

SPI Japan 2011

組織の実績ベースライン、および、 改善活動の効果の検定手順の確立

2011年10月27日

住友電気工業株式会社

中塚 康介

中村 伸裕

はじめに

- 高成熟度のプロセス領域では統計的手法の適用が必要となる
 - 組織のプロセス実績の分布を検定する
 - 各プロジェクトの品質改善活動の効果を検定する
- 統計的手法適用における問題
 - 種々の統計手法がそのまま適用可能なプロセス実績ベースラインとなるとは限らない
 - 統計的な検定は数学的な知識が必要となり「難しい」ととらえられがち
 - プロセス実績ベースラインの変換と正規性検定による検証で、正規分布に関する統計処理を適用可能にする
 - 検定を手順化・ツール化し、容易に利用できるようにする



Agenda

- 改善の体制
- プロセス実績ベースラインの検定
- 各プロジェクトの品質改善活動の検定
- 検定のツール化・標準化

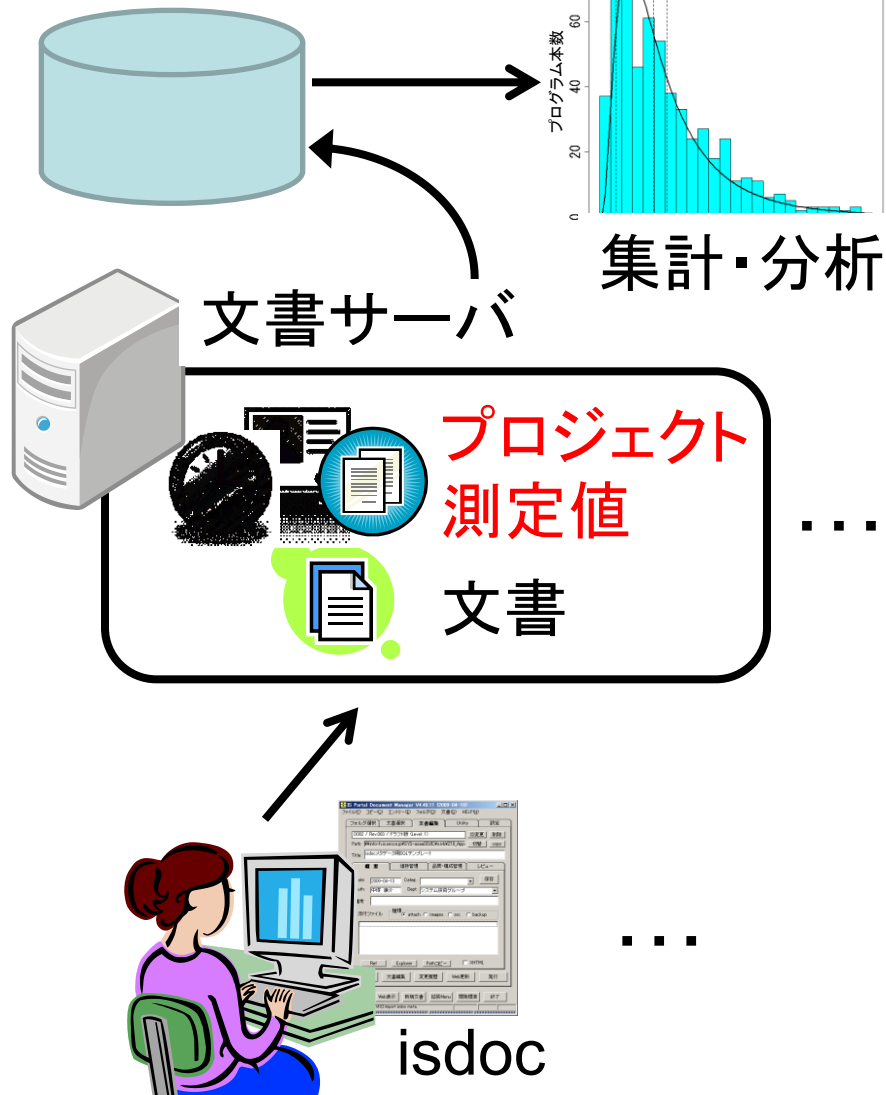
背景:改善の体制

- 品質改善活動のワーキンググループが主体
 - 住友電気情報システム株式会社 (以下SIS)
 - 品質改善推進グループ
 - システム開発部隊
 - 住友電気工業株式会社 (以下、住友電気)
 - システム技術グループ
- 上記体制で2011年6月CMMI Ver 1.3 レベル5の達成を確認
- システム技術グループで統計的分析手法の調査、統計的分析を実施
 - ワーキンググループには2名参加
 - 統計の専門ではない

