

順位による誤差分散を考慮した ソフトウェア信頼度成長曲線の適合度検定

石岡 恒憲
(株)リコー

ソフトウェア研究所

ソフトウェアのバグの発見順位による誤差分散を考慮した信頼度成長曲線の適合度検定法を提案する。本検定法ではまず始めに、種々提案されているソフトウェア信頼度成長曲線を一旦、規準化2重指数分布に置換する。次に規準化2重指数分布の順序統計量の各確率密度分布が正規分布に近似できることを示し、この性質を用いて信頼度成長曲線の適合度検定問題を正規性の検定に帰着させることで実現する。順位相関を考慮した適合度検定であるだけでなく、情報量損失がないため高い検出力をもつ。適合度検定のためのアルゴリズムを具体例を用いて説明し、またバグ推定値の信頼限界の算出方法についても記述する。

A Test of Fitness for Software Reliability Growth Curve,
with Observation Error by Its Order

Tsunenori Ishioka

Ricoh Co., Ltd.
Software Research Center

Koishikawa 1-1-17, Bunkyo-ku,
Tokyo, 112, Japan

This paper presents a technique to test of fitness for software reliability growth curve, with observation error by its order. This technique consists of 1) translation of various software reliability growth curves to normalized double exponential distribution, and 2) application of fitness test of normal distribution, to which the p.d.f of order statistics under the double exponential distribution approximates. This test have not only consideration of the order correlation, but also high power of detection because of no lack of its information. The algorithm of this test is explained with a definite sample, and confidence limit of estimated total bag number can be calculated.

はじめに

開発・生産中のソフトウェアのバグの出現を予測することは、生産物としてのソフトウェアの品質管理、生産管理、および原価管理にとってさまざまな意味で重要である。そのため、ソフトウェア信頼度成長曲線のモデルとしてGompertz曲線やLogistic曲線の他に、指数型モデルや非定常ポアソン過程(Non-Homogeneous Poisson Process)に基づく遅延S字型モデル、加速S字型モデルなどが提案^{5,6,7)}され、実用に供されている。また適合性についての事例研究や解析^{5,6)}が行なわれており、さらに母数推定のための尤度方程式⁷⁾も導かれている。

この信頼度成長モデルを適用してバグ数を推定する場合、実務的な立場からはモデルの選択とその妥当性が問題となる。種々の成長曲線の中からKullback-Leiblerの情報量やAIC(Akaike's Information Criterion:赤池の情報量規準)に基づいて最良なモデルを選択することはできるが、これらの情報量規準でわかるのはあくまでもモデルの相対的な比較である。したがってその成長曲線が現実のデータに適合しているか否かといった絶対的な評価は、モデルの適合度検定を行わなくてはならない。

ところがソフトウェア信頼度成長曲線に関して、バグ・データが順序統計量、かつ、打ち切りデータ(censoring data)であるといった信頼性データのもつ本質的な性質を考慮した適合度検定法は、まだ提案されていない。したがって、適合度検定に検出力の弱い χ^2 検定やKolmogorov-Smirnov検定を用いているのが実情であろう。大場⁸⁾は観測されたバグ・データと推定した成長曲線が傾向的にうねるといった成長曲線の形状の妥当性を評価するために、連の総数に基づく2項検定を提案している。しかしこれとしても、観測データの推定した成長曲線の上側から下側への遷移確率を1/2と仮定しており、バグ・データの順位相関を無視していることに注意すべきである。

そこで本稿では、規準化した2重指数分布の順序統計量の各確率密度分布が正規分布に十分近似できることを示し、この性質を用いて順位による誤差分散を考慮した各種ソフトウェア信頼度成長曲線の適合度検定法を提案する。本検定法の妥当性を示す理論的背景を述べ、適合度検定として検出力の強いNeymanのスムーズ・テスト⁹⁾やShapiro-Wilkの検定⁹⁾が利用できることを示す。つまり信頼度成長曲線の適合度検定問題を、正規性の検定に帰着させている。また適合度検定のための具体的なアルゴリズムをその実施例を用いて説明する。

2. ソフトウェア信頼度成長モデル

ソフトウェア信頼度成長モデルとして、以下のモデルがよく知られている^{5,6,7)}(表1)。

Weibull,遅延S字型、加速S字型の各モデルの母数は、表1で示される平均値関数 $h(t)$ を持つ非定常ポアソン過程(Non Homogeneous Poisson Process)に基づいて最尤法により求めることができる。

テスト時刻 t までに発見される累積エラー数を表わす計数過程を $\{c(t), t > 0\}$ により定義する。観測時刻を t_1 , 時間区間 $[0, t_1]$ に発見したバグ数を y_1 とすれば

$$\{c(t_1)=y_1, c(t_2)=y_2, \dots, c(t_k)=y_k\}$$

の同時確率密度関数、すなわち尤度関数は

$$L = \prod_{i=1}^k \frac{\{h(t_i) - h(t_{i-1})\}^{y_i - y_{i-1}}}{(y_i - y_{i-1})! \exp[-h(t_k)]} \quad (1)$$

ただし

$$t_0 = 0, y_0 = 0$$

で与えられる。したがって、それぞれの平均値関数に応じて(1)の対数尤度を最大とする母数が最尤推定値となる。最尤推定値はバグ・データ数が十分大きいときは、漸近的に不偏の多次元正規分布に従う。

