

開発環境特性の統計分析を用いた エラー予測モデルの評価

高橋宗雄 釜池祐治
(NTT 電気通信研究所)

本報告では、先に筆者らが提案した、設計仕様の変更度、開発要員の熟練度などの開発環境特性要因の統計分析を用いたエラー予測モデルについて、実際のソフトウェア開発で得られたデータを用いてその予測精度を評価する。さらに、モデルの説明変数である環境特性要因を変化させたときの予測精度への影響を分析し、予測精度の向上法について検討を加える。その結果、モデルの精度向上に有効な環境特性要因の選定条件を明らかにし、これに基づいて導出したモデルの予測誤差は観測値に対して25%程度であることを示す。

A PRACTICAL EVALUATION OF A MODEL FOR PROGRAM ERROR PREDICTION

Muneo Takahashi and Yuji Kamayachi
NTT Electrical Communication Laboratories

This paper describes a practical evaluation of a model to estimate the number of errors detected in the testing phase of development. The model is derived based on the statistical analysis of environmental factors such as frequency of program specification change. In order to test the application of the model to the actual software development, the predicting reliability of the model is then examined by using the data collected during the development of large-scale software systems. Additionally, an effect on the predicting reliability in increasing the independent variables of the model is also examined. Result of the examinations indicates that the model estimates the program errors with error of about 25% against the actual program errors.

1. まえがき

プログラムのエラー予測モデルに關しては、この提案でプログラムを設計する際には、高精度の予測モデルを必要とする。プログラムのエラー予測モデルは、プログラムの開発経験に基づいて、環境特性要因の分析から、プログラムのエラーに影響を与える環境特性要因として、ソフトウェア要員の熟練度、プログラミング技法、プログラムの改造規模の比率などが指摘されている⁽²⁾⁽³⁾。筆者らは、先に、これら指摘や大規模ソフトウェアの開発経験に基づいて、環境特性要因の仮説をたて、大規模ソフトウェアの開発データを用いて、単位的に分析した⁽¹⁾。その結果、エラー率に影響を与える環境特性要因として、表1に示す6個の要因が明らかとなり、このうち、設計仕様の変更度、ソフトウェア要員の経験年数およびドキュメント密度については危険率5%以上の有意性が認められた。

2. エラー予測モデル

2.1 環境特性要因の分析

従来から、プログラムのエラーに影響を与える環境特性要因として、ソフトウェア要員の熟練度、プログラミング技法、プログラムの改造規模の比率などが指摘されている⁽²⁾⁽³⁾。筆者らは、先に、これら指摘や大規模ソフトウェアの開発経験に基づいて、環境特性要因の仮説をたて、大規模ソフトウェアの開発データを用いて、単位的に分析した⁽¹⁾。その結果、エラー率に影響を与える環境特性要因として、表1に示す6個の要因が明らかとなり、このうち、設計仕様の変更度、ソフトウェア要員の経験年数およびドキュメント密度については危険率5%以上の有意性が認められた。

2.2 モデルの設定

(1) モデル式

プログラムのエラー数は、そのプログラムのエラー率を推定し、次式により求める。

$$B = f \cdot s \tag{1}$$

- ここで、
- B = エラー数 (件)
- s = プログラム規模 (KS)
- f = 推定エラー率 (件 / KS)

* プログラムの規模とエラー数および環境特性要因との間には強い相関があることから、分析においてはプログラム規模の影響を除くため、エラー数、環境特性要因をプログラム規模でそれぞれ正規化した。

表 1 分散分析結果

		F 値	寄 与 率 (%)
プログラム種別	(PTYP)	2.47	16.07
設計仕様の変更度	(SCHG)	5.79**	33.14
プログラムの経験年数	(ISKL)	3.91*	23.15
プログラム構成比	(ICON)	2.44	9.04
開発技法の水準	(TLVL)	2.87	16.20
ドキュメント密度	(DOCC)	3.70*	21.86

