

プログラム作成能力のSN比による分析

2K-6

高田 圭[†] 高橋 宗雄^{††} 山内 成志^{†††} 矢野 宏[†]
 電気通信大学 桐陰学園横浜大学 三菱電機㈱ 電気通信大学

1. はじめに

ソフトウェアの生産性や品質に影響を及ぼす主要な人的要因であるプログラム作成能力については、これまでも多くの研究が行われている。しかし、プログラム作成能力は本質的には人間の問題に帰着するため、具体的な数値に基づく分析例はそれほど多くない。そこで、ハードウェアの開発で大きな効果をあげている品質工学に期待し、品質工学的手法を導入して実験及び解析を行った結果を示す。

まず、品質工学で評価することの信頼性を確かめるため、第一段階として、エラー発見能力とそれに関する要因を定量的に分析、評価する方法を開発し、実験を行った。因子として、表1の11種類を選択し、解析方法は標準SN比というものを使った。結果としては、教育時間は長いほうが良い、教育方法は講義で行ったほうが良いなど、ほとんどの因子が経験とあわせてみても納得のいくものとなった。

第一段階実験は確認実験も行われており、ほぼ再現性が得られている。

表1 エラー発見能力評価の実験の因子

因子	水準	1	2
A.教育時間		4h	8h
B.教育方法		講義	ソフト
C.レビュー人数		1人	2人
D.レビュー時間		30分	20分
E.チェックリスト		有り	無し
F.国語力		高い	普通
G.事務処理能力		高い	普通
H.目標値		あり	なし
I.知的能力		高い	普通
J.作業中断		あり	なし
K.エラー数		少	多

Analysis of ability evaluation by quality engineering

[†] University of Electro-Communications

^{††} Toin University of Yokohama

^{†††} Mitsubishi Electric Corp.

本稿では、この実験からさらに1歩踏み込み、プログラム作成実験を行い、プログラム作成能力に関する要因とプログラムの信頼性(単位規模あたりのエラー数)との関係を分析した結果を示す。

2. 解析方法

まず、被験者に仕様書を渡し、エディタでプログラムを作成してもらい、被験者が納得のいくプログラムができた時点でプログラム(データ1)をセーブする。次に、ここで初めてコンパイラにかけ、被験者自身にデバッグを行ってもらい(データ2)。ここでデータ1とデータ2のプログラムを比較し、異なっている部分をすべてエラーとして、データを取った。

解析方法(SN比の計算)は以下の通りである。

表2 データの取り方

行目	1	2	3	……	n
$y_{正}$	y_1	y_2	y_3	……	y_n

(y はその行が正しければ1, エラーならば0.)

データ2の行数 n に対し、データ1のプログラムが正しいものを記述した割合 p は

$$p = \frac{y_1 + y_2 + \dots + y_n}{n} \quad (1)$$

全変動は、 y が0と1なので、

$$S_T = y_1^2 + y_2^2 + \dots + y_n^2 = np \quad (2)$$

信号の効果は、式(1)の変動より、

$$S_p = \frac{p^2}{(\text{係数の2乗和})} = \frac{\left(\frac{y_1 + y_2 + \dots + y_n}{n}\right)^2}{\left(\frac{1}{n}\right)^2 + \left(\frac{1}{n}\right)^2 + \dots + \left(\frac{1}{n}\right)^2} = \frac{(y_1 + y_2 + \dots + y_n)^2}{n} = np^2 = S_m \quad (3)$$

誤差変動は

$$S_e = S_T - S_p = np - np^2 = np(1-p) \quad (4)$$

0, 1 データのSN比は信号の変動と誤差変動の比で表わされ、測定し加法性を持たせるために、デシベルで表現すると次のようになる。

$$\eta = 10 \log \frac{S_p}{S_e} = 10 \log \frac{np^2}{np(1-p)} \quad (5)$$

$$= 10 \log \frac{p}{(1-p)} = -10 \log \left(\frac{1}{p} - 1 \right) \quad [db]$$

3. 実験

L₁₂直交表を使い、因子としては表4の10種類を選択した。能力テストについては市販されているIQテストを使用した。言語はC言語を使用し、プログラムの規模は100ステップ程度である。被験者は10時間ほどC言語の教育を受けた工学系の学生6人を選択した。この被験者をL₁₂直交表の水準に基づき、実験を行った。

次にエラーを、表5の4種類に分類する。このエラーを新たにL₈直交表に割り付け、水準1を解析に使用した場合、水準2を使用しなかった場合とし、選択した因子とエラーの数及び種類の関係の解析をL₁₂直交表とL₈直交表の直積により行った。すなわちエラーの重要性の評価である。

4. 結果とまとめ

得られたSN比を分散分析にかけ、それらの結果から以下のようなことが定量的に示された。

仕様の量(H)は量が少ないほうが、数理能力は高いほうが、SN比が高くなった。これは経験と一致するものである。しかしながら、国語力(F)は、「普通」のほうが「高い」よりSN比が高くなったなど、興味の引く結果が出た因子も存在した。

エラーの種類では、論理エラーの2つの因子(A', B')が、この実験においては重要な因子であるという結果が得られた。また、IとB'においては交互作用も見られた。

今回はこのような結果となったが、被験者に学生を使っていること、能力テストの値が信用できないなどの問題が残っている。

なによりも、この方法によりプログラム作成能力に関与する要因とプログラムの信頼性(単位規模あたりのエラー数)との関係を定量的に分析できる可能性を示すことができたことが最大の結果である。

表3 L₁₂直交表

no.	A	B	C	D	E	F	G	H	I	e	J
1	1	1	1	1	1	1	1	1	1	1	1
2	1	1	1	1	1	2	2	2	2	2	2
3	1	1	2	2	2	1	1	1	2	2	2
4	1	2	1	2	2	1	2	2	1	1	2
5	1	2	2	1	2	2	1	2	1	2	1
6	1	2	2	2	1	2	2	1	2	1	1
7	2	1	2	2	1	1	2	2	1	2	1
8	2	1	2	1	2	2	2	1	1	1	2
9	2	1	1	2	2	2	1	2	2	1	1
10	2	2	2	1	1	1	1	2	2	1	2
11	2	2	1	2	1	2	1	1	1	2	2
12	2	2	1	1	2	1	2	1	2	2	1

表4 L₈直交表

no.	A	B	C	D	E	F	G
1	1	1	1	1	1	1	1
2	1	1	1	2	2	2	2
3	1	2	2	1	1	2	2
4	1	2	2	2	2	1	1
5	2	1	2	1	2	1	2
6	2	1	2	2	1	2	1
7	2	2	1	1	2	2	1
8	2	2	1	2	1	1	2

表5 因子の選定と各水準

因子	水準	
	1	2
A.プログラムの種類	事務	数理
B.教育方法	講義	ソフト
C.論理的技法	F.C	PAD
D.締切り	3日	2日
E.コーディングガイドライン	有り	無し
F.国語力	高い	普通
G.事務処理能力	高い	普通
H.仕様の量	少	多
I.知的能力	高い	普通
J.数理能力	高い	普通

表6 エラーの選定と水準

因子	水準	
	1	2
A'.論理エラー	mistake	使用 未使用
B'.	unknown	使用 未使用
C'.構文エラー	mistake	使用 未使用
D'.	unknown	使用 未使用

参考文献

高田・高橋・山内・矢野、「コンピュータープログラマの能力評価の基礎研究」第25回官能検査シンポジウム、日科技連、1995