背景：組織の実績値の収集

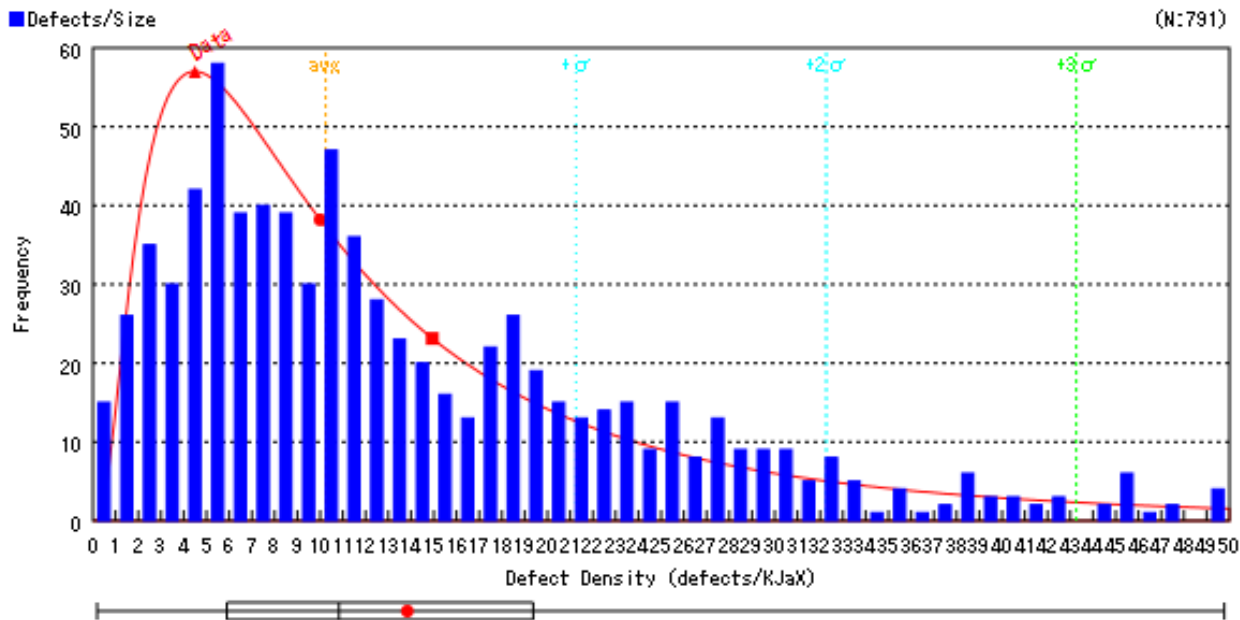
- 自社開発の文書作成・構成管理・定量管理ツールを使用
- 測定メトリクス数: 312
- 集計規模
 - システム数: 464システム
 - 外部仕様書: 156,897ページ
 - プログラム: 16,727本
- 中塚康介、中村伸裕、“組織レベルの開発実績収集・分析”, SPI Japan 2009, 2009