3. 誤差分散を考慮した適合度検定法

3.1 信頼度成長曲線の2重指数分布への変換

ソフトウェア信頼度成長曲線のうちWeibullモデルを考えてみよう。Weibull分布の累積分布関数 $F(t)$ は、

$$F(t) = 1 - \exp[-(t/\eta)^m] \quad (2)$$

で示される。したがって、Weibullモデル

$$h(t) = N \{1 - \exp[-(t/\eta)^m]\} \quad (3)$$

とは、累積分布関数 $F(t)$ のノン・パラメトリックな推定値が

$$\hat{F}(t) = \hat{h}(t)/N \quad (4)$$

で示されることを記述しているにすぎない。

Weibull分布において、

$$x = m \cdot \ln(t/\eta) \quad (5)$$

なる変換を施せば、(2)は、

$$F(x) = 1 - \exp[-\exp(x)] \quad (6)$$

と変換される。(6)は一般の2重指数分布

$$F(x) = 1 - \exp[-\exp\{(x-u)/b\}] \quad (7)$$

において、

尺度母数(scale parameter) : $b=1$,

位置母数(location parameter) : $u=0$

とした規準化2重指数分布である。

ここでWeibull分布などで示される信頼度成長モデルの分布と規準化2重指数分布の累積分布関数とを区別するために、以下、前者(2)を $F_N(t)$ 、後者を $F_D(x)$ とおく。

表1 ソフトウェア信頼度成長曲線

モデル	平均値関数
Logistic	$h(t) = \frac{N}{1 + m \cdot \exp[-at]}$ ($a > 0, m > 0$)
Gompertz	$h(t) = N \cdot a \cdot b^t$ ($b < 1, a < 1$)
Weibull	$h(t) = N \{1 - \exp[-(t/\eta)^m]\}$ ($m > 0, \eta > 0$)
指数型	$h(t) = N \{1 - \exp[-\Phi t]\}$
遅延S字型	$h(t) = N \{1 - (1 + \Phi t) \exp[-\Phi t]\}$
加速S字型	$h(t) = N \frac{1 - \exp[-\Phi t]}{1 + (1/r - 1) \exp[-\Phi t]}$ ($0 \leq r \leq 1$)

ただし、

t : 時刻,

N : ソフトウェアに潜在する総エラー数,

Φ : 単位時間当たりのエラー発見率,

r : エラー発見率の増加率.

さて(6)を変形すると、

$$x = \ln \ln [1 / (1 - F_0(x))] \quad (8)$$

となる。したがって縦軸に

$$y = \ln \ln [1 / (1 - F_0(x))] \\ = \ln \ln [1 / (1 - h(t)/N)], \quad (9)$$

横軸に

$$x = m \cdot \ln(t/\eta) \quad (10)$$

をとれば、信頼度成長曲線は、原点を通る傾き1の直線に変換できることがわかる(図1)。

観測されるのはtの点列(t_1, t_2, \dots, t_k)と、それに対応する発見総バグ数であるから、バグ・データが仮定した信頼度成長曲線に適合するならば、(9),(10)より得られる変換値のプロット(x_j, y_j) ($j=1, \dots, k$)は上記の直線にのる。このときプロットは、縦軸方向に誤差をもつことに注意すべきである。

表1に示す信頼度成長モデルにおいて以下の変換を施すことにより、それぞれの分布関数は標準化2重指数分布の分布関数に変換されるから、Weibull分布同様に信頼度成長曲線は原点を通る傾き1の直線に変換できる:

Logistic:

$$x = \ln[\ln\{1 + m \cdot \exp[-at]\} + at + \ln m], \quad (11)$$

Gompertz:

$$x = \ln \ln [1 / (1 - a^{-t})], \quad (12)$$

指数型:

$$x = \ln(-\Phi t), \quad (13)$$

遅延S字型:

$$x = \ln(2\Phi t - \ln \Phi t), \quad (14)$$

加速S字型:

$$x = \ln[\Phi t + \ln\{r + (1-r)\exp[\Phi t]\}]. \quad (15)$$

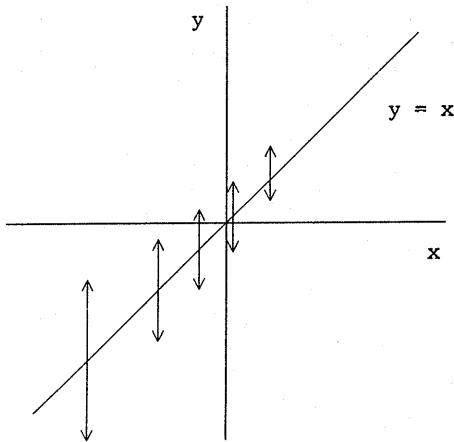


図1 線形化した信頼度成長曲線

3.2 標準化2重指数分布の順序統計量の分布

一般に、分布関数F(x)に従うn個中i番目の大きさの順序統計量の確率密度関数は、

$$g_{nd}(x) = n \binom{n-1}{i-1} f(x) \{F(x)\}^{i-1} \{1-F(x)\}^{n-i} \quad (16)$$

となる。そこで、F(x)に標準化2重指数分布を仮定すると(13)の分布の平均ならびに分散は、

$$\mu_{nd} = n \binom{n-1}{i-1} \sum_{j=1}^i (-1)^j \binom{i-1}{j-1} \frac{\gamma + \ln(n-i+j)}{n-i+j}, \quad (17)$$

$$\sigma_{nd}^2 = n \binom{n-1}{i-1} \sum_{j=1}^i (-1)^{j-1} \binom{i-1}{j-1} \frac{1}{n-i+j} [(\gamma^2 + \pi^2/6)$$

$$+ 2\gamma \ln(n-i+j) + \{\ln(n-i+j)\}^2] - \mu_{nd}^2 \quad (18)$$

となる(誘導は付録Aを参照)。ここに、 γ はEulerの定数($\gamma=0.5772156\dots$)である。

この σ_{nd}^2 の値を標本の大きさ $n=5, 10, 20$ の場合について示したのが、表2である。この表から $g_{nd}(x)$ の分布の分散は、順位iによって1桁はオーダーが異なることがわかる。

