ソフトウェア工学 140－18
（２００３．３．７）

統合モデルを用いた推定残存欠陥数の信頼限界について

古山 恒夫
東海大学　開発工学部 1)

概要：ソフトウェア信頼度成長モデルのひとつである統合モデルを用いて残存欠陥数を推定する際に、その信頼限界も合わせて計算する方法を提案する。この方法は、与えられた累積欠陥データを生み出す可能性のあるモデルインスタンスすべての確率と、モデルインスタンスごとに求められる推定残存欠陥数から得られる各推定欠陥数の発生確率分布を組み合わせる。この方法を用いて信頼限界を計算した結果、ポアッソン分布から推定した従来の方法より信頼限界が広くなること、信頼限界に最も大きな影響を与えるのは総欠陥数に対する欠陥検出比率であること、推定曲線のＳ字型が強くなるに従って信頼限界は狭くなること、累積欠陥データの凹凸は信頼限界に全く影響を与えないことが明らかとなった。

Confidence Limits of the Number of Remaining Faults Estimated Using a Manifold Growth Model

Tsuneo Furuyama
Tokai University, School of High Technology for Human Welfare

This paper proposes a new method that gives confidence limits of the number of remaining faults. In the new method, probabilities for all instances of a software reliability growth model that may cause the given accumulated fault data, and probability distributions of the estimated number of remaining faults from the all instances are combined to calculate the final probability distribution of the estimated number of remaining faults. The simulation results show that the reliability range given by the new method is wider than that by the existing method using the Poisson distribution, and the most effective parameter is the fault detection ratio that is the ratio of the number of detected faults so far and the total number of faults.

１．はじめに

ソフトウェア信頼度成長モデル（SRGM）を用いて残存欠陥数を推定する方法は、テスト工程において製品の信頼度を評価する有力な方法である。この方法では、例えば与えられた累積欠陥データから最尤推定法により総欠陥数を推定し、その推定値と現時点までに検出した欠陥数の差から残存欠陥数を推定する。しかし、これまでは多くの場合、推定残存欠陥数が関心の中心で、その推定値がどの程度信頼できるか、言い換えると推定値の信頼限界に対しては議論の対象となることは少なかった。実際の累積欠陥データは、従来の典型的なモデルに近いものからどのモデルにも似ていないものまでさまざまなものがある。従って、同じ最尤推定値が得られてもその有意な信頼限界は累積欠陥データで異なると思われる。

SRGM に関連するソフトウェア信頼度については、これまで２つの概念が提唱されている。ひとつは、テストが時刻 t まで進行しているときに、時間区間（t，t+x）においてソフトウェア故障の発生しない確率 $R(x|t)$であり、ソフトウェア信頼度と呼ばれている[1][2]。しかし、この方法は残存欠陥数の信頼限界を求めるときには、直接には役に立たない。

もうひとつは SRGM における推定総欠陥数の信頼限界を示すものであり、SRGM のうち非定常ポアッソン過程モデル（NHPP モデル）に属するモデルを用いた場合の総欠陥数あるいは残存欠陥数の推定値からその信頼限界を求める

1) 〒410-0395 沼津市西野 317

ものである[2]。この方法は、NHPP モデルの平均値関数ｙ＝H(t)の信頼限界を、各時刻において H(t) がポアッソン分布をなすという仮定のもとで、次の計算式で計算している。

$$H(t) \pm K_r \sqrt{H(t)} \quad (1)$$

ここで Kr は標準正規分布の 100(1+α)／2 パーセント点であって、α＝0.95 の場合は Kr＝1.96 である。

しかし、この方法は次のような理由から必ずしも現実感覚にそぐわない。

(1) 累積欠陥のグラフがｙ＝H(t)にぴったりの場合も、得られた平均値関数ｙ＝H(t)に対して上下に大きく振れている場合も全く同じ信頼限界をもつことになる。
(2) ｔが大きくなるに従って信頼限界が大きくなり、ほぼ収束しているようなデータに対しても大きな信頼限界を持つことになる。
(3) 推定総欠陥数Ｎは、累積欠陥数の最新値（以下単に最新値と呼ぶ）y_n より小さな値にはなりえないにもかかわらず、N の信頼限界の下方が y_n より小さな値として計算されることがある。

