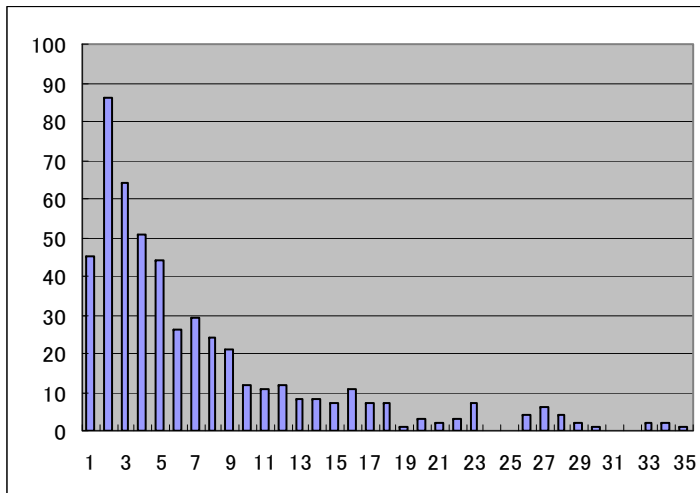
(4) 必要なら変数変換を

- ・プロジェクトの量的データ(工数・規模など)の分析
≠
正規分布
- ・プロジェクトの量的データの多くは
対数変換により正規分布に

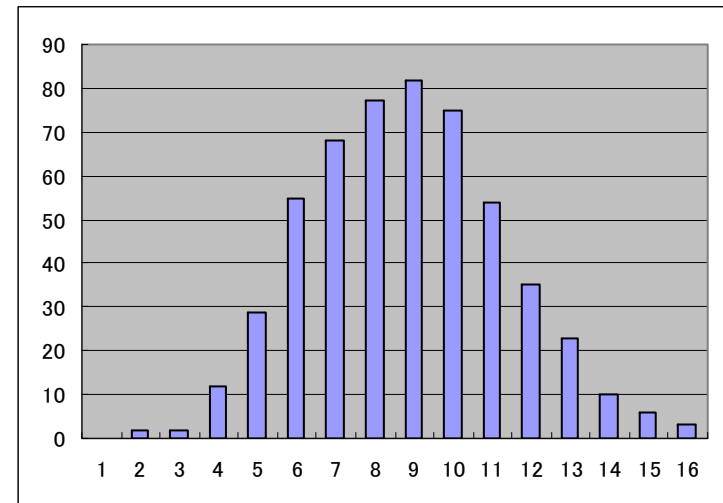
プロジェクトデータの典型的分布

([5]の原データより作成)

規模・工数・工期などの分布例
(右裾が長く延びる非対称形)



対
数
変
換



正規分布



(5) まず2変数間の分析を

例) 目的変数に影響を与える要因の分析方法

- ・相関分析・回帰分析
- ・分散分析

注意1) 目的変数に影響を与えない説明変数
が多くなると、分析結果の精度が低下

注意2) 回帰分析における高い相関を持つ
複数の説明変数 → 結果が不安定



(6) 基本測定量を分析対象に

例) プロジェクト管理者 (PM) のスキルレベルと
生産性の関係の分析

方法1) 目的変数: 生産性 = 工数 / 規模 (←加工)

説明変数: PMのスキルレベル

→ 割算によって規模情報が失われる

方法2) 目的変数: 工数

説明変数: 規模とPMのスキルレベル
(2変数)

一般に方法2の方がよい

重回帰分析における要因の選択方法

- (1)すべての組み合わせを調べる総あたり法
- (2)説明変数が1つも含まれない場合から始めて
順次1つずつ説明変数を追加していく方法
- (3)説明変数の候補すべてを含めた場合から
開始して不適切な説明変数を1つずつ
減らしていく方法

プロジェクトデータの分析では、経験上(2)がよい？

変量間の関係を調べる方法は 尺度によって異なる

- ・量的変数(比尺度)同士: 相関係数
 ／回帰分析など
- ・量的変数(比尺度)と質的変数: t 検定
 ／分散分析など
- ・質的変数同士: χ^2 検定など
- ・質的変数(順位尺度)同士: χ^2 検定
 ／順位相関など



(7) 相関係数の信頼度に留意せよ

- ・データ数の少ない相関分析はあてにならない
- ・データ数が多ければ相関係数Rの絶対値が低くても、2つの変数間に関係がある
 - 目的変数に与える説明変数の影響度
= 寄与率
= 相関係数の二乗
 - 目安: データ数30で $R=0.36$ なら相関あり