表2 標準化2重指数分布に従う順序統計量の確率密度分布 $g_{nd}(x)$ の分散 σ^2

標本数	順位	分散	標本数	順位	分散
n	i	σ_{nd}^2	n	i	σ_{nd}^2
5	1	1.644934	20	1	1.644934
	2	0.645153		2	0.645153
	3	0.395397		3	0.395397
	4	0.284558		4	0.284558
	5	0.308497		5	0.222364
10	6	0.284864	6	0.182710	
	7	0.155326	7	0.155326	
	8	1.644934	8	0.135370	
	9	0.645859	9	0.120270	
	10	0.397015	10	0.108539	
	11	0.287388	11	0.099271	
	12	0.226855	12	0.091890	
	13	0.189578	13	0.086036	
	14	0.165814	14	0.081495	
	15	0.151918	15	0.078190	
16	0.148792	16	0.076196		
17	0.171429	17	0.075848		
		18	0.078077		
		19	0.085783		
		20	0.113145		

また、 $g_{nd}(x)$ の各分布を $n=5$ の場合に実際に示したのが、図2である。この図から $g_{nd}(x)$ の各分布が、 $i=1$ の場合を除き正規分布に近似できることが認識できる。

ここで正規分布からのズレを測る尺度として、3次と4次のキュムラントを分散 σ^2 で規格化した

$$\beta_1 = K_{(3)} / \sigma^3, \quad (19)$$

$$\beta_2 = K_{(4)} / \sigma^4 \quad (20)$$

を用いることにする。 β_1 は非対称性の尺度であり歪度(skewness)と呼ばれる。 β_2 は尖度(kurtosis)と呼ばれるが、そこから連想される分布の中央部の尖り方より、むしろ分布の裾の長さの尺度に用いられる。正規分布のもとでは

$$\beta_1 = \beta_2 = 0$$

である(文献によっては(β_2+3)を尖度と呼ぶ)。

一方、 χ^2 分布はその自由度 Φ が大きくなるにつれて正規分布に近づくことはよく知られている。事実、 χ^2 分布の歪度 β_1 、および尖度 β_2 は

$$\beta_1 = 2\sqrt{2}/\sqrt{\Phi}, \quad (21)$$

$$\beta_2 = 12/\Phi \quad (22)$$

で示され、 $\Phi \rightarrow \infty$ のとき $\beta_1 \rightarrow 0$ 、 $\beta_2 \rightarrow 0$ となる。

そこで $g_{nd}(x)$ の各分布が正規分布に十分近似できることを、その歪度および尖度に対する χ^2 分布の相等自由度が十分大きいことで証明する。

ここで規格化2重指数分布に従う n 個中 i 番目の大きさの順序統計量 $g_{nd}(x)$ の分布の原点回りの3次及び4次のモーメントは、次式で示される(誘導は付録Aを参照)：

$$\begin{aligned} \mu_{nd(3)} &= n \binom{n-1}{i-1} \sum_{j=1}^i (-1)^{j-1} \frac{1}{n-i+j} \binom{i-1}{j-1} [\Gamma'''(1) \\ &\quad - 3\Gamma''(1) \cdot \ln(n-i+j) + 3\Gamma'(1) \cdot \{\ln(n-i+j)\}^2 \\ &\quad - \{\ln(n-i+j)\}^3], \quad (23) \\ \mu_{nd(4)} &= n \binom{n-1}{i-1} \sum_{j=1}^i (-1)^{j-1} \frac{1}{n-i+j} \binom{i-1}{j-1} [\Gamma''''(1) \\ &\quad + 6\Gamma'''(1) \cdot \{\ln(n-i+j)\}^2 + \{\ln(n-i+j)\}^4 \\ &\quad - 4\Gamma''(1) \cdot \ln(n-i+j) - 4\Gamma'(1) \cdot \{\ln(n-i+j)\}^3] \quad (24) \end{aligned}$$

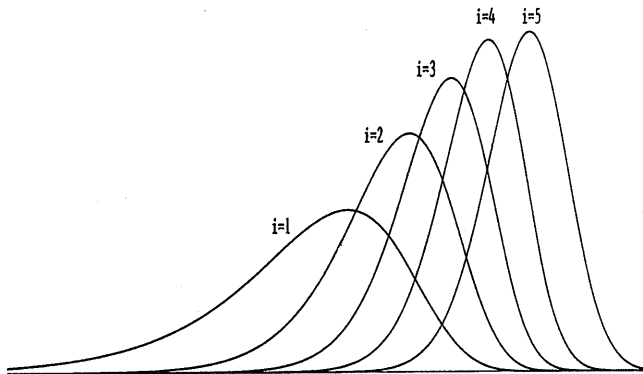


図2 規格化2重指数分布に従う順序統計量の確率密度分布
($n=5, i=1, 2, \dots, 5$)

