

SPI Japan 2008

ソフトウェア生産性データの可視化 についての一考察

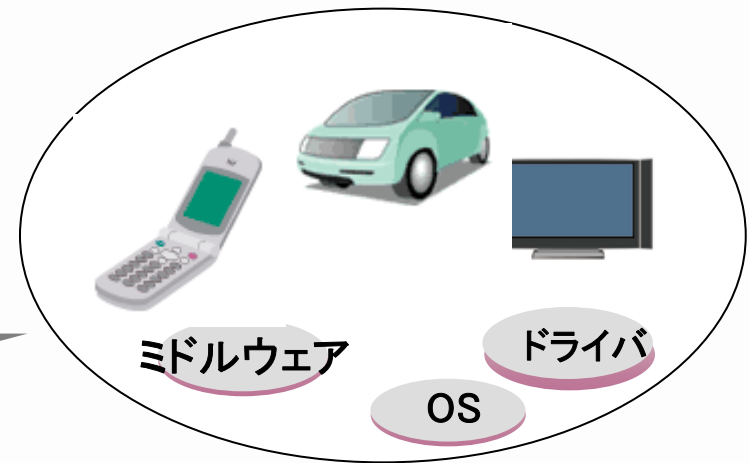
2008年 11月26日

(株)日立製作所 井奥 章
(株)日立製作所 小川 秀人
(株)日立製作所 森本 義章
(株)ルネサステクノロジ 館内 嗣治

- 1 背景(集団全体の評価に関わる問題認識と工夫)
- 2 生産性の評価に関わる留意点
- 3 当該組織での生産性評価の現状と課題
- 4 生産性の評価方法の提案
- 5 提案の適用
- 6 現場の意識の変化
- まとめ

1 背景（集団全体の評価に関わる問題認識と工夫）

- 多品種化・多様化
 - システム系と単品系
 - 顧客特有カスタム品と汎用品
 - 新規、改造、移植
- これらが同一組織内で同時進行



個々のマネジメント

だけでなく

組織全体のマネジメント・評価

も重要

品質データの数値にも多様さ(“ばらつき”)が生じるなか、
いかにマネジメントするか、どのような指標で可視化するのか

1.1 集団全体の評価(要約)に対する問題意識

■一般的な指標(算術平均)は、極端な値(外れ値, 例外値)の影響を強く受ける。分布形状が非対称では、平均=“真ん中”という感覚とずれる。

■分布の実態を表現できる“ものさし”を提供したい。

年収の例

- Aさん 1億5000万円
- Bさん 1200万円
- Cさん 800万円
- Dさん 450万円
- Eさん 350万円
- Fさん 350万円
- Gさん 300万円
- Hさん 300万円
- Iさん 220万円
- Jさん 220万円

算術平均: 1919万円

- 全体の半数が, 300万円周辺に分布。
- 実感としては, ”平均300~350万円”

全数の半分が
“平均”以下に分布する
とは限らない

1.2 全体を要約するための従来技術

■従来技術 統計学的知見を参照し、“ものさし”の使い分けに注目

一般的な用法

分布の特徴	代表値の指標	分布形状(ばらつき)の指標
対称	算術平均値	分散・標準偏差,四分位偏差
非対称	中央値(*1)	四分位偏差, 範囲(最大-最小)