閑話休題

P値について

- ・ある(帰無)仮説が成立つ確率
例)ある相関係数のP値が0.04
→2つの変数間に相関がないという仮説は危険率5%で成り立たない
- ・相関の有無は相関係数の値よりもP値の方が重要
- ・P値の確認→相関係数の確認の手順で



(8) 検定／推定的前提条件を忘れずに

分析結果が統計的に有意(意味がある)か？
あるいは推定値の信頼幅はどれくらいか？

→ t-検定、F検定、 χ^2 検定などで確認

前提条件は「母集団が正規分布であること」
(相関分析や回帰分析も含めて)

ただし、ヒストグラムを人間の目で見ても、正規分布とみなせれば、検定を行わなくても概ね大丈夫

閑話休題

検定の目安

危険率5%で有意性の有無を判断する目安

1) 相関係数: 標本数30、相関係数0.36 \div 1/3以上

→母集団に相関あり

2) t 検定: 2つの標本

(平均 μ_1, μ_2 、標準偏差 σ_1, σ_2 、標本数 n_1, n_2 以上)

t 値 = $|\mu_1 - \mu_2| / \sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2} > 2.0$ が成立すれば、2つの

母集団の平均値に差あり

標本数が無限大の場合は1.96、 n_1, n_2 が10ずつの場合でも2.1を超えれば有意であることから

→”2”という数字がおおまかな目安

(9) 疑わしきは経過観察

- ・相関係数RのP値が0.1(10%) (> 0.05 (5%))
の場合に2つの変数の間に相関はないと断定
すべきかどうか？

「母集団のパラメータ(例えば平均値)の推定精度はデータ数nの平方根に比例して向上する」

→2つの母集団同士で本当に相関がある場合は、
標本数を増やすに従ってP値は漸次小さくなる

→nが小さいときはnが大きくなるまで経過観察



(10) 統計学的基準に振り回されない

統計学的基準は単なる約束事(共通の土台の提供)

例)「危険率5%で相関係数が有意である」の意味

＝「2つの変数が独立で本当は相関がない場合でも、計算で求めたような相関係数が得られる可能性(分析ツールではP値として出力)があるが、それが20回に1回(5%)と小さいから母集団に相関があると言ってもよい」