* 危険率 5 % 有意
** 危険率 1 % 有意

プログラムのエラー率は、表1の環境特性要因がプログラムの開発において複雑に作用しあうことにより変動する。この現象は多次元回帰であり、エラー率と環境特性要因との関係はエラー率を目的変数、環境特性要因を説明変数とする以下の線形重回帰式で表わすことができる。

$$\hat{f} = a_0 + a_1 f_1 + a_2 f_2 + \dots + a_n f_n \quad (2)$$

ここで、

\hat{f} = 推定エラー率 (件/KS)

f_1, f_2, \dots, f_n = 環境特性要因

a_0, a_1, \dots, a_n = パラメータ

ここでは、環境特性要因として分析の結果、危険率 5 % 以上で有意と判定された設計仕様の変更度 (f_1)、プログラムの経験年数 (f_2) およびドキュメント密度 (f_3) を用いて最小 2 乗法によりパラメータを推定する。

推定結果を表 2、表 3 に示す。

表 2 から \hat{f} は

$$\hat{f} = 67.98 + 0.46f_1 - 9.69f_2 - 0.08f_3 \quad (3)$$

と表わせる。

(2) モデルの検定

モデルの妥当性を判断するため、重回帰分析の結果得られた重相関係数 (R) と分散分析表について考察する。

重相関係数 R は、説明変数の観測値と重回帰式 (3) による推定値との相関係数であり、この値が高い程モデルの精度は良好であるといえる。

このモデルでは表 2 からわかるように $R = 0.78$ であり、従来、経験的に必要とされている値 (0.85) (4) を満足していないが、R の

大小は直ちにモデルの良否を決定するものではない。これについては4章で考察する。

分散分析表はモデルの有意性を検定するものであり、このモデルは表3に示すように危険率1%以上で有意であるといえる。

表2 重回帰分析結果

	Regression Coefficient	Standard Deviation	T ratio	Degree of freedom	Probability
SCHG	0.4579	0.130	3.516	20	0.0022
ISKL	-9.687	3.714	-2.608	20	0.0166
DOCC	-0.083	0.0419	-1.974	20	0.0624

Constant = 67.98
r = 0.7754
r square = 0.6012

表3 分散分析表

	Sum of Square	Degree of freedom	Mean Square	F ratio
Due				
Regression	3256.6	3	1085.5	10.05**
Residual	2160.5	20	108.02	
Total	5417.1	23		

** Significant at the 0.01 level

3. モデルの予測精度の評価

実際のソフトウェア開発に適用した場合のモデルの予測精度を確認するため、大規模ソフトウェアの開発プロジェクトから12個のサンプルプログラムを選定して評価を行った。

3.1 評価用プログラム

評価に用いたプログラムは、言語プロセッサの分野から8本、制御プログラムの分野から4本の計12本を選定した。これらのプログラムの規模は、約5Ks~100Ksの範囲に分布している。

3.2 評価方法

提案モデルと、従来一般的に用いられている以下の2つのモデルについて、観測値に対する予測値の信頼度(精度)を求め、比較評価した。

(1) 平均エラー率モデル

ソフトウェア開発で一般的に用いられている予測方法であり、経験に基づいた単位規模当りのエラー数(エラー率)の平均を用いてエ

ラー数を予測するモデル（図1の点線参照）。

(2) 単回帰モデル

エラー数とプログラム規模との間に強い相関があることから⁽⁵⁾⁽⁶⁾、エラー数を目的変数、プログラム規模を説明変数とする単回帰式を用いてエラー数を予測するモデル（図1の実線参照）

3. 3 評価結果

各モデルの予測精度を表4に示す。提案モデルは、従来一般的に使用されている平均エラー率モデルに比べ予測精度が高く、観測値に対しては26%（平均）程度の誤差で予測可能である。

4. モデルの精度向上法

モデルの予測誤差の原因として、①観測データに測定誤差がある、②モデルに導入した3つの環境特性要因だけではエラー変動を十分説明できない、ことがあげられる。①はエラーデータ測定法の問題であり、ここではモデルの精度向上法の観点から②について観察する。表1によると、モデルに導入した3つの環境特性要因の他に危険率5%以上の有意性はないが、プログラム種別、開発技法の水準がエラー率に影響を与えていると考えられる。そこでこれらの要因をモデルに導入することにより予測精度がどの程度向上するかについて検討する。