(*1) 中央値

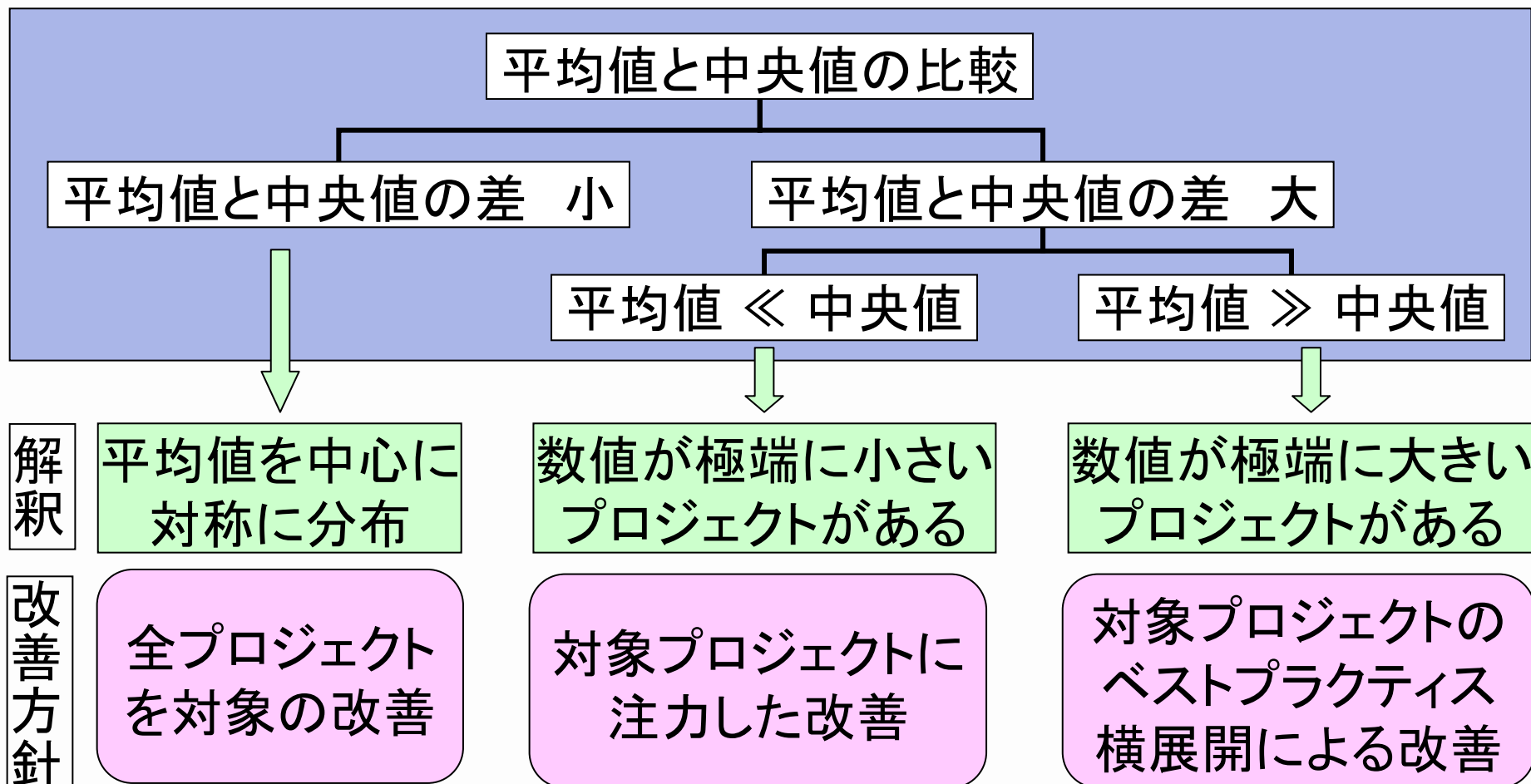
データ群を、同数の上下2つの分布にわける境界(中央)に存在する値

- ✓ 外れ値など極端な値に影響されにくい
- ✓ 算術平均と比べて数理的操作が得意でない

工夫: 複数の代表値を併用し, 中心的傾向と分布形状の表現を両立させる

■ 平均値(算術平均)と中央値からの状況解釈に基づく評価

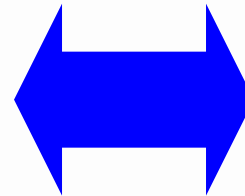
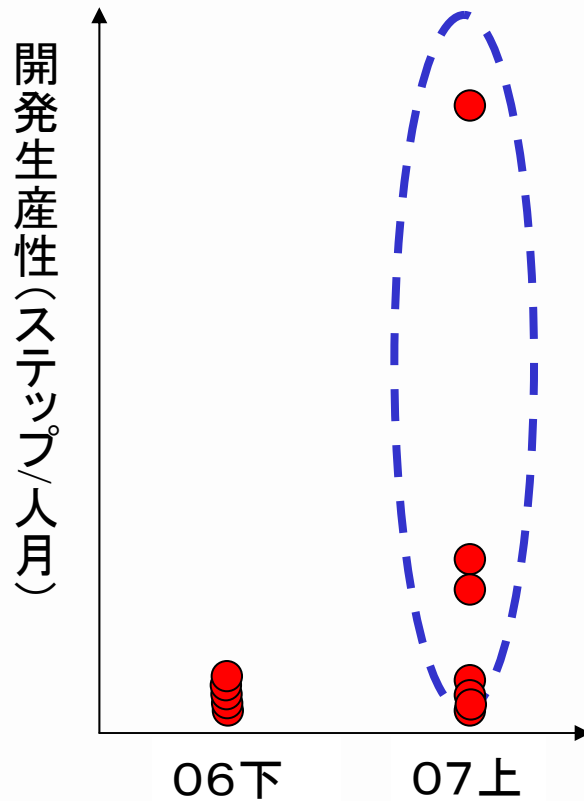
▶ ばらつきの指標「分散」を用いずとも, 分布状況を表現できる



(注) 数値が高いことが望ましい指標の場合

■組織の実態

- ① 中心的傾向は少し悪化
- ② 個別には生産性の良いプロジェクトは07上の方が多い

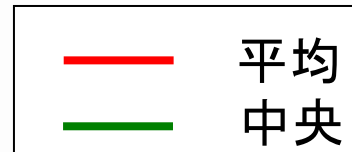