プロジェクト運営と統計的有意水準

プロジェクト運営をする場合：厳密な線引き（統計的有意水準）は横目で見ても

例) P値が6%の場合：「本当は相関がない可能性が6% = 1/16程度」と解釈して利用

P値が10%の場合：「本当は相関がない可能性が10% = 1/10程度」と解釈して利用



(11) 分析結果は利用者の立場で

- ・ 数学的な有効桁数と読者の認知レベルのバランス
- ・ 有効数字は1桁減らして採用

数学的な有効桁数と 読者の認知レベルのバランス

分析結果の報告では、数学的な有効桁数と報告書の
読者が認識・記憶できる桁数とを考慮して決定

例) 計算結果の有効数字が4桁以上あることは稀

- ・精度が1/100のデータを100個集めて平均をとっても、その精度は1桁しか上らず高々3桁
- ・4桁の数字を記憶できる経営者や管理者は・・・

→ 最終結果の有効数字は2桁から3桁で十分



有効数字は1桁減らして採用

分析対象のプロジェクト群 ≠ 自分のプロジェクト

→ 論文や報告書の分析結果は1桁丸めて利用

例) 生産性が3.62(単位は省略)という報告

→ 3.6までは信用

生産性が3.6という報告

→ 4という数値を利用



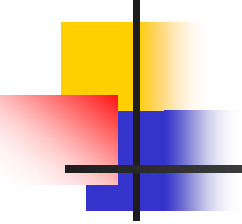
(12)分析結果はひとつの指針

- 定量化は銀の弾丸ではない
- 勘ピュータに与える基礎データ
- 分析結果に魂を与えるのはPM
- 3割打者と2割5分の打者のどちらを使う？



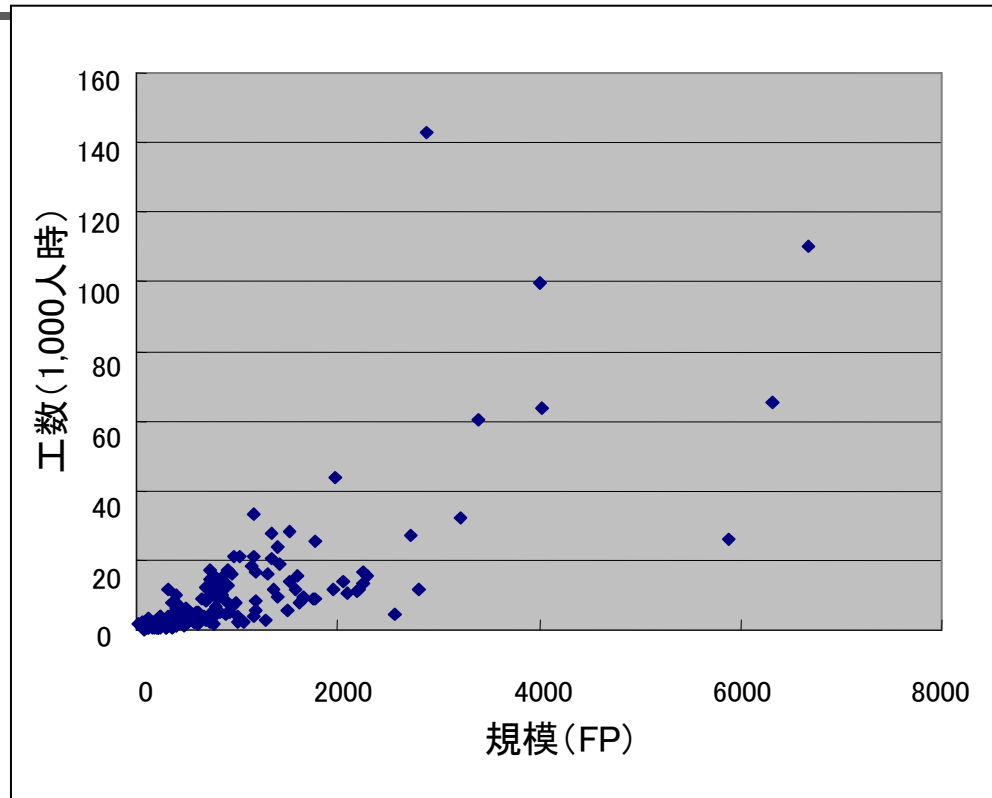
(13)誤った分析事例

- ・対数変換せずに回帰分析
 - ・・・回帰分析の前提を満たしていない
- ・相関係数だけ議論して信頼幅を議論しない
 - ・・・正規分布の仮定とデータ数による信頼限界がある
- ・分散を考慮せずに平均値だけで良し悪しを議論
 - ・・・分散が大きければ平均値に差があっても2つの集団に差があるとは言えない



データの分析事例 (回帰分析)