ただし、

$$\Gamma'(1) = -\gamma = -0.5772156649,$$

$$\Gamma''(1) = \gamma^2 + \pi^2/6 = 1.978111994,$$

$$\Gamma'''(1) = -(\gamma^3 + \gamma\pi^2/2 + \ln^3 \Gamma(1))$$

$$= -5.444874456,$$

$$\Gamma''''(1) = \gamma^4 + \gamma^2\pi^2 + 4\gamma\ln^2 \Gamma(1) + 3\pi^4/20$$

$$= 23.561457369$$

である。またキュムラントと原点回りのモーメントには次の関係が成り立つ：

$$K_{(3)} = \mu_{(3)} - 3\mu_{(4)} \cdot \mu_{(1)} + 2\mu_{(2)}^3, \quad (25)$$

$$K_{(4)} = \mu_{(4)} - 4\mu_{(3)} \cdot \mu_{(1)} - 3\mu_{(2)}^2 + 12\mu_{(2)} \cdot \mu_{(1)}^2 - 6\mu_{(1)}^4. \quad (26)$$

以上より(19)、(20)の関係を用いて、標本数 $n=5, 10, 20$ としたとき $g_{nd}(x)$ の歪度、尖度ならびにそれぞれに対する χ^2 分布の相等自由度を算出し、表3に示す。表3より $g_{nd}(x)$ の各分布が正規分布に十分近似できることがわかる。 n を大きくすれば、この傾向はさらに顕著である。

表3 確率密度分布 $g_{nd}(x)$ の歪度および尖度

標本 n	順位 i	確率密度分布 $g_{nd}(x)$		χ^2 分布の 相等自由度	
		歪度 β_1	尖度 β_2	Φ_1	Φ_2
5	1	-1.1396	2.4000	6	5
	2	-0.7728	1.1724	13	10
	3	-0.5957	0.7211	23	17
	4	-0.4649	0.4684	37	26
	5	-0.3028	0.2419	87	50
10	1	-1.1395	2.4000	6	5
	2	-0.7786	1.1841	13	10
	3	-0.6161	0.7545	21	16
	4	-0.5195	0.5432	30	22
	5	-0.4513	0.4162	39	29
	6	-0.3963	0.3286	51	37
	7	-0.3457	0.2607	67	46
	8	-0.2912	0.2010	94	60
	9	-0.2183	0.1397	168	86
	10	-0.0840	0.0542	1135	221
20	1	-1.1396	2.4000	6	5
	2	-0.7798	1.1867	13	10
	3	-0.6199	0.7608	21	16
	4	-0.5273	0.5541	29	22
	5	-0.4653	0.4334	37	28
	6	-0.4198	0.3543	45	34
	7	-0.3843	0.2983	54	40
	8	-0.3552	0.2564	63	47
	9	-0.3304	0.2235	73	54
	10	-0.3085	0.1968	84	61
	11	-0.2885	0.1744	96	69
	12	-0.2696	0.1548	110	78
	13	-0.2509	0.1375	127	87
	14	-0.2318	0.1214	149	99
	15	-0.2111	0.1062	180	113
	16	-0.1873	0.0912	228	132
	17	-0.1577	0.0759	322	158
	18	-0.1169	0.0601	586	200
	19	-0.0515	0.0442	3019	271
	20	0.8095	0.0249	1221	483

3.3 順位による誤差分散の標準正規分布への変換

3.1では、Weibullモデル(指数型モデル)、遅延S字型モデル、加速S字型モデルが、適当な変換を施すことにより標準化2重指数分布に変換できる、すなわち線型式に変形できることを示した。また3.2では、標準化2重指数分布の順序統計量の各分布が正規分布に近似できることを示した。

したがって、観測値の変換値 y_j ($j=1, \dots, k$)を(17)、(18)で示した各順序統計量 $g_{nd}(x)$ の平均 μ_{nd} ($=\mu_{nd}(t_j)$)、分散 σ_{nd}^2 ($=\mu_{nd}(t_2) - \mu_{nd}(t_1)$)を用いて標準化する

$$z_j = (y_j - \mu_{nd}) / \sigma_{nd}$$

は標準正規分布 $N(0, 1^2)$ に従うことがわかる。ここで

$$y_j = \ln \ln \{1 / (1 - F_n(x_j))\} \\ = \ln \ln \{1 / (1 - F_n(t_j))\}$$

であるから、信頼度成長曲線の仮定が正しければ、最尤法など適当な方法で推定した種々の信頼度成長曲線の累積分布関数 $\hat{F}_n(t_j)$ を用いて、

$$[\ln \ln \{1 / (1 - \hat{F}_n(t_j))\} - \mu_{nd}] / \sigma_{nd} \sim N(0, 1^2) \\ (j=1, 2, \dots, k) \quad (27)$$

- ただし、j: 時間区間順位、
- k: 総時間区間数、
- n: 総バグ数、
- i: バグ順位

となる。ここで具体例で示そう。

表4に示すバグが発見されたと仮定する。ここで時刻 t_j はバグの発見された正確な時間ではなく、時間区間 $[0, t_j]$ までにバグが発見されたことを示す時間の区切りであることに留意する。

このバグ・データに遅延S字型モデル

$$h(t) = N(1 - (1 + \Phi t) \exp[-\Phi t]) \quad (28)$$

を仮定し最尤法で母数を推定すると、

$$\hat{N} = 31.2, \\ \hat{\Phi} = 0.504$$

である。縦軸に(9)を、横軸に(14)をとり、 $j=1, \dots, 10$ における各プロットを結んだのが図3である。回帰直線であるべき $y=x$ も同時に記す。

表4 遅延S型モデルをあてはめたバグ・データ

区 間 順 位	時 刻	バ グ 順 位	観測値 の変換値	標準化2重 指数分布		標準化 した 変換値
				平均	分散	
j	t_j	i	y_j	μ_{nd}	σ_{nd}^2	z_j
1	1	2	-2.346	-3.027	0.6450	0.848
2	2	6	-1.170	-1.675	0.1818	1.184
3	3	18	-0.527	-0.243	0.0530	-1.233
4	4	19	-0.093	-0.158	0.0620	0.261
5	5	23	0.231	0.183	0.0698	0.182
6	6	23	0.489	0.183	0.0698	1.159
7	7	28	0.701	0.657	0.0465	0.204
8	8	28	0.881	0.657	0.0465	1.038
9	9	29	1.038	0.775	0.0507	1.167
10	10	30	1.175	0.912	0.0545	1.127