全社実績リポジトリ



プロセス組織ベースライン

- プログラム単位行数あたりの作り込み欠陥の分布
 - 組織の品質目標達成のため、プロジェクトが品質計画、品質の監視と制御に使用



- この分布をどう扱えば良いか？
 - 例：プロジェクトの改善活動の効果はどう確認？

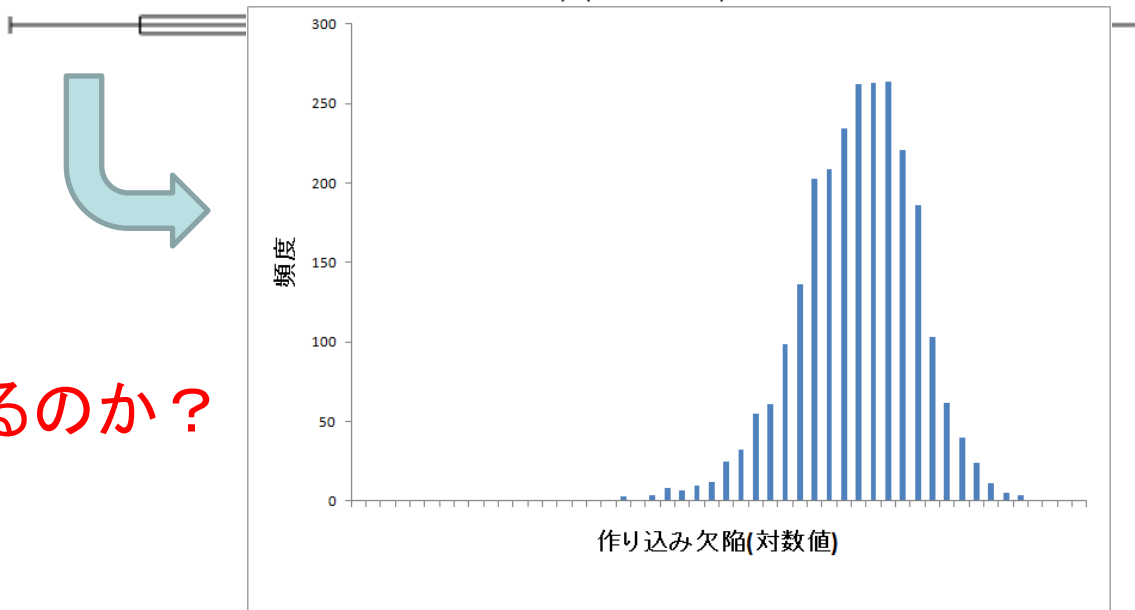
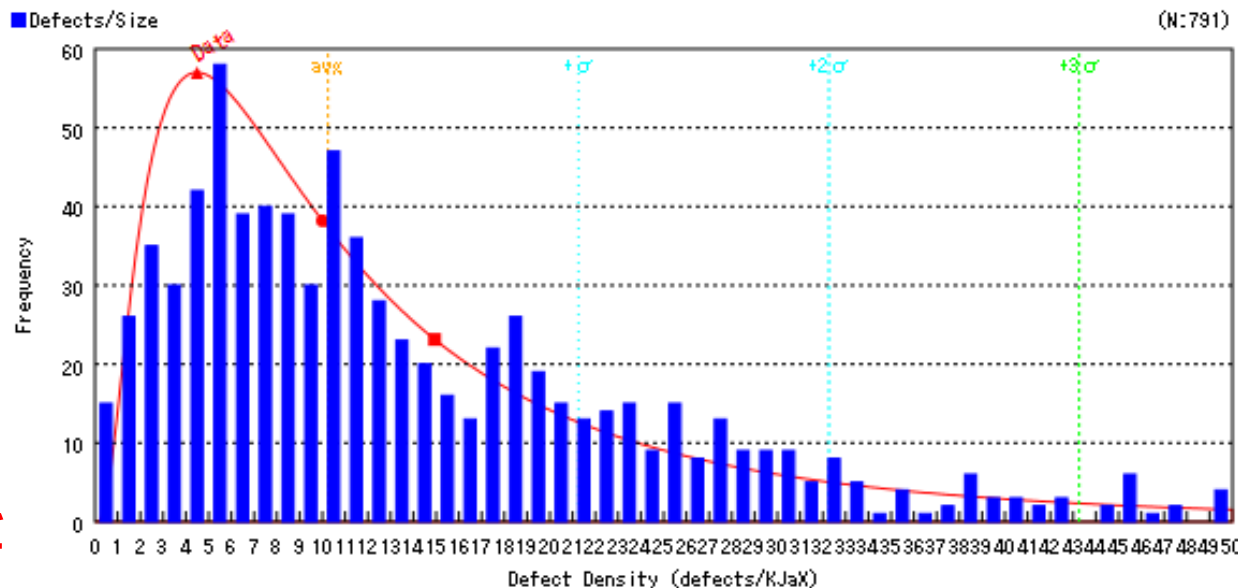
組織の実績値の分析

- プログラム単位行数あたりの作り込み欠陥の分布

- 対数変換することで正規分布に近い形となっているようである

- 正規分布ならば多くの統計手法が適用可能

⇒ 正規分布であると言えるのか？



プロセス実績ベースラインの正規性検定

- 正規分布であることを示すためには正規性検定を行う
 - Kolmogorov-Smirnov検定(KS検定)、正規性に関するShapiro-Wilk検定、Anderson-Darling検定、D'Agostino-Pearson検定(K2検定)、...