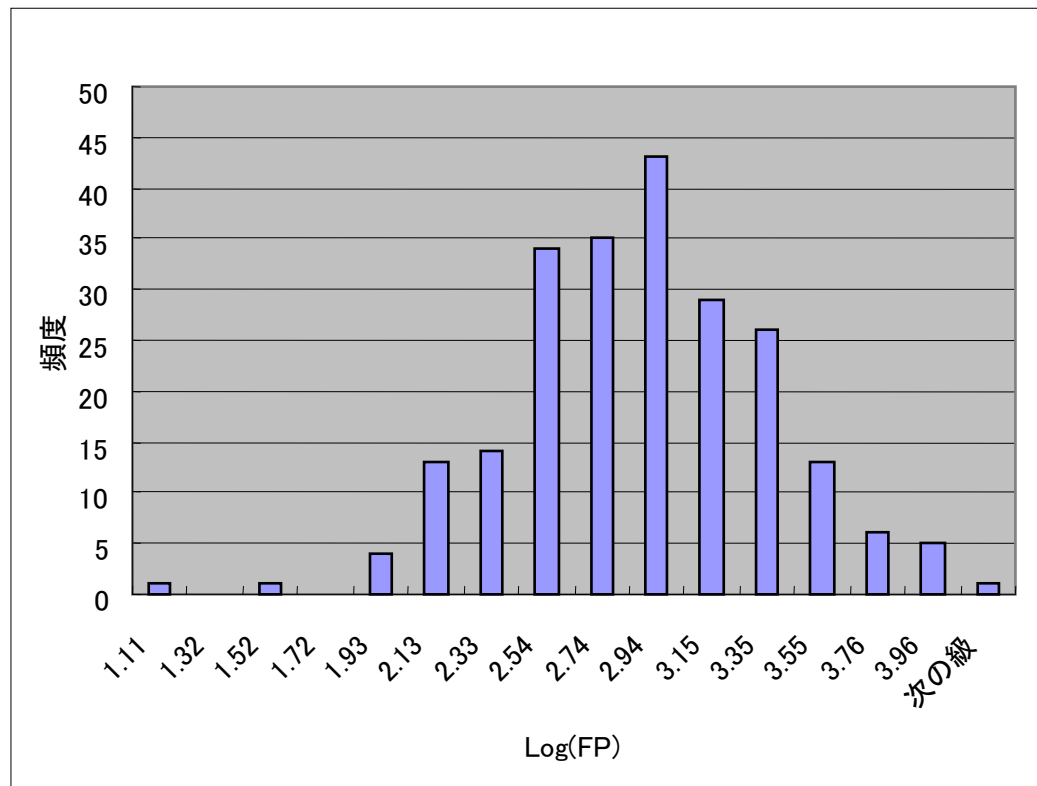
(1) 生データの分布の確認



このままでは玉葱の皮むき

([4]より引用)

(2) 対数変換によるヒストグラム確認



正規分布とみなせる

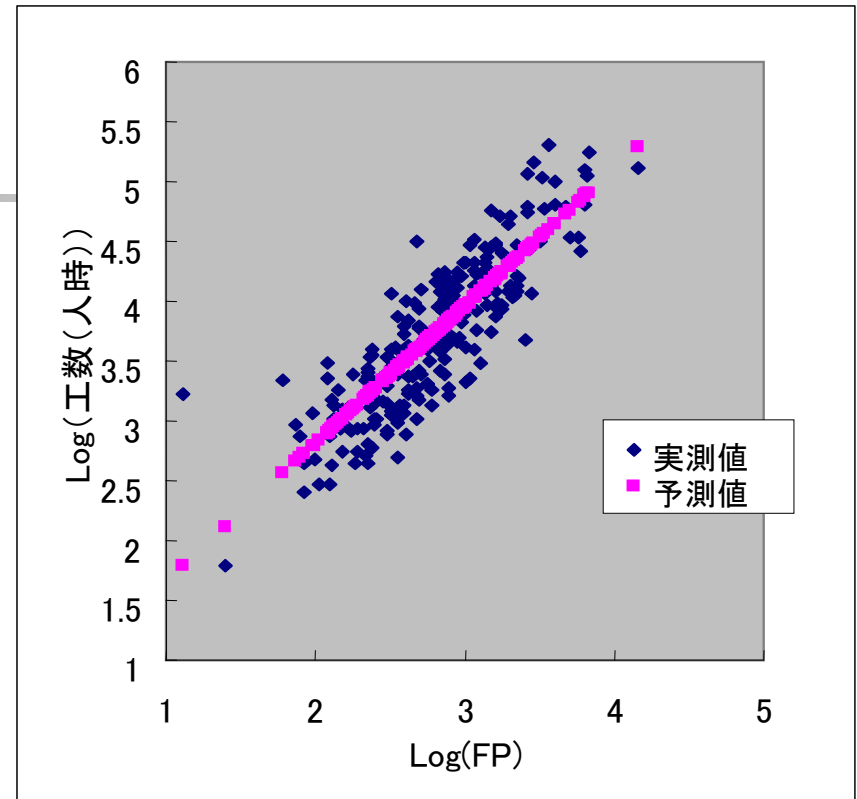
([4]より引用)

(3) 散布図の作成 と回帰分析

$$\begin{aligned}\text{Log(工数)} &= 1.15 \text{ Log(FP)} + 0.513 \\ \text{人時} &= 10^{0.513} \times \text{FP}^{1.15} \\ &= 3.26 \times \text{FP}^{1.15}\end{aligned}$$

相関係数は0.85(寄与率=0.72)

→工数の72%は規模要因で決定



([4]より引用)