この考え方を現実に即して累積欠陥データに適用するには、次のような方法が考えられる。すなわち、推定残存欠陥数にだけポアッソン分布の考え方を適用し、時刻ｔおける推定残存欠陥数の信頼限界を、

$$(N - H(t)) \pm K_r \sqrt{N - H(t)} \quad (2)$$

で表すものである。Kr は式(1)におけるものと同じものである。

しかし、この方法も次のような疑問には答えていない。問題点を例で示そう。「今、袋に赤色と白色の玉が全部で 1,000 個入っているものとする。赤色の玉と白色の玉の比率は不明である。最初に袋から 10 個の玉を取り出したところ、赤色 3 個で白色が 7 個であった。その玉をすべて袋に戻してまた 10 個の玉を取り出したとき、赤色の玉が何個取り出される可能性が最も高いか」という問題を考える。これは単純に考えて３個と答えて誤りではない。しかし、問題が、「赤色の玉がｎ個取り出される確率分布を計算せよ」という場合は話しは単純でなくなる。単純に赤色の出現確率が 0.3 であるとして 2 項分布で計算するのは誤りである。袋の中の赤色の玉の比率が 0.3 であることは最も確からしいことではあるが、それ確実に保証されたものではないからである。ここで考えるべきことは、まずこれまでの（この例では１回）の試行から、袋の中の赤色の玉の比率を推定し、それぞれの推定確率に従って、２項分布を用いて赤色の玉がｎ個取り出される確率分布を計算して総和を求めるべきである。

今回提案する方法は、これと同じ考え方に基づいたものである。

２．新しい信頼限界の推定方式

２．１　基本的考え方

(1) 最初に、ある時刻までの累積欠陥データから、推定モデルのインスタンスごとに、与えられた累積欠陥データが得られる確率を計算する。モデルは一般にいくつかのパラメータに支配されているので、具体的にはそれぞれのパラメータの値の組ごとに、累積欠陥データを生み出す確率を計算する。
(2) 次に、各モデルのインスタンスに対して、残存欠陥数の最尤推定値を求める。
(3) 最尤推定値を平均値とするポアッソン分布から、インスタンスごとの残存欠陥数の確率分布を求める。
(4) インスタンスごとの残存欠陥数の確率分布に、各インスタンスが得られる確率を掛けて、残存欠陥数ごとに和をとる。これによって推定残存欠陥数ごとの確率分布を求めることができる。

ここで注意しなければならないことは、計算対象となるモデルのインスタンスとして、H（∞）の値が現時点までの累積欠陥数よりも小さいものが含まれていてもよいことである。これは、平均値関数 H(t)は検出欠陥数 N(t)そのものではなく、あくまでも N(t)の期待値にすぎないという事実に基づいている。したがって H(∞)＜N(∞)あるいはその逆であってもかまわない。

２．２　式による表現

（１）総欠陥数の推定

時刻ｔが $t_0, t_1, \ldots, t_n$ のときの実際の累積欠陥数 y を $y_0, y_1, \ldots, y_n$、推定モデル $y=M(t_i; a)$（a はモデルパラメータ）から得られる累積欠陥数の推定値 $\tilde{y}$ を $\tilde{y}_0=M(t_0;a)$, $\tilde{y}_1=M(t_1;a), \ldots,$

$\tilde{y}_n$ =M(tₙ;a) とする。 y_0, y_1,…, y_n 及び $\tilde{y}_0$=M(t_0;a), $\tilde{y}_1$=M(t_1;a),…, $\tilde{y}_n$=M(t_n;a)に関してある評価関数の値を最小とするような最適なモデルパラメータを推定することにより最適な推定モデルのインスタンスを定める。そのモデルインスタンスから最も確からしい総欠陥数 $\tilde{y}_\infty$ = M(t_∞;a)を推定する。

一般に y_0 = 0 となるモデル、例えば指数型モデルや遅延 S 字型モデルでは非定常ポアッソン過程の仮定が成立するため、次の同時確率密度関数を尤度関数 L として、L を最大とするパラメータを求めることにより、モデルを決定し、総欠陥数を推定している。

$$L = P\{M(t_1)=y_1, M(t_2)=y_2, ..., M(t_n)=y_n\}$$
$$= \prod_{i=1}^{n} \frac{\{M(t_i;a)-M(t_{i-1};a)\}^{y_i-y_{i-1}}}{(y_i-y_{i-1})!} \quad (3)$$
$$\exp\{-M(t_n;a)\}$$