- 統計ツールには正規性検定が実装されている
 - 1. Excelなどから統計ツールにデータをコピー
 - 2. 正規性検定を実行
 - 3. 結果を確認(基本的には値の大小で判定できる)

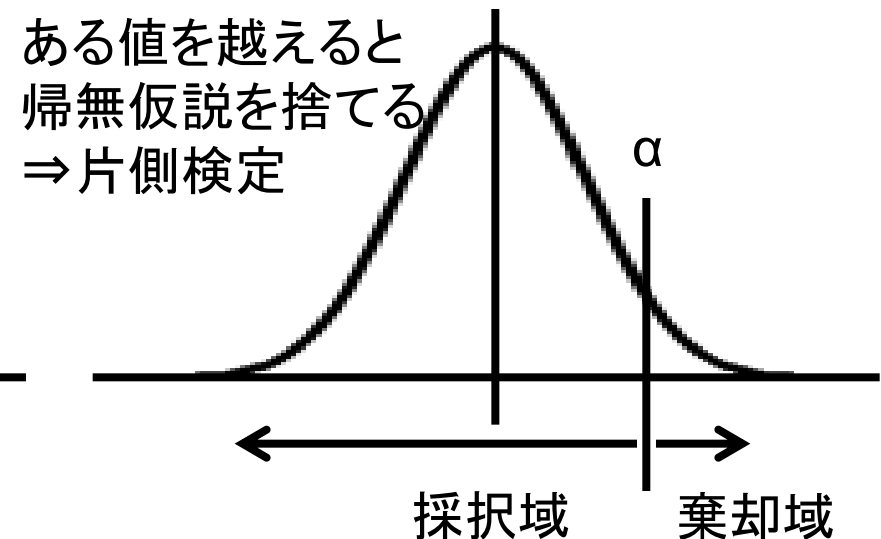
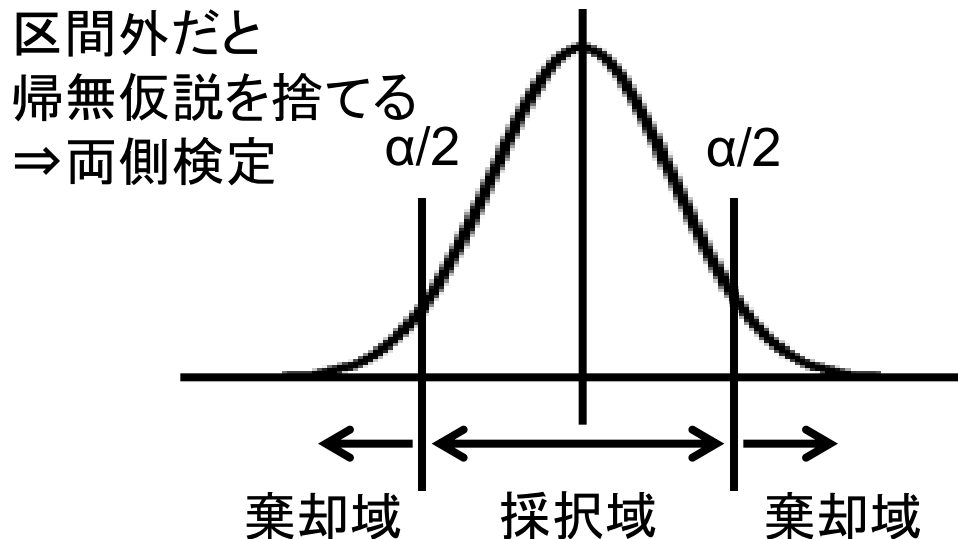
- 検定手法によって性能・条件が異なるので、事前に確認する
 - 正規分布を正規分布と判定するか
 - 正規分布でないものを正規分布でないと判定するか

(補足) 検定について: 仮説

- **仮説**: 確認したいことを統計的に表現したもの
 - 例「プロセス改善後は欠陥数の分布が変化した」
- **帰無仮説**: 確認したい仮説とは逆の、否定したい仮説
 - 例「プロセス改善後は欠陥数の分布が変化していない」
- **帰無仮説の棄却**: 統計的に帰無仮説が起こりえないことを示す
 - 棄却されると「変化した」と主張できる
 - 棄却されない場合は「変化したとは言えない」
 - ≠「変化していない」