さてNは発見される総バグ数であるので、整数値でなくてはならない。推定した総バグ数よりも多くのバグが発見される危険を無くすために

$$\hat{N} = 32$$

として、すなわち小数部を切り上げて以下の議論を進める。

j=1の場合に注目する。時刻 t_j までに発見されたバグは2個であるから、 $[t_0=0, t_1=1]$ で発見されたバグは、n=32個中2番目に発見されたバグと考えることができる。(17)、(18)の式を用いることにより、32個中2番目にバグが発見されたときの $\ln \ln \{1 / (1 - F_n(t_1))\}$ の期待値(平均)、および分散はそれぞれ

$$\mu_{32,2} = -3.027, \\ \sigma_{32,2}^2 = 0.6450 = (0.8031)^2$$

となる。一方、観測値の変換値 y_1 は、

$$y_1 = \ln \ln \{1 / (1 - F_n(t_1))\} \\ = \ln \ln \{1 / (1 + \Phi t_1) \exp[-\Phi t_1]\} \\ = \ln \ln \{1 / (1 + 0.504) \exp[-0.504]\} \\ = -2.346$$

となる。よって、

$$z_1 = (y_1 - \mu_{32,2}) / \sigma_{32,2} \\ = 0.848$$

を得る。

j=2の場合も同様に、32個中6番目にバグが発見されたとみなすことにより、 z_2 が算出できる。遅延S字型モデルの仮定が正しければ数列 $(z_1, z_2, \dots, z_{10})$ は標準正規分布に従う。したがって、ソフトウェア信頼度成長曲線の適合度検定を正規性への検定に帰着させることができる。

ここでは簡便のためn=32としたが、より厳密をきすならばn=31.2の μ_{nd} と σ_{nd}^2 の値は、をn=31とn=32の場合のそれぞれを1次補間することにより容易に求めることができる。もっとも μ_{nd} と σ_{nd}^2 の値はnよりもiに依存することがわかっているため、補間することに多くの効果は期待できない。

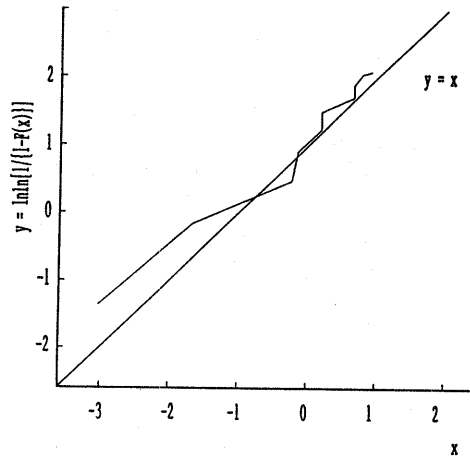


図3 遅延S型モデルをあてはめた信頼度成長曲線

4. 正規性の検定

本章では、3章の手順で得られた点列 z_j ($j=1, \dots, k$) からなる母集団分布が、標準正規分布とみなせるかの検定—正規性の検定—について述べる。

そこで検定すべき仮説を

$$H_0: \text{点列 } z_j (j=1, \dots, k) \text{ は標準正規分布からの標本である}$$

とする。

4.1 Neymanのスムーズ・テスト³⁾

はじめに

$$U_j = \Phi(z_j)$$

とおく。ここに $\Phi(z)$ は標準正規分布 $N(0, 1^2)$ の累積分布関数である。 U_j は $(0, 1)$ 上の確率変数で互いに独立となるから、正規性の仮説は

$$H_0': U_j (j=1, \dots, k) \text{ の分布は } (0, 1) \text{ 上の一様分布である}$$

と同値となる。

そこでNeymanは

$$\Psi_p^2 = \frac{1}{k} \sum_{j=1}^k \left\{ \sum_{i=1}^k \pi_i(U_j) \right\}^2 \quad (29)$$

を検定統計量に選んだ。ここで p は、 U_i ($i=1, \dots, k$)の分布が $(0, 1)$ 上の一様分布でないとしたときに示される分布の未知母数の数である。また $\{\pi_i(u); i=0, 1, \dots\}$ は $(0, 1)$ 上の直交多項系であり

$$\begin{aligned} \pi_0(u) &= 1, \\ \pi_1(u) &= \sqrt{3}(2u-1), \\ \pi_2(u) &= \sqrt{5}(6u^2-6u+1), \\ \pi_3(u) &= \sqrt{7}(20u^3-30u^2+12u-1), \\ \pi_4(u) &= 3(70u^4-140u^3+90u^2-20u+1), \\ &\dots \end{aligned} \quad (30)$$

で示される。

対立仮説の分布はその分布の母数を種々に変え、一様分布から滑らかにずれた $(0, 1)$ 上の分布の族を与える。

Ψ_p^2 ($p=1, 2, \dots$)のうちの検定統計量を用いたら良いかという問題は、現在のところまだ得られていないようである。若干のシミュレーション結果によると、 $p=3, 4$ のいずれを用いても差はあまりないようである³⁾。

ここで表4に示す点列 z_j ($j=1, \dots, 10$)について U_j ($=\Phi(z_j)$)を求めると

$$\{U_j; j=1, \dots, 10\} = \{0.8018, 0.8818, 0.1088, 0.6030, 0.5722, 0.8768, 0.5808, 0.8504, 0.8784, 0.8701\} \quad (31)$$