(4) 信頼幅の計算

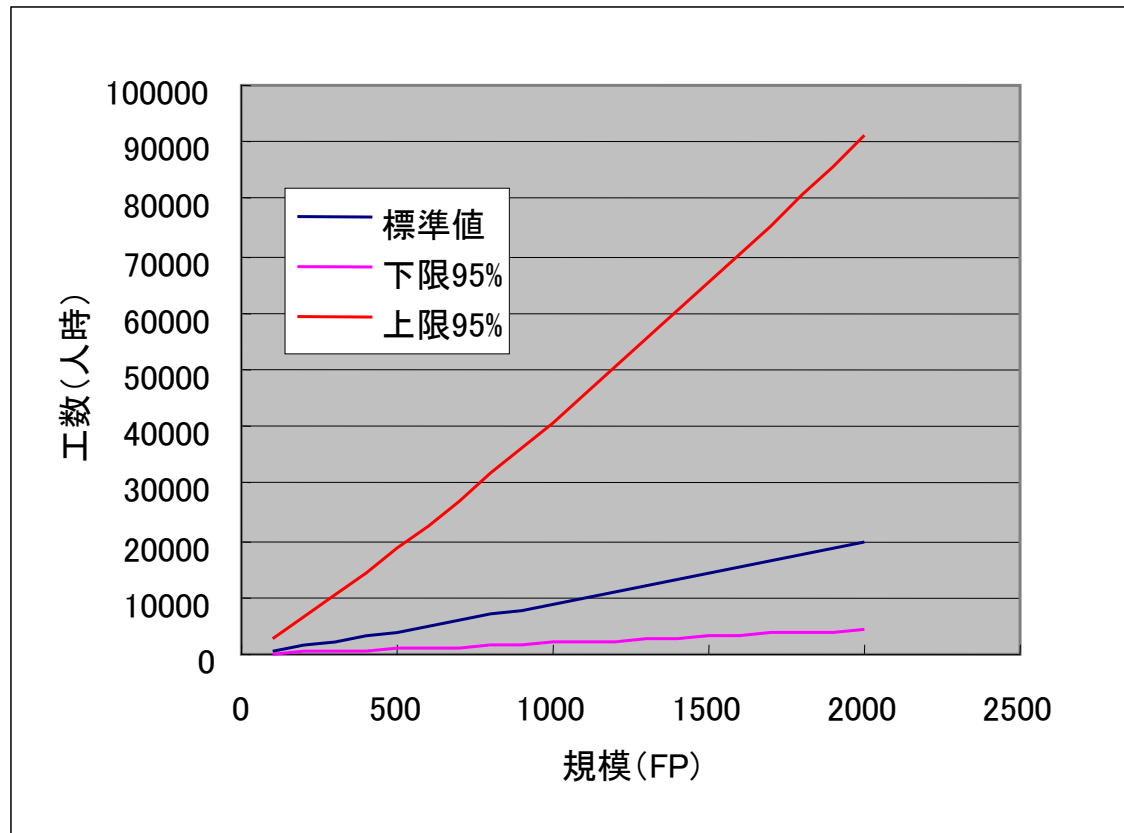
- 回帰の95%信頼幅を計算
(計算式は標準的な統計学の教科書を参考に)