4. 1 分析実験

(1) 実験モデルの導出

3章で示したモデルにプログラム種別および開発技法の水準をモデルの説明変数として導入した場合のモデル式と各モデルの重相関係数および有意性の検定結果を表5に示す。

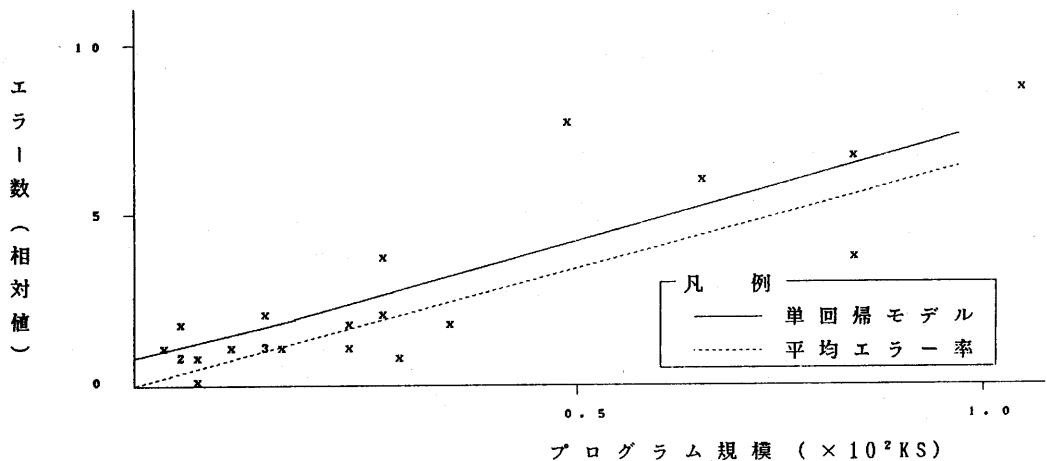


図1 プログラム規模とエラー数の関係

表 4 予測精度の評価結果

モデル サンプルプログラム		平均エラー率	単回掃モデル	提案モデル
		モデル		
言語 プロセッサ	P 1	0.43	0.49	0.14
	P 2	1.32	1.44	0.16
	P 3	0.27	0.27	0.64
	P 4	0.22	0.34	0.24
	P 5	0.41	0.38	0.44
	P 6	0.21	0.11	0.11
	P 7	0.36	0.30	0.004
	P 8	0.02	0.02	0.21
制御 プログラム	P 9	0.20	0.23	0.07
	P 10	0.66	1.24	0.15
	P 11	0.61	1.05	0.36
	P 12	2.23	2.06	0.56
平均		0.58	0.66	0.26

(注) 予測精度 = (観測値 - 予測値) / 観測値

(2) 予測精度の向上度評価

表 6 に評価結果を示す。表 6 から次のことが推察される。

- i) M1 にモデルの説明変数としてプログラム種別を導入した M2 では、重相関係数は高くなるが (表 5)、モデルの予測精度はそれに比例して向上しない。
- ii) さらに、M2 に開発技法の水準を追加した M3 では重相関係数、予測精度ともほとんど変化がない。

4. 2 予測精度の向上条件

前述の評価結果によれば、説明変数を増せば重相関係数 (R) をある程度まで高くできるが、R を高くすることが必ずしもモデルの精度を高くするとはかぎらない。この理由は、説明変数が多くなれば、①自由度が減少するので残差 (誤差) の変動 (SE) が大きくなる、②寄与率の小さい説明変数が寄与率の大きい説明変数の回帰変動を減少させる方向に作用し、その結果、回帰式により説明できる変動 (SR) が小さくなる可能性があることによるためと考えられる。

表 7 はこれらの様子を示したものである。したがって、最小の説明変数で最大の精度を得るためには、分散分析の結果寄与率が高い要因から選定することが有効であり、危険率 5% 以上で有意であることが選定の目安といえる。