となる。統計数値表のほか、ほとんどの初等統計学のテキストには付表に正規分布の上側確率

$$Q(u) = 1 - \Phi(u)$$

が掲載されているので、 U_j は容易に求められる。

(29), (30), (31)を用いて Ψ_p^2 ($p=1, \dots, 4$)を求めると、

$$\begin{aligned} \Psi_1^2 &= 4.916, \\ \Psi_2^2 &= 5.192, \\ \Psi_3^2 &= 6.850, \\ \Psi_4^2 &= 9.367 \end{aligned} \quad (32)$$

となる。

一方、 Ψ_p^2 ($p=1, \dots, 4$)の有意点は表5に示される。

したがって危険率1%で H_0 を検定した場合、 Ψ_1^2 , Ψ_2^2 , Ψ_3^2 , Ψ_4^2 のいずれを用いても H_0 は棄却される。

つまり点列 z_j は正規分布に適合しないことがわかる。

表5. Neymanのスムーズ・テストの有意点

危険率 検定統計量	1%	5%	10%
Ψ_1^2	0.299	0.173	0.122
Ψ_2^2	1.644	0.974	0.702
Ψ_3^2	6.928	4.142	3.014
Ψ_4^2	9.332	6.113	4.727

4.2 Shapiro-Wilkの検定³⁾

点列 z_j を昇順に並べた

$$z_{(1)} < \dots < z_{(k)}$$

は、 H_0 のもとで標準正規分布 $N(0, 1^2)$ からの標本の順序統計量となる。

Shapiro-Wilkは標本の標準偏差 σ の最良線形不偏推定量

$$\hat{\sigma} = \frac{1}{k} \sum_{j=1}^k a_j \cdot z_{(j)} \quad (33)$$

と

$$S^2 = \sum_{j=1}^k (z_{(j)} - \bar{z})^2 = \sum_{j=1}^k z_{(j)}^2 - \left(\sum_{j=1}^k z_{(j)} \right)^2 / k \quad (34)$$

を用いて、

$$W = (c \cdot \hat{\sigma})^2 / S^2 \quad (35)$$

の形の検定統計量を提案した。ただし定数 c は

$$\sum_{j=1}^k (c \cdot a_j)^2 = 1 \quad (36)$$

より定められる。ここで

$$b_j = c \cdot a_j \quad (j=1, \dots, k) \quad (37)$$

とおくと

$$b_j + b_{k-j+1} = 0 \quad (j=1, \dots, k) \quad (38)$$

の関係を用いて W は

$$W = \left[\frac{k}{2} \sum_{j=1}^k b_{k-j+1} (z_{(k-j+1)} - z_{(j)}) \right]^2 / S^2 \quad (39)$$

と表わすことができる。

彼らは種々の非正規分布の下での W の期待値をモンテカルロ実験で求め、それが正規分布の下での W の期待値 w_0 より小さくなる傾向があることから

$$W < w_0 \quad (40)$$

のとき H_0 を棄却することを示した。

表4に示す点列 z_j ($j=1, \dots, 10$)についてShapiro-Wilkの検定量 W を求めてみよう。点列 z_j を昇順に並べた点列 $z_{(j)}$ は

$$\begin{aligned} \{z_{(j)}; j=1, \dots, 10\} &= \{-1.233, 0.182, 0.204, 0.261, 0.848 \\ &1.038, 1.127, 1.159, 1.167, 1.184\} \end{aligned} \quad (41)$$

となる。 W を求めるための重みは $k=10$ のとき

$$\{b_j; j=1, \dots, 5\} = \{0.5739, 0.3291, 0.2141, 0.1224, 0.0399\}$$

である。したがって

$$\begin{aligned} W &= \left(\sum_{j=1}^5 -b_j (z_{(11-j)} - z_{(j)}) \right)^2 / S^2 \\ &= 0.7752 \end{aligned} \quad (42)$$

を得る。

$k=10$ の場合のShapiro-Wilkの検定の有意点は、表6に示される。表6より下側1%点は0.781であるから、 H_0 は棄却される。つまりNeymanのスムーズ・テスト同様、点列 z_j は正規分布に適合していないことがわかる。

なお3)には、Shapiro-Wilkの検定統計量を計算するための重み b_i の表と検定の有意点が、 $k=3(1)50$ について与えられている。

なおこの表4のデータを用いてKolmogorov-Smirnov検定を行うと、危険率10%でも H_0 は棄却されない。

表6. Shapiro-Wilkの検定の有意点

α	.99	.95	.90
$k=10$.781	.842	.869

ただし α は z_1, \dots, z_k が正規分布からの標本である場合のWの分布の上側100 α %点を与える。

5. バグ推定値の信頼限界の算出

Neymanのスムーズ・テストあるいはShapiro-Wilkの検定の結果、正規性の仮説が棄却されないならば、任意の観測時刻におけるバグ推定値の信頼限界を求めることができる。

$$y = \ln \ln [1 / (1 - F_0(x))] \\ = \ln \ln [1 / (1 - h(t) / N)] \quad \text{再掲(9)}$$

の上限/下限は y の分布が正規分布に近似できるから

$$y_U = \mu_{nd} + K_\alpha \sigma_{nd}, \quad (43)$$

$$y_L = \mu_{nd} - K_\alpha \sigma_{nd} \quad (44)$$

と書ける。ただし μ_{nd} および σ_{nd} は(17)、(18)で示される標準化2重指数分布の順序統計量の平均、および分散である。また K_α は

$$\alpha/2 = \int_{K_\alpha}^{\infty} \phi(x) dx \quad (45)$$

を満たす標準正規分布のパーセント点であり、 α は危険率と呼ばれる。ここで $\phi(x)$ は標準正規分布の確率密度関数である。片側の信頼限界を求めるときには、(45)の $\alpha/2$ を α に置き換えれば良い。