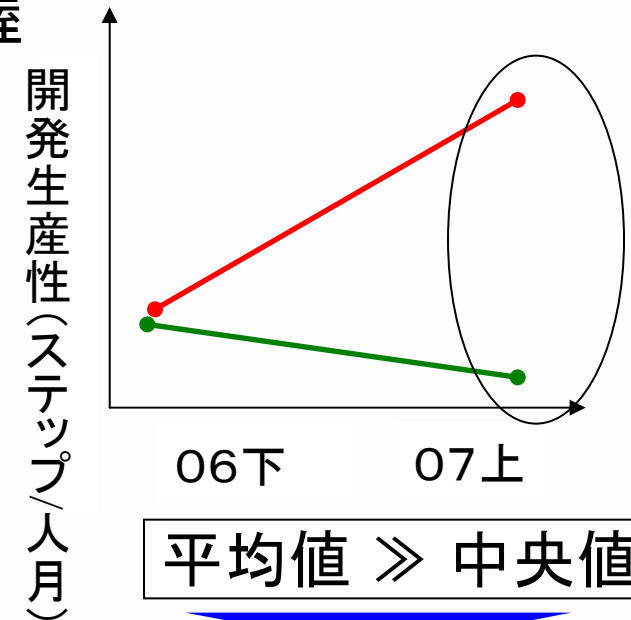


■施策:統計的表現の併用

- ① 中央値 ② (算術) 平均値

■可視化の効果

- 平均値と中央値の併用だけで、組織全体の傾向(分布)も把握



生産性の数値が極端に大きいプロジェクトがある

2 生産性の評価に関する留意点

■ 組織全体の生産性（開発“効率”）

要約の仕方で倍違う

- ① 全案件(プロジェクト)の**生産性の算術平均** = 0.444
- ② $\frac{\text{全案件(プロジェクト)の**規模の総和**}}{\text{全案件(プロジェクト)の**工数の総和**}}$ = 0.212

案件名	規模	工数	生産性(規模／工数)
A	120	1450	0.083
B	140	420	0.333
C	210	660	0.318
D	100	390	0.256
E	128	458	0.280
F	101	340	0.296
G	10	6	1.667
H	43	512	0.084
I	68	100	0.680