これは、ある累積欠陥データが与えられたときに、実測値（累積欠陥数の列）が起きる確率を最大にするモデルパラメータを求めよ、という問題に相当する。これは次のように表すことができる。

「$P(M(t_i;N,\beta,\gamma)\,|\,y(t_i)) \to \max$ となる N, β, γを求める」

（2）推定総欠陥数の信頼限界の推定

累積欠陥数の最新値を y_n、推定モデルのインスタンスを $y = M(t_i;\tilde{N},\tilde{\beta},\tilde{\gamma})$ とすると、そのモデルインスタンスでの推定残存欠陥数 $y_+ (= y_\infty - y_n)$ の確率分布は、

$$P(y_+) = \frac{(\tilde{N} - M(t_n;\tilde{N},\tilde{\beta},\tilde{\gamma}))^{y_+}}{y_+!} \quad (4)$$
$$\times \exp\{-(\tilde{N} - M(t_n;\tilde{N},\tilde{\beta},\tilde{\gamma}))\}$$

となる。つまり、推定総欠陥数 N の確率分布は、y_n より小さい値ではゼロ、y_n 以上では N－y_n の値に対して平均値 $\tilde{N} - M(t_n;\tilde{N},\tilde{\beta},\tilde{\gamma})$ のポアッソン分布に従うものとなる。その期待値は

$$N = \tilde{N} - M(t_n;\tilde{N},\tilde{\beta},\tilde{\gamma}) + y_n \quad (5)$$

となる。推定モデルでの最新値 $M(t_n;\tilde{N},\tilde{\beta},\tilde{\gamma})$ が実際の最新の累積欠陥数 y_n と等しい場合は、N は推定モデルのインスタンスでの推定総欠陥数 $\tilde{N}$ に等しい。

個々のモデルインスタンスに対する N の信頼限界は、一般に

$$N \pm K_r \sqrt{\tilde{N} - M(t_n;\tilde{N},\tilde{\beta},\tilde{\gamma})} \quad (6)$$

となるが、実際の信頼限界を求めるためには、推定モデルのインスタンスごとの発生確率を計算し、これと式(4)で示した推定モデルのインスタンスごとの y_+ の分布とを掛け合わせて全体としての y_+ の分布を求め、その分布から N の信頼限界を求めなければならない。

２．３　実際の推定式

以下では、推定モデルとして統合モデルを用いる。統合モデルは既存の代表的な SRGM をカバーすることができるため、さまざまな累積欠陥データに対して適合することができる[3][4]。統合モデルの一般解のうち y_0 = 0 となるものは、次の式で表される。

$$y = N(1-e^{-\beta t_i})^{\frac{1}{\gamma}} \quad (7)$$

ただし、N は総欠陥数、βは欠陥の検出速度を表すパラメータ、γは欠陥検出の過程を表すグラフの形状（モデルインスタンスの形状）を規定するパラメータである。

（1）仮定

実際の推定式を導出するにあたって次の仮定を設ける。

(A) 総欠陥数 N、欠陥の検出速度を表すパラメータβ、モデルの曲線の形状を規定するパラメータγはそれぞれ独立に決定される。

(B) 最尤推定から大きくはずれるパラメータ N、β及びγに対しては、そのパラメータに対するモデルから実測値が実際に発生する確率をゼロとみなす。

(C) パラメータ N、β及びγの発生確率は、最尤推定に近い範囲ではそれぞれ一様分布をする。

式(7)を次のように変形する。

$$\frac{y}{N} = (1-e^{-\beta t_i})^{\frac{1}{\gamma}} \quad (8)$$

この式から、N はグラフの縦軸である累積欠陥数を正規化するパラメータであり、βはグラフの横軸である時間を正規化するパラメータであることがわかる。これらのことから、上記(A)の仮定には無理のないことがわかる。