したがってバグ推定値の上限/下限は(9)、(43)、(44)より

$$\hat{N}_U = \hat{h}(t) / \{1 - \exp[\exp[y_U]]\}, \quad (46)$$

$$\hat{N}_L = \hat{h}(t) / \{1 - \exp[\exp[y_L]]\} \quad (47)$$

(UとLの添字が逆になることに注意)となるから、発見されない潜在的なバグも含めて残存バグ数の上限/下限は

$$\hat{N}_U - \hat{h}(t) = \hat{h}(t) \cdot \{1 / (\exp[\exp[y_U]] - 1)\}, \quad (48)$$

$$\hat{N}_L - \hat{h}(t) = \hat{h}(t) \cdot \{1 / (\exp[\exp[y_L]] - 1)\} \quad (49)$$

となる。

ここで \hat{N} は

$$\hat{N} = \lim_{t \rightarrow \infty} \hat{h}(t)$$

で示されるが、顕在する全てのバグが発見されたときの t は有限であることに留意すべきである。つまり有限な時刻 t において、 $t \rightarrow \infty$ としたときの総バグ数 N を推定している。

さてここで、発見されない潜在的なバグがどの程度あるかを考えてみよう。

(48)(49)の右辺の $\{\}$ の中身は、危険率 α を定めると、総バグ数 n とバグの発見順位 i にのみ依存する。一方 $h(t)$ は発見バグ数の平均値関数であることから、信頼度成長曲線の母数と時刻 t に依存する。ところが信頼度成長曲線の母数の最尤推定値は漸近的に不偏な推定量であるから、全てのバグが発見されたとき、すなわち $i=n$ のときには

$$E[h(t)] = n \quad (50)$$

と考えられる。つまり全てのバグが発見されたとき

の t は十分に大きいという仮定ができる。

以上の議論より、全てのバグが発見されたときの残存バグの全体に対する上限/下限は(48)(49)において $i=n$ とおくことにより得られる。これを危険率 $\alpha=0.01, 0.05, 0.10$

のとき

$$n=i=10(10)40$$

について示したのが表7である。表中()は(50)の仮定を用いて実際にいくつのバグが残存するかを示している。

これより危険率の多寡にかかわらず、発見総バグ数 n が増えるにつれて、発見されない残存バグの全体に対する比率の上限は減少することがわかる。その減少のしかたは発見総バグ数にほぼ反比例するため、残存バグの実際の数は $n=10(10)40$ のいずれもほぼ一定なのは興味深い。

もっとも n を40以上にした場合には、残存バグの全体に対する比率がどのように変化するかはまったくわからないことを断っておく。

表7. 残存バグの全体に対する比率 (上段:上限, 下段:下限)

発見バグ数 n	危険率 α		
	0.01	0.05	0.10
10	0.66 (6.6)	0.43 (4.3)	0.34 (3.4)
	0.00 (0.0)	0.00 (0.0)	0.00 (0.0)
20	0.31 (6.3)	0.21 (4.2)	0.16 (3.3)
	0.00 (0.0)	0.00 (0.0)	0.00 (0.1)
30	0.21 (6.3)	0.14 (4.2)	0.11 (3.3)
	0.00 (0.0)	0.00 (0.0)	0.00 (0.1)
40	0.16 (6.4)	0.10 (4.2)	0.08 (3.3)
	0.00 (0.0)	0.00 (0.0)	0.00 (0.1)

6. まとめ

種々提案されてきたソフトウェア信頼度成長曲線の適合度を、順位による誤差分散を考慮して検定する方法について述べた。本検定法は信頼度成長曲線の適合度検定問題をNeymanのスムーズ・テストやShapiro-Wilkの検定に帰着させている。標本をそのまま用いるため、Kolmogorov-Smirnov検定や χ^2 検定のような情報量損失がなく一般に高い検出力を示すことがわかっている。またバグ推定値の信頼限界も容易に求めることができる。

ただShapiro-Wilkの検定では、検定する正規分布の平均と分散は未知であっても構わないので、より適切な検定統計量が存在するかもしれないことを付記する。

参考文献

- 1) Kenett, R. and Pollak, M.: A Semi-Parametric Approach to Testing for Reliability Growth, with Application to Software Systems, IEEE Trans. Reliability, vol R-35, No.3, pp304-311 (1986).
- 2) 大場充: ソフトウェア信頼度成長曲線の適合度検定, 第6回ソフトウェア生産における品質管理シンポジウム発表報文集, 日本科学技術連盟, pp157-162 (1986).
- 3) 柴田義貞: 正規分布, p307, 東京大学出版会(1981)
- 4) Gumbel, E.J. 河田(訳): Statistics of Extremes

極値統計学, 生産技術センター新社(1978).

5) Yamada, S., Ohaba, M. and Osaki, S.: S-Shaped Reliability Growth Modeling for Software Error-Detection, IEEE Trans. Reliability, vol R-32, No.4, pp475-478, 484(1986).

6) Yamada, S., Ohaba, M. and Osaki, S.: S-Shaped Software Reliability Growth Models and their Applications, IEEE Trans. Reliability, vol R-33, No.3, pp289-292(1984).

7) Goel, A.L. and Okumoto, K.: Time-dependent Error-Detection Rate Models for Software Reliability and Other Performance Measures, vol R-28, No.2, pp206-211(1979).