- ソフトウェア工学の世界では
 - 95%の信頼幅は大きい
→ 特異なプロジェクトの検出に

 - 50%の信頼幅のグラフも加える
→ 平均的なプロジェクトの確認に

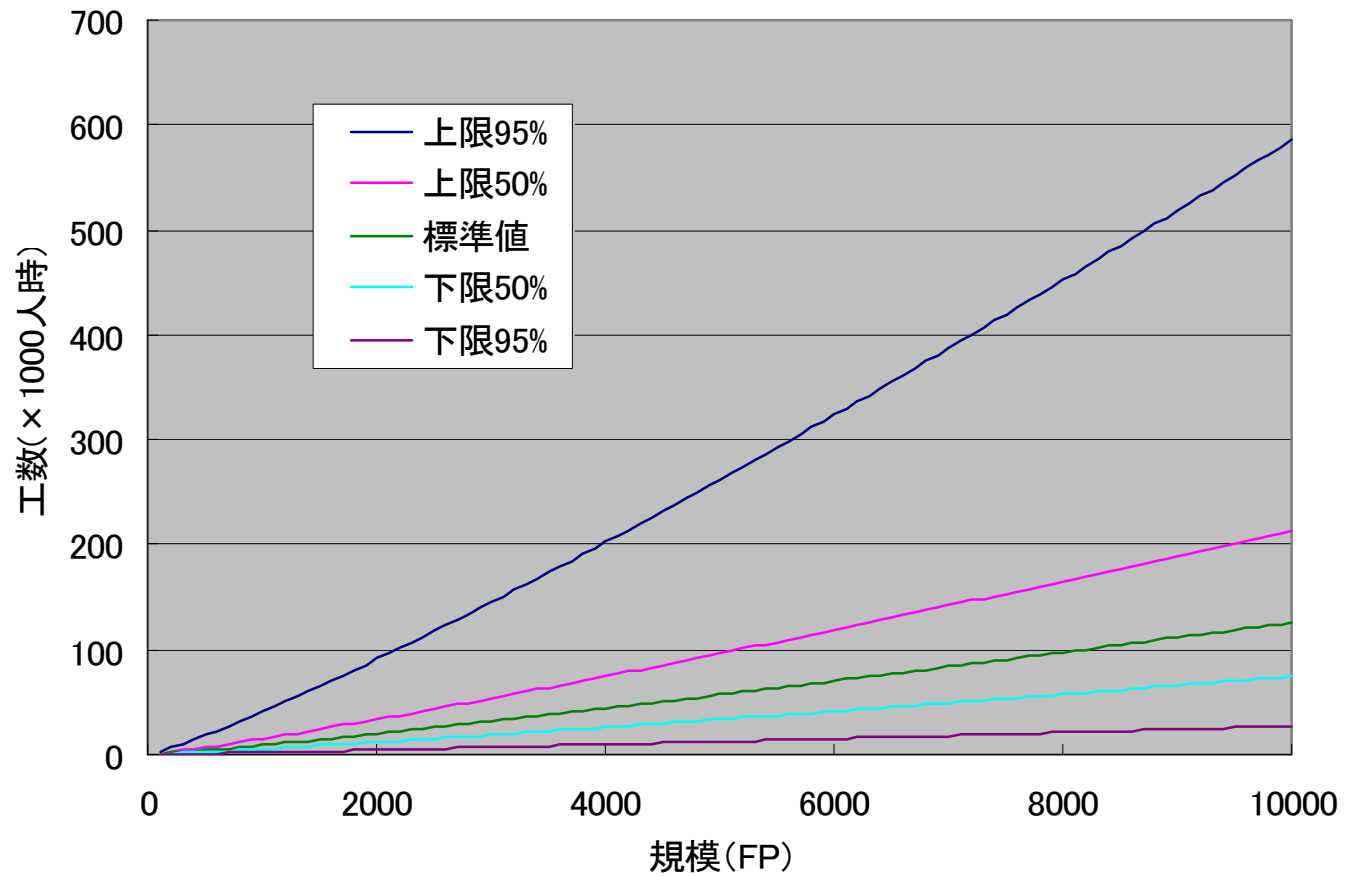
(5)もとの世界で結果を表示

規模あたりの工数の予測曲線と信頼幅



([4]より引用)

50%の信頼幅を加えたグラフ





おわりに

(1) 定量化の目的と効果

・・・定量化はなぜ重要か

- 現実を正確に把握
- 勘と経験の補強
- 定性モデルを数値による詳細化へ

(2) 定量化の枠組み

・・・何をどのように測定するか

- コミットメント(確約)の確立と維持
- 測定ニーズの特定と優先順位付け(*)
- 測定量の選択と特定(*)

(*)についてはさまざまな書籍が参考になる



(3) データの分析方法

・・・測定結果の分析と利用方法

→ 統計学の利用

- ・大学の教養レベルの統計学で十分
- ・統計学の心を知る必要

→ 現実のデータに応じた手法の修正

- ・現実のデータは千変万化
- ・修整した手法による分析の試行錯誤

→ 分析結果はひとつの指針

- ・魂を入れるのはPM



参考文献

- [1] 鶴保征城, 富野 壽監訳: ソフトウェア開発の定量化手法 (第2版), 共立出版, 1998.
 - ・・・定量化の目的・効果や定量化システムの構築
- [2] JIS X-0141:2009 (ISO/IEC 15939:2007) システム及びソフトウェア技術—測定プロセス, 平成21年5月20日改正.
 - ・・・測定プロセスの枠組み
- [3] John McGarry他著, 古山恒夫, 富野 壽監訳: 実践的ソフトウェア測定, 共立出版, 2004.
 - ・・・情報ニーズや測定量の例と選択方法の指針
- [4] 古山恒夫: プロジェクトデータ分析の指針と分析事例, SECジャーナル, Vol. 1, No. 3, pp. 6-13, 2005.
 - ・・・データ分析上の留意事項

参考文献(続き)



[5](独)情報処理推進機構ソフトウェアエンジニアリングセンター:ソフトウェア開発データ白書2005,日経BP社,2005.

- ・・・豊富なプロジェクトデータと測定項目(測定量)

[6](独)情報処理推進機構ソフトウェアエンジニアリングセンター編:定量的品質予測のススメ、オーム社、102p.(2008).

- ・・・上流工程からの信頼性予測を定量的に行うための実践的方法