(B)の仮定は、あるモデルインスタンス

$M(t_i; N, \beta, \gamma)$ が与えられたときに実際の累積欠陥データ $y(t_i)$ が得られる確率 $P(y(t_i) \mid M(t_i; N, \beta, \gamma))$ はβ及びγがある範囲以外（$N_0 < N, \beta < \beta_0, \beta_1 < \beta$ 及び $\gamma < \gamma_0, \gamma_1 < \gamma$）ではゼロとみなせる、というものである。理論上、パラメータＮのとりうる値の範囲は[0, ∞]、パラメータβとγの値のとりうる範囲は（−∞, ∞）であるが、これでは数値積分を行うことができないので、(B)は計算の都合上設けた仮定である。ただし、この仮定は２．４で示すように実際の累積欠陥データの調査結果から無理な仮定でないことがわかる。

(C)の仮定は、次のように表すことができる。

$$N = \frac{1}{N_0} \tag{9}$$

$$\beta = \frac{1}{\beta_1 - \beta_0} \qquad (\beta_0 \le \beta \le \beta_1) \tag{10}$$

$$\gamma = \frac{1}{\gamma_1 - \gamma_0} \qquad (\gamma_0 \le \gamma \le \gamma_1) \tag{11}$$

ここでは範囲の境界を(B)の仮定で用いたと同じ値（$N_0, \beta_0, \beta_1, \gamma_0, \gamma_1$）を用いている。(C)の仮定は重要で、一様分布でなければ信頼限界を計算するにあたってその確率分布をさらに考慮しなければならない。

（２）信頼限界の推定式

実測値 $y(t_i)$ が与えられたときに、推定モデルのインスタンス $M(t_i; \tilde{N}, \tilde{\beta}, \tilde{\gamma})$ の推定総欠陥数 $\tilde{N}$ と推定モデルの最新値 $\tilde{y}_n$（＝$M(t_n; \tilde{N}, \tilde{\beta}, \tilde{\gamma})$）との差が y_+ となる確率 $P(y_+ \mid y(t_i))$ は次の式で与えられる。

$$P(y_+ \mid y(t_i)) = \iiint_{y_+ = N - \tilde{y}_n} P(M(t_i; N, \beta, \gamma) \mid y(t_i))\, dN d\beta d\gamma$$

$$= \frac{1}{P(y(t_i))} \iiint P(y(t_i) \mid M(t_i; N, \beta, \gamma)) \times P(M(t_i; N, \beta, \gamma))\, dN d\beta\, d\gamma$$

（∵ベイズの定理より）

$$= \frac{1}{P(y(t_i))} \iiint P(y(t_i) \mid M(t_i; N, \beta, \gamma)) \times P(N) P(\beta) P(\gamma)\, dN d\beta\, d\gamma$$

（∵仮定(A)より）

$$= \frac{1}{P(y(t_i))} \cdot \frac{1}{\mathrm{N}_0(\beta_1 - \beta_0)(\gamma_1 - \gamma_0)} \times \int_{\gamma_0}^{\gamma_1} \int_{\beta_0}^{\beta_1} \int_0^{N_0} P(y(t_i) \mid M(t_i; N, \beta, \gamma))\, dN d\beta\, d\gamma \tag{12}$$

（∵仮定(B)及び(C) より）

ここで、βの範囲について考えてみる。βは欠陥の検出速度（次元は時間の逆数 1/T）であるが、数式上は時間軸方向のスケールファクタである。実際の累積欠陥データでは、さまざまなプロジェクトがあり、1/βが数時間から数週間や数ヶ月にわたることがある。従って、積分範囲となる β_0 と β_1 の値は広い範囲となり、定めにくい。また、統合モデルでは、累積欠陥データが発散する場合も扱うことができる。このとき β はゼロまたは負の値をとる。そのため、時定数 T = 1/βという考え方を用いることはできない。すなわち、$\beta \to 0$ で $T \to \infty$ となり、T は一様分布でなくなる。

一方、$\beta' = \beta \cdot t_n$ を考えると、これは式(7)の指数関数のべき乗の値であり、推定曲線の飽和度を表す値とみなすことができる。したがって、その値のとりうる範囲も β そのものに比べ限られている。そこで