(補足) 検定について: 採択・棄却

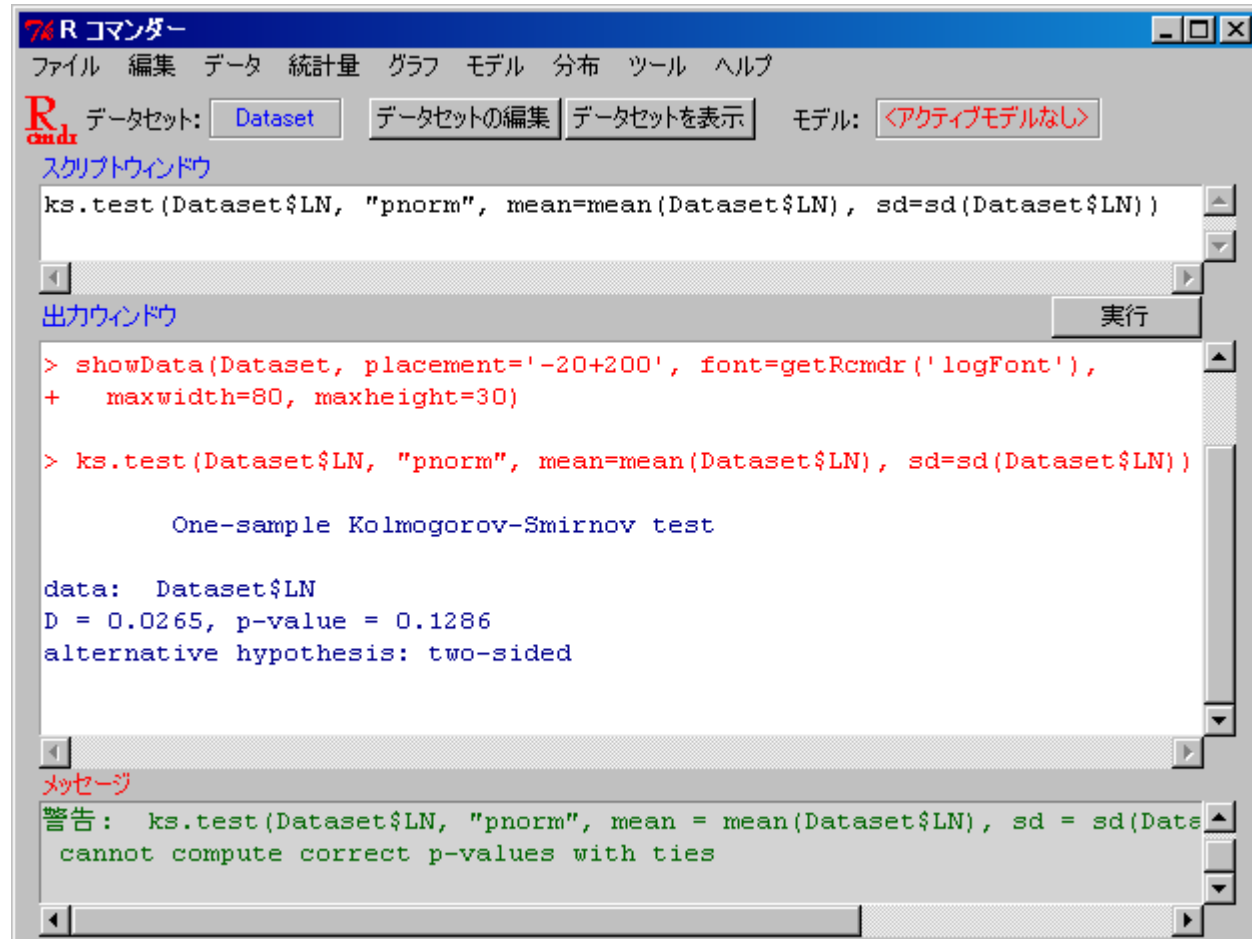
- **有意水準 α** : 帰無仮説の棄却を決める確率
 - $\alpha = 10\%$ 、 5% 、 1% などを利用
- **棄却域**: 帰無仮説を捨てる領域
 - 帰無仮説が「起こりえない」ことを表す領域
- **採択域**: 帰無仮説を捨てない領域
 - 帰無仮説が「起こりうる」ことを表す領域



正規性検定の実施

■ 統計ツールとしてRを使用

- オープンソースソフトウェア
- KS検定(`ks.test`)やShapiro-Wilk検定(`shapiro.test`)といった正規性検定が実装されている
- GUIのツールRcmdrがある



The screenshot shows the Rcmdr GUI window titled "R コマンドー". The menu bar includes "ファイル", "編集", "データ", "統計量", "グラフ", "モデル", "分布", "ツール", and "ヘルプ". The main area contains a "データセット:" field with "Dataset" selected, and a "モデル:" field with "<アクティブモデルなし>". Below this is a "スクリプトウィンドウ" containing the command: `ks.test(Dataset$LN, "pnorm", mean=mean(Dataset$LN), sd=sd(Dataset$LN))`. An "出力ウィンドウ" shows the execution of this command, including a `showData` call and the output of `ks.test`:
One-sample Kolmogorov-Smirnov test
data: Dataset\$LN
D = 0.0265, p-value = 0.1286
alternative hypothesis: two-sided
A "メッセージ" window at the bottom displays a warning: "警告: ks.test(Dataset\$LN, "pnorm", mean = mean(Dataset\$LN), sd = sd(Dataset\$LN)) cannot compute correct p-values with ties".