表 5 モデル式と有意性の検定結果

モデル		モデル式 (注1)	重相関係数	F 比
提案モデル (M1)		$\hat{r} = 67.98 + 0.46f_1 - 9.69f_2 - 0.08f_3$	0.7754	10.05 **
実験モデル	SCTGの導入 (M2)	$\hat{r} = 64.16 + 0.43f_1 - 7.73f_2 - 0.09f_3 - \begin{cases} 0 \\ 7.0 \\ 2.94 \\ 7.5 \end{cases}$ (注2)	0.7999	6.81 **
	TLVLの導入 (M3)	$\hat{r} = 68.96 + 0.43f_1 - 7.60f_2 - 0.08f_3 - 25.15f_4 - \begin{cases} 0 \\ 6.2 \\ 2.3 \\ 7.2 \end{cases}$ (注2)	0.8004	5.60 **

(注1) f : エラー率 f₁ : SCHG f₂ : ISKL f₃ : DOCC f₄ : TLVL

(注2) 分野別定数 : 通信制御、言語、データベース、制御プログラム

表 6 予測精度の向上度

モデル		M ₁	M ₂	M ₃
サンプルプログラム				
言語プロセッサ	P 1	0.14	0.34	0.10
	P 2	0.16	0.04	0.25
	P 3	0.64	0.48	0.49
	P 4	0.24	0.28	0.21
	P 5	0.44	0.35	0.45
	P 6	0.11	0.20	0.18
	P 7	0.004	0.08	0.07
	P 8	0.21	0.21	0.21
制御プログラム	P 9	0.07	0.01	0.02
	P 10	0.15	0.28	0.28
	P 11	0.36	0.23	0.32
	P 12	0.56	0.43	0.54
平均		0.26	0.25	0.26

表 7 説明変数の数と残差の変動との関係

比較項目 \ モデル	M ₁	M ₂	M ₃
説明変数の数	3	4	5
残差の自由度	20	19	18
残差の変動	8.93×10	9.21×10	9.6×10
回帰による変動	1.18×10^3	6.27×10^2	5.38×10^2

5. あとがき

本報告では、設計仕様の変更度などの開発環境特性の統計分析を用いたエラー予測モデルについて、実際のソフトウェア開発で得られたデータに基づき評価結果を示した。さらに、モデルの説明変数である環境特性要因を変化させたときの予測精度への影響を分析し、予測精度の向上法について考察した。

主要な結果は以下のとおりである。

- (1) 提案モデルは観測値に対し25%程度の誤差で予測可能であり、従来のプログラム規模ベースの予測方法に比べて予測精度が約30%向上する。
- (2) モデルに導入する環境特性要因を増加させても予測精度はそれに比例して向上しない。環境特性要因の選定は分散分析の有意水準による判断が有効であり、その水準は危険率5%以上とすることが妥当である。今後は、提案モデルを品質管理のための作業要領に組み入れて、実際のプロジェクトに適用しながらリファインをはかる予定である。

参考文献

- (1) M.Takahashi, Y.Kamayachi, "An Empirical Study of a Model for Program Error Prediction," Proc. 8th ICSE, PP. 330-336, Aug. 1985
- (2) L.M.Ottenstein, "Quantitative Estimate of Debugging Requirements," IEEE Trans. on Software Eng., VOL.SE-5, No.5, pp.504-514, Sept. 1979
- (3) V.R.Basili and D.H.Hutchens, "An Empirical Study of Syntactic Complexity Family," IEEE Trans. on Software Eng., VOL.SE-9, NO.6, PP.664-672, Nov. 1983
- (4) 中村, "多変量解析入門", 日刊工業新聞社, pp.16-17, 1979.
- (5) F.Akiyama, "An Example of Software System Debugging," Proc. IFIP Congress, PP.353-358, Aug. 1971.
- (6) T.Sunohara, A.Takano, K.Uehara and T.Ohkawa, "Program Complexity Measure for Software Development Management," Proc. 5th International Conf. on Software Eng., pp.100-106, Mar.1981