付録

A. $g_{nd}(x)$ の1-4次モーメントの誘導

一般に、分布関数 F に従う n 個中 i 番目の大きさの順序統計量の確率密度関数は、

$$g_{nd}(x) = n \binom{n-1}{i-1} f(x) (F(x))^{i-1} (1-F(x))^{n-i} \quad (A.1)$$

となる。この分布の原点回りの k 次モーメント $\mu_{nd}(k)$ は

$$\mu_{nd}(k) = \int_0^{\infty} x^k \cdot g_{nd}(x)$$

で示されるから、 $F(x)$ に標準化2重指数分布を仮定すると

$$\begin{aligned} \mu_{nd}(k) &= \int_0^{\infty} x^k \cdot n \binom{n-1}{i-1} \exp[x] \cdot \exp[-\exp[x]] \cdot \\ &\quad \{1 - \exp[-\exp[x]]\}^{i-1} \cdot \{\exp[-\exp[x]]\}^{n-i} dx \\ &= n \binom{n-1}{i-1} \int_0^{\infty} x^k \cdot \exp[x] \cdot \\ &\quad \{1 - \exp[-\exp[x]]\}^{i-1} \cdot \{\exp[-\exp[x]]\}^{n-i+1} dx \\ &= n \binom{n-1}{i-1} \sum_{j=0}^{i-1} (-1)^j \binom{i-1}{j} \int_0^{\infty} x^k \exp[x] \cdot \\ &\quad \{\exp[-\exp[x]]\}^{n-i+j+1} dx \quad (A.2) \end{aligned}$$

となる。ここで

$$\exp[x] = y$$

とおくと、

$$x = \ln y,$$

$$\exp[x] dx = dy$$

であるから、

$$\mu_{nd}(k) = n \binom{n-1}{i-1} \sum_{j=0}^{i-1} (-1)^j \binom{i-1}{j} \int_0^{\infty} (\ln y)^k \cdot \exp[-y] \cdot y^{n-i+j+1} dy \quad (A.3)$$

となる。 j を $(j-1)$ と置き換えることにより、

$$\mu_{nd}(k) = n \binom{n-1}{i-1} \sum_{j=1}^i (-1)^{j-1} \binom{i-1}{j-1} \int_0^{\infty} (\ln y)^k \cdot \exp[-y] \cdot y^{n-i+j} dy \quad (A.4)$$

を得る。ここで

$$y^{(n-i+j)} = z$$

とおくと

$$\mu_{nd}(k) = n \binom{n-1}{i-1} \sum_{j=1}^i (-1)^{j-1} \binom{i-1}{j-1} \frac{1}{n-i+j}$$

$$\int_0^{\infty} (\ln \frac{z}{n-i+j})^k \cdot \exp[-z] dz \quad (A.5)$$

となる。さて

$$\begin{aligned} \int_0^{\infty} (\ln z)^k \exp[-z] dz &= \Gamma'''(1) \\ &= \gamma^4 + \gamma^2 \pi^2 + 4\gamma \ln''(1) + 3\pi^4/20, \quad (A.6) \end{aligned}$$

$$\int_0^{\infty} (\ln z)^3 \exp[-z] dz = \Gamma'''(1)$$

$$= -(\gamma^3 + \gamma \pi^2/2 + \ln'''(1)), \quad (A.7)$$

$$\int_0^{\infty} (\ln z)^2 \exp[-z] dz = \Gamma''(1) = \gamma^2 + \pi^2/6 \quad (A.8)$$

$$\int_0^{\infty} (\ln z) \exp[-z] dz = \Gamma'(1) = -\gamma, \quad (A.9)$$

$$\int_0^{\infty} \exp[-z] dz = 1 \quad (A.10)$$

であるから、 $g_{nd}(x)$ の1-4次モーメントは以下で示される。

$$\begin{aligned} \mu_{nd(1)} &= n \binom{n-1}{i-1} \sum_{j=1}^i (-1)^{j-1} \frac{1}{n-i+j} \binom{i-1}{j-1} [\Gamma'(1) \\ &\quad - \ln(n-i+j)], \quad (A.11) \end{aligned}$$

$$\begin{aligned} \mu_{nd(2)} &= n \binom{n-1}{i-1} \sum_{j=1}^i (-1)^{j-1} \frac{1}{n-i+j} \binom{i-1}{j-1} [\Gamma''(1) \\ &\quad - 2\Gamma'(1) \cdot \ln(n-i+j) + \{\ln(n-i+j)\}^2], \quad (A.12) \end{aligned}$$

$$\begin{aligned} \mu_{nd(3)} &= n \binom{n-1}{i-1} \sum_{j=1}^i (-1)^{j-1} \frac{1}{n-i+j} \binom{i-1}{j-1} [\Gamma'''(1) \\ &\quad - 3\Gamma''(1) \cdot \ln(n-i+j) + 3\Gamma'(1) \cdot \{\ln(n-i+j)\}^2 \\ &\quad - \{\ln(n-i+j)\}^3], \quad (A.13) \end{aligned}$$

$$\begin{aligned} \mu_{nd(4)} &= n \binom{n-1}{i-1} \sum_{j=1}^i (-1)^{j-1} \frac{1}{n-i+j} \binom{i-1}{j-1} [\Gamma''''(1) \\ &\quad + 6\Gamma''(1) \cdot \{\ln(n-i+j)\}^2 + \{\ln(n-i+j)\}^4 \\ &\quad - 4\Gamma'''(1) \cdot \ln(n-i+j) - 4\Gamma'(1) \cdot \{\ln(n-i+j)\}^3]. \quad (A.14) \end{aligned}$$