正規性検定の実行結果例

```
> ks.test(Dataset$LN, "pnorm", mean=mean(Dataset$LN),  
sd=sd(Dataset$LN))
```

One-sample Kolmogorov-Smirnov test

data: Dataset\$LN

D = 0.0265, p-value = 0.1286

alternative hypothesis: two-sided

- プログラム単位行数あたりの作り込み欠陥の対数値をKS検定
 - データ量を変えながら一様分布、正規分布などを用いてKS検定が使えることを確認
- p値が0.05未満であれば、正規分布に従うという帰無仮説が棄却される。0.05より大きければ棄却されない($\alpha = 0.05$)
 - 注: 0.05より大きくても正規分布でないかもしれない

正規性検定の手順化

- プロセス実績ベースラインの作成は限られている
 - 年間数回
 - 品質改善のグループが実施



- 以下があれば運用できる
- 統計ツールのインストール・利用手順
 - 住友電工・SISではRを利用
- データ抽出方法
 - 測定リポジトリから実績を抽出するVIEWを作成
- 検定手法
 - 開発標準に記述

プロセス実績ベースラインの発行標準

- 正規性検定の手順を含んだ形でプロセス実績ベースラインの発行標準を作成

| 組織プロセス実績 | | |
|---------------|----------------|------|
| 文書ID | タイトル | |
| OPP.PPB.C [-] | PPB 作成手順 | Ver. |
| OPP.PPB.C.1 | 作込欠陥分布 作成手順 | Ver. |
| OPP.PPB.C.2 | レビュー区分別実績 作成手順 | Ver. |
| OPP.PPB.C.3 | 成果物相模変換係数 作成手順 | Ver. |

正規性検定の手順を含んだ作成手順

- この標準を元に、6月と11月にプロセス実績ベースラインが発行される

標準化された正規性検定手順

< M > TOP | CMMI | SIS システム開発ホームページ (開発標準) | SIS開発標準(品質改善推進G) | 作込欠陥分布 作成手順

[履歴][UTIL][操作]

作込欠陥分布 作成手順

| | | | | | | | | | |
|----|-------------|---------|-----|-------|------|---------------------|-----|------------|---|
| ID | OPP.PPB.C.1 | Ver.001 | 発行者 | 中村 伸裕 | 発行部署 | 品質改善推進G | 発行日 | 2011-01-19 | X |
| | | Rev.001 | 作成者 | 加藤 孝則 | 作成部署 | QCD改善推進部品品質改善推進グループ | 作成日 | 2010-12-28 | |

全文

< >

添付資料

実施要領

成果物補足説明

テラリング基準

評価指標

参照文書

被参照文書

関連システム

7. 分布の検定

正規分布への近似度を測定するため

欠陥数/規模の対数を求めRにて検定 (コルモゴロフ・スミルノフ検定) する

(対象は作込欠陥分布のみ)

- 1. 検定対象のデータ抽出

6-1 で使用したデータの欠陥率 (欠陥件数/規模)

の対数 (自然対数) を導出しCSVファイルに保存する

- 2. 検定ツールへの取込および検定

(1) 統計ツールRを起動する

(2) 7-1のデータをツールに取り込む

(3) 検定用の関数を起動する

関数名: kstest

引数1: "pnorm" (=正規分布)

引数2: (導出した「欠陥率の対数」の列)

(4) 返ってきた計算結果をisdocに記録する

※結果として表示されている

----- (検定) -----

各プロジェクトの改善効果の検定

- 各プロジェクトでは作り込み欠陥を減らす改善活動を実施
⇒ 効果があるものを組織に展開

- 「効果がある」とはどういうことか？
⇒ プロセス実績ベースラインとプロジェクトの実績を比較して差があるか、差があるとすれば良くなっているか



- 統計的に判断するため、検定を行う

改善効果の検定

- プロセス実績ベースラインとプロジェクトの実績に差があるか
⇒ 2つの分布の平均値に差があるかどうかを調べる
 - 作り込み欠陥の比較の場合、プロジェクト < プロセス実績ベースラインであれば改善されていると言える