数値は仮想的なデータ

生産性のような“率”の指標では、算術平均だけでなく、②を使いやすい。

その意味合いへの認識が希薄な使用も見られ、注意を要する。

2.1 生産性の評価指標の解釈

②

全案件(プロジェクト)の規模の総和
全案件(プロジェクト)の工数の総和

■一見すると

【解釈1】 複数の案件を全体として
一つの仮想的なプロジェクトと見立てた場合の
開発生産性に相当

個々の案件の生産性の分布(ばらつき・形状の変化)の情報が失われる

■実は

【解釈2】 全案件の**生産性の加重平均**に相当
工数が“重み”となる。

少数の大工数案件の生産性が、多数の小工数案件の生産性を覆い隠す

2.2 加重平均であることの説明(1)

- 加重平均は、複数の変数の総合点やその平均の比較をおこなう場合、各変数の比重を一律でなく変化をつけて算出する方法である。
- 各変数に設定する比重の変化を重み・ウェイトという。

定義式: n 個のデータ x_1, x_2, \dots, x_n に対する加重平均は、

$$\frac{\sum_{i=1}^n w_i x_i}{\sum_{i=1}^n w_i} \quad (\text{ただし, } w_i \geq 0, i=1, 2, \dots, n) \quad \dots \text{(式1)}$$

w_i を重みと呼ぶ。

便宜上、重みは $\sum_{i=1}^n w_i = 1$ となるように配分されることが多く、

$$(\text{ただし } \sum_{i=1}^n w_i = 1, w_i \geq 0, i=1, 2, \dots, n) \quad \dots \text{(式2)}$$

•(加重平均についての概略は参考文献[1][2][3][4]を参考に記載した)。

2.2 加重平均であることの説明(2)

$$\begin{aligned} \text{PRJ(プロジェクト)の規模} &= (\text{PRJの規模} / \text{PRJの工数}) * \text{PRJの工数} \\ &= \boxed{\text{PRJの開発生産性} * \text{PRJの工数}} \quad \text{なので,} \end{aligned}$$

全案件(プロジェクト)の規模の総和

を展開すると,

全案件(プロジェクト)の工数の総和

$$\begin{aligned} &= \left\{ \begin{aligned} &\{ \text{PRJ1の開発生産性} * \text{PRJ1の工数} \} / (\text{全プロジェクトの工数の総和}) \\ &+ \\ &\{ \text{PRJ2の開発生産性} * \text{PRJ2の工数} \} / (\text{全プロジェクトの工数の総和}) \\ &+ \\ &\{ \text{PRJ3の開発生産性} * \text{PRJ3の工数} \} / (\text{全プロジェクトの工数の総和}) \\ &+ \\ &\{ \text{PRJ4の開発生産性} * \text{PRJ4の工数} \} / (\text{全プロジェクトの工数の総和}) \\ &+ \end{aligned} \right. \\ &= X_1 W_1 + X_2 W_2 + X_3 W_3 + \dots = \sum_{i=1}^n W_i X_i \\ &\left\{ \begin{aligned} &X_i : \text{PRJ}_i \text{の開発生産性, } W_i : \text{PRJ}_i / \text{全PRJの工数の総和} \\ &(\text{全PRJの工数の総和に対するPRJ}_i \text{の工数の重み}) \end{aligned} \right. \end{aligned}$$

