

統計によるソフトウェアの品質および工数の予測

石橋雄一（スタットラボ）

プロジェクトごとの、規模・工数・テスト項目数・検出バグ数などのデータを統計的に分析することにより、工数やバグ数の増加要因を調べたり、予測式をつくることにより工数見積もりやバグ数の予測を行うことができます。

対象となるデータ

分析を行うためのデータは以下のように、プロジェクトごと、工程ごとの項目です。

分類	項目	
ソフト特性	開発タイプ 開発形態 システム構成 利用目的 処理形態 利用形態 ：	開発タイプであれば、新規・機能拡張・保守・カスタマイズの4種類。各項目にいくつかのカテゴリがあります。
開発規模	開発規模の実績値 ファンクションポイント数 ソースコード行数 母体規模 再利用率	ファンクションポイント（F P 値）は機能の大きさを計測する基準である。開発規模をF P 値で計測しない場合、言語ごとの規模からF P 値に変換することもできます。
開発工数	プロジェクト全体 要求分析 機能設計 詳細設計 製造 単体試験 結合試験 システム試験 ユーザ向けドキュメント リリース後サポート ：	左側の項目に対してそれぞれ以下の入力項目があります。 工期開始日 工期終了日 要員数（人数） 工数（人時） レビュー工数（人時） 成果物量（文書頁） 成果物量（テスト項目数）
属性	プロジェクト管理者の経験 業務分野の経験 分析・設計経験 言語・ツール利用経験 ：	これらは管理者へのアンケート項目で、各項目は5段階の評点となっています。

データの件数は、ソフトの特性のばらつき方にもよりますが、少なくとも100プロジェクト位必要となります。また、最初から全ての工程のデータが揃っている必要はなく、工程全体のデータから分析を行うことができます。

何ができるか

	要因分析	予測
品質（バグ数）	プロジェクトごとのバグ数増加の要因を洗い出し、バグ減少のための改善提案を行います。	プロジェクト開始時に、抽出予定バグ数予測を出します。これと実績値と比較し、この差が何によるのかを分析します。
テスト項目数	プロジェクトごとのテスト項目数とバグ数との関連を分析することにより、バグ検出のためのテスト項目が用意されているかどうか分析します。	テスト項目数の予測値をもとにして、適切なテスト項目数を提示します。
工数	工数を増加させる要因を洗い出し、工数を削減するための改善提案を行います。	予測式により新規プロジェクトの工数見積もりを行います。予測された工数と実績値との差が何によるものなのかを分析します。

分析・予測の精度を上げるために

一つの企業あるいは組織に閉じた分析	データや分析結果は外部には公表しませんので、信頼性の高いデータを用いて、精度の高い分析を行うことができます。
部門の違いを考慮した分析	企業や組織内でも、部門間には計測方法や開発するソフトの種類、管理方法など様々な違いがあります。この部門の違いを考慮しますので、分析・予測の精度を上げることができます。

効果・実績

企業内データの分析	リリース後バグ数の分析	リリース後バグ数と、テスト項目数および検出バグ数などの関係を分析することにより、リリース後にバグが多く発生するのはどのようなケースなのかを調べることができました。
	バグ数・工数の分析・予測	バグ数や工数を増加させる要因を洗い出し、工数の予測式を重相関係数が約0.8の精度で作ることができました。要因の一つとして、テスト体制の整備が大きく影響しており、また部門間での整備状況もかなり異なることがわかりました。
JASMIC 定量化情報提供サービスセンター	現在、構造計画研究所では、インターネット上で幅広い企業から提供されたソフトウェアの実績データの蓄積と工数・品質などの予測を行っています。	

なお、この分析は株式会社構造計画研究所の定量化情報提供サービスセンタープロジェクトとして、東海大学開発工学部教授の古山恒夫先生のご協力のもとに行われました。