- 1. 分散の検定(F検定)を行う
- 2. 分散が等しい → 等分散の平均値の検定(Studentのt検定)
分散が等しくない → 不等分散の平均値の検定(Welchのt検定)
- 3. 平均値に差があると言えない → 改善されていると言えない
平均値が差がある →
 - プロジェクトの実績 < 組織の実績 → 改善されている
 - プロジェクトの実績 > 組織の実績 → 改善されていない

改善効果の検定の手順化

- 改善活動は各プロジェクトで行われる
 - 途中結果の効果の検定など頻繁に行う
 - 必ずしも統計に詳しいわけではなく、時間もない



- 簡単に改善効果の検定を行えるツールが必要
 - Excelで実装
 - ExcelにはF検定(FTEST)、t検定(TTEST)がある
 - 組織の実績ベースラインを埋め込み

改善効果検定Excelシート

| | A | B | C | D | E | F | G | H | I | J |
|----|------|---|-----------------------|-------------------------|---|---------|----------|---|---|---|
| 5 | 分布1 | 組織の実績ベースラインを指定します。「シート」欄にシート名、「列」に列名(AやBなど)を指定してください。 | | | | | | | | |
| 6 | | | | | | | | | | |
| 7 | | 分布1 | シート | PG開発作込欠陥(CI除く)2007~2010 | | | | | | |
| 8 | | | 列 | L | | | | | | |
| 9 | | | | ※ 赤枠内を記入下さい。 | | | | | | |
| 10 | | | | | | | | | | |
| 11 | | | 平均 μ (AVERAGE) | 2.051966 | | | | | | |
| 12 | | | 分散 $V=\sigma^2$ (VAR) | 0.926973 | | | | | | |
| 13 | | | 標準偏差 σ (STDEV) | 0.962794 | | | | | | |
| 14 | | | | | | | | | | |
| 15 | | | | | | | | | | |
| 16 | 分布2 | プロジェクトの実績値を指定します。「シート」欄にシート名、「列」に列名(AやBなど)を指定してください。 | | | | | | | | |
| 17 | | | | | | | | | | |
| 18 | | 分布2 | シート | GOS出荷PG無効・UT・CI除 | | | | | | |
| 19 | | | 列 | F | | | | | | |
| 20 | | | | ※ 赤枠内を記入下さい。 | | | | | | |
| 21 | | | | | | | | | | |
| 22 | | | 平均 μ (AVERAGE) | 1.989936 | | | | | | |
| 23 | | | 分散 $V=\sigma^2$ (VAR) | 0.375936 | | | | | | |
| 24 | | | 標準偏差 σ (STDEV) | 0.613136 | | | | | | |
| 25 | | | | | | | | | | |
| 26 | | | | | | | | | | |
| 27 | | | | | | | | | | |
| 28 | | 分散の検定 | FTEST | 0.110766 | > | 0.05 | | | | |
| 29 | | | 対立仮説：分散に差がある | | | | | | | |
| 30 | | | 帰無仮説：分散に差がない | | | | | | | |
| 31 | | | 分散に差があると言えない→等分散 | | | | | | | |
| 32 | | | | | | | | | | |
| 33 | | 平均値の検定 | TTEST 両側 等分散 | 0.831034 | > | 0.05 | | | | |
| 34 | | | 対立仮説：平均値に差がある | | | | | | | |
| 35 | | | 帰無仮説：平均値に差がない | | | | | | | |
| 36 | | | 平均値に差があると言えない | | | | | | | |
| 37 | | | | | | | | | | |
| 38 | | 平均値の大小 | 分布1の平均値 | 2.051966 | > | 分布2の平均値 | 1.989936 | | | |
| 39 | | | | | | | | | | |
| 40 | 検定結果 | 平均値に差があると言えない | | | | | | | | |

改善効果検定Excelシートの利用(1)