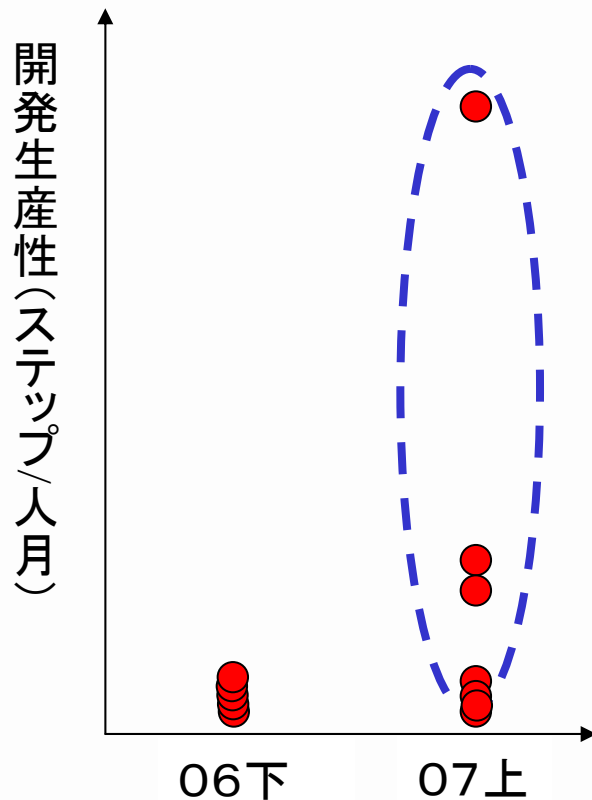
3 当該組織での生産性評価の現状と課題

■「組織の生産性」

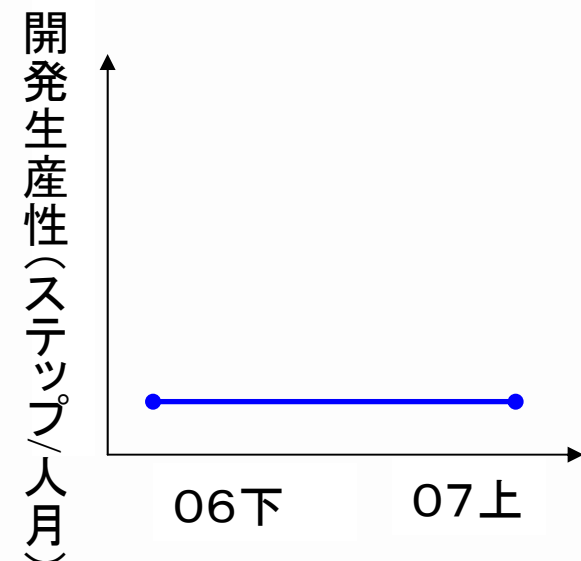
$\Sigma(\text{プロジェクト規模}) \div \Sigma(\text{プロジェクト工数})$ で評価

→ 下記の2つの期間で横ばいの結果となる。

➤ 現場の直感: 07上の方が高い生産性のプロジェクトが多いはず

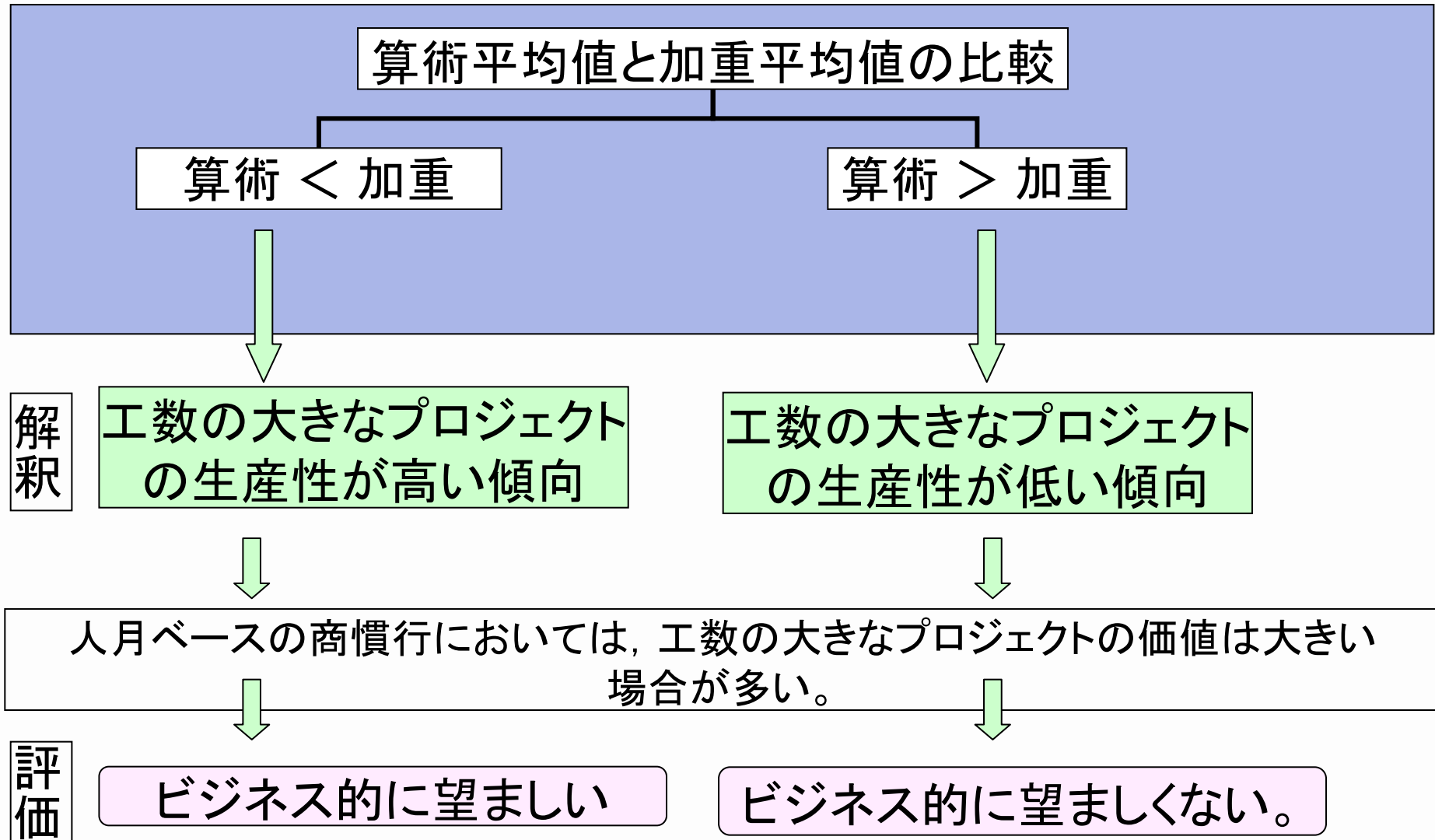


やはり、個々のプロジェクトの生産性の分布情報が失われている



4 生産性の評価方法の提案

■算術平均値と加重平均値からの状況解釈に基づく評価



5 提案の適用

■組織の実態

①人月の大きなプロジェクトに
生産性がよくないものが存在(07上)

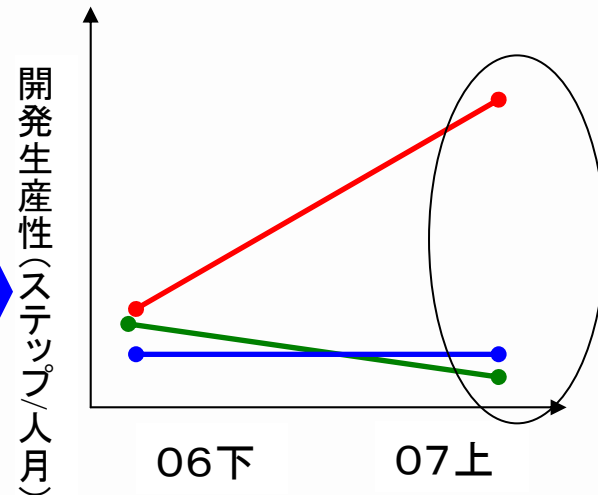
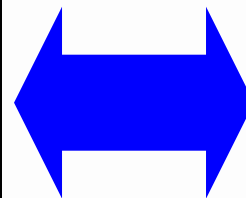
■提案:統計的表現の併用