1. プロジェクトの実績値をExcelシートにコピー
 - 対数正規分布用の計算式はシートに記述済み
2. シート名と列を指定
 - プロセス実績ベースラインはシートに含まれ、選択するだけ

| 分布1 | 組織の実績ベースラインを指定します。「シート」欄にシート名、「列」に列名(AやBなど)を指定してください。 | |
|-----|---|-------------------------|
| 分布1 | シート | PG開発作込欠陥(CI除く)2007~2010 |
| | 列 | L |
| | | ※ 赤枠内を記入下さい。 |
| | 平均 μ (AVERAGE) | 2.051966 |
| | 分散 $V=\sigma^2$ (VAR) | 0.926973 |
| | 標準偏差 σ (STDEV) | 0.962794 |
| 分布2 | プロジェクトの実績値を指定します。「シート」欄にシート名、「列」に列名(AやBなど)を指定してください。 | |
| 分布2 | シート | GOS 出荷PG無効・UT・CI除 |
| | 列 | F |
| | | ※ 赤枠内を記入下さい。 |
| | 平均 μ (AVERAGE) | 1.989936 |
| | 分散 $V=\sigma^2$ (VAR) | 0.375936 |
| | 標準偏差 σ (STDEV) | 0.613136 |

改善効果検定Excelシートの利用(2)

- 検定結果が自動計算される
 - 結果は「平均値に差があると言えない」「分布1の方が大きい」「分布2の方が大きい」の3通り
 - 作り込み欠陥の分析の場合
 - 分布1(プロセス実績ベースライン) > 分布2(プロジェクト)
- ⇒ 改善されている

| | | | | | |
|--------|------------------|----------|---|---------|----------|
| 分散の検定 | FTEST | 0.110766 | > | 0.05 | |
| | 対立仮説：分散に差がある | | | | |
| | 帰無仮説：分散に差がない | | | | |
| | 分散に差があると言えない→等分散 | | | | |
| 平均値の検定 | TTEST 両側 等分散 | 0.831034 | > | 0.05 | |
| | 対立仮説：平均値に差がある | | | | |
| | 帰無仮説：平均値に差がない | | | | |
| | 平均値に差があると言えない | | | | |
| 平均値の大小 | 分布1の平均値 | 2.051966 | > | 分布2の平均値 | 1.989936 |
| 検定結果 | 平均値に差があると言えない | | | | |

改善効果の検定の事例

- 各プロジェクトで行われた改善活動の結果を検定

| 改善グループ | 結果 |
|--------|--------------|
| A | 改善されていると言えない |
| B | 改善されている |
| C | 改善されている |
| D | 改善されていると言えない |

- B、Cの取り組みをさらに決定分析し、1つ(デシジョンテーブルの利用)を採用、試行
 - 評価基準：作込欠陥密度(平均・最頻値)、不良率(目標値以上の欠陥のある割合)、実施・教育コスト、アンケート結果等
 - 試行の効果も検定
 - 一部、改善効果が見られないものがあったが、さらなる分析によりそれらはチーム編成の影響を受けていることが判明

検定の標準化

- 各プロジェクトで改善効果の検定が行えるよう標準化する
 - 開始のための条件
 - 検定手法の説明
 - 検定Excelシートを含めたツールの説明

改善効果の検定手順

【目的】

改善活動により改善されていることを検証するため、改善活動結果と、それに対応する組織の実績ベースラインとの差を検定する。

【添付資料】

- ・添付ファイルはありません。

検定用のExcelシートについては下記文書を参照すること。

- ・ [\[OPM.5TEST.1\]](#)

【適用範囲】

完了、または、継続中であるが結果が得られている改善活動

【入 力】

- (1) 改善活動結果
- (2) (1)に対応する組織の実績ベースライン

【成果物】

検定結果

検定結果は以下のいずれかとなる

統計的管理の教育

- 住友電工で統計的品質管理に関する教育が実施されている
 - SQC入門（主として工場の品質管理向け）

- SISにおいても、上記に加えて以下を実施
 - ワーキンググループ内での説明
 - 品質改善推進者に対する検定の説明
 - SIS全社品質大会において、プロセス実績ベースライン、検定について説明

- 開発者・改善担当者が単にツールとして使うだけでなく、背景となる統計知識についても教育
 - ツール自身も開発者からの提案で改善されている

まとめ

- 高成熟度のプロセス領域では統計的な検定が必要となる
 - 組織のプロセス実績がどんな傾向になるのか
 - 各プロジェクトの品質改善活動は効果があったのか

- 以下の取り組みを実施
 - プロセス実績ベースラインの分布の推定と検定
 - ➡ ■ 正規分布を仮定した統計処理の適用が可能になった
 - プロジェクトの改善活動の効果の検定とツール化
 - ➡ ■ プロジェクトで容易に改善効果を判断できるようになった
 - ➡ ■ 効果を確認して、改善活動を展開できるようになった