①算術平均 ②加重平均

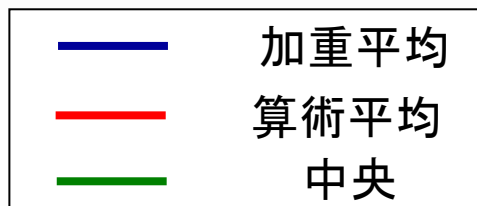
■可視化の効果

➤工数の大きな案件の生産性の傾向を直観的に把握

	生産性 (組織全体の算術平均値を1とする)
工数最大の案件	0.13
工数2位の案件	0.30



算術平均 ≫ 加重平均

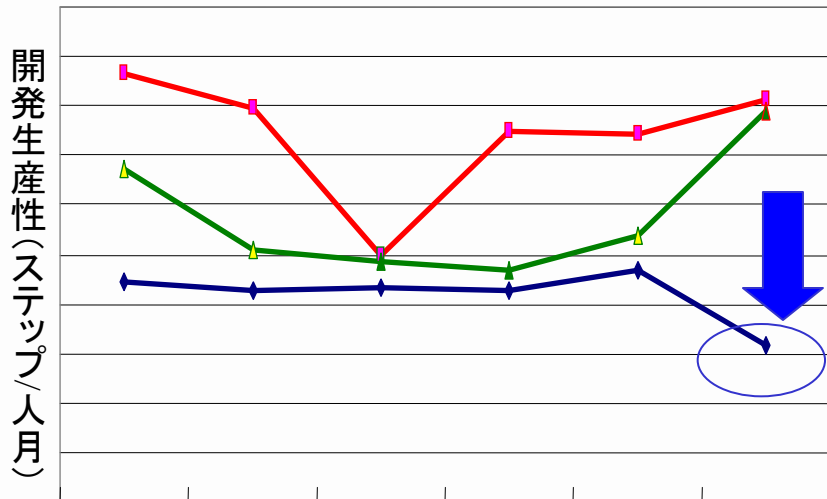


工数の大きなプロジェクトに
生産性の数値が小さい傾向
→ 全体を押し下げている

6 現場の意識の変化(定着)

組織全体の生産性

◆加重平均 ■算術平均 ▲中央値



製品群Kを除いて可視化

◆加重平均 ■算術平均 ▲中央値



5/下 6/上 6/下 7/上 7/下 8/上(予)

5/下 6/上 6/下 7/上 7/下 8/上(予)

■調査: 08上に加重平均の低下理由

■推測: 工数大のプロジェクトにて
生産性の数値が小

■事実確認: 製品群Kが該当

✓一部のサンプルが全体の評価
を大きく左右したことを確認

✓客観的な改善指針の抽出へ

■ 組織全体の生産性の指標選択のガイドを提示

① 全案件(プロジェクト)の生産性の算術平均

- 個々のプロジェクトの改善努力を平等に扱う

② $$\frac{\text{全案件(プロジェクト)の規模の総和}}{\text{全案件(プロジェクト)の工数の総和}} \dots\dots \text{全案件の生産性の加重平均}$$

- 工数の大きなプロジェクトの改善努力を重視する(ビジネス視点を重視)

※注意: 少数の大工数のプロジェクトの生産性が

多数の小工数のプロジェクトの生産性を覆い隠す

■ 複数の指標の併用で表現力を, 簡単に高める

- 算術平均値(①)と加重平均値(②)の大小比較も有効
→ 開発工数の大きなプロジェクトの生産性の状況を直観的に把握。

■ 今回の取組みを通じての教訓

平均することによって

- 消してよいのは「誤差」
- 消していけないのは「個性」

(出所:参考文献[3])

皆様の組織でのプロセス改善活動においては、いかがでしょうか？

ご議論・ご助言いただければ幸いです。

ご清聴ありがとうございました。

■ 参考文献

- [1] 門倉貴史：統計数字を疑う，光文社，2006
- [2] 岩淵千明：あなたもできるデータの処理と解析，福村出版，1997
- [3] 上田尚一：統計学の基礎，朝倉書店，2002
- [4] 吉田耕作：直感的統計学，日経BP社，2006

他業界にみる関連事例 加重平均と算術平均の相互関係に基づく状況把握(「NT倍率」)[*1]

- NT倍率:日経平均株価を東証株価指数(TOPIX)で割った値
- 日経平均は単純平均
 - 株価水準の高い銘柄に左右されやすい
 - ハイテク関連株の影響大
- TOPIXは銘柄ごとの時価総額による加重平均
 - 大型株の影響を受けやすい
 - 内需関連株(銀行株など)の影響大
- NT倍率の上昇
 - ハイテク関連株が相対的によく買われている
- NT倍率の低下
 - 内需関連株が相対的によく買われている

[*1] 参考文献[1]を参考に記載した。