$\beta t_n = \beta'$（時定数で正規化した最新値の時間）

とおくことにより、式(12)を次のように変形する。

$$P(y_+ \mid y(t_i))$$

$$= \frac{1}{P(y(t_i))} \cdot \frac{1}{\mathrm{N}_0(\beta_1 - \beta_0)(\gamma_1 - \gamma_0)} \times \int_{\gamma_0}^{\gamma_1} \int_{\beta_0}^{\beta_1} \int_0^{N_0} P(y(t_i) \mid M(t_i; N, \beta, \gamma))\, dN d\beta\, d\gamma$$

$$= \frac{1}{P(y(t_i))} \cdot \frac{1}{\mathrm{N}_0(\beta'_1 - \beta'_0)(\gamma_1 - \gamma_0)} \times \int_{\gamma_0}^{\gamma_1} \int_{\beta'_0}^{\beta'_1} \int_0^{N_0} P(y(t_i) \mid M(t_i; N, \beta', \gamma))\, \mathrm{d}N d\beta'\, d\gamma \tag{13}$$

ただし、$\beta_0 t_n = \beta'_0 \le \beta\, t_n = \beta' \le \beta'_1 = \beta_1 t_n$ である。

２．４　パラメータの確率密度関数の分布

実際の 59 のプロジェクトデータをもとに、βとγの分布を調査した。一般に、βは「時定数 T」の逆数であり、実際には測定単位および組織ごとに異なるので、分布調査ではβの代わりにβ* t_n（＝t_n／T）を用いた。β* t_n を横軸に、γを縦軸にとった散布図を図１に示す。β

* t_n は[−5, 30]、γは[0, 4]の範囲内に収まっている。なお、βが負の場合は、累積欠陥データが発散することを意味している。

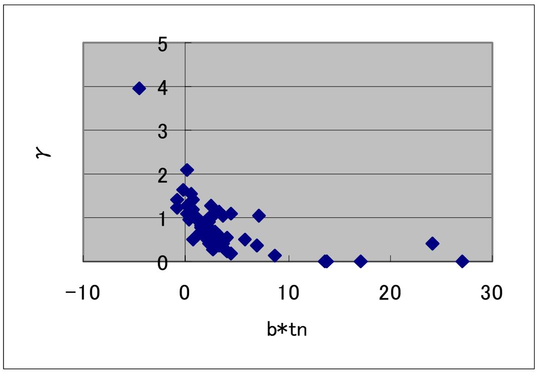

図1 ($\beta *t_n$, γ)の散布図

図１から極端な値を持つプロジェクトを除いたものについて、β* t_n の分布、γの分布を調べた結果を図２と図３に示す。

β* t_n の分布における特徴を以下に示す。

(1) 値が負のもの（収束しないプロジェクト）が３つある。

(2) β＝０のラインを「自然に」越えている。すなわち、収束するプロジェクトから発散するプロジェクトまで連続的に分布している。

(3) β* t_n の分布をみると、80%以上のプロジェクトは[−1, 5]の範囲内でほぼ一様分布をしている。

γの分布における特徴を以下に示す。

(4) γの大きなプロジェクト４つ（γ＞2 のものを加えると６つ)、γがゼロに近いプロジェクト４つ以外の 80%以上のプロジェクトが[0.1, 1.5]の範囲内でほぼ一様分布している。

以上のことから、一般には、β* t_n は[−1, 5]、γは[0, 1.5]の範囲より広い範囲、すなわち、β* t_n は[−2, 10]、γは[0, 3]の範囲を定義域として確率計算すればよいと思われる。

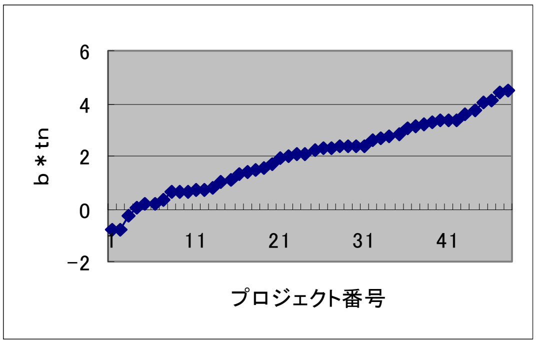

図２ β＊ t_n の分布(−1＜ $\beta *t_n$ ＜５の範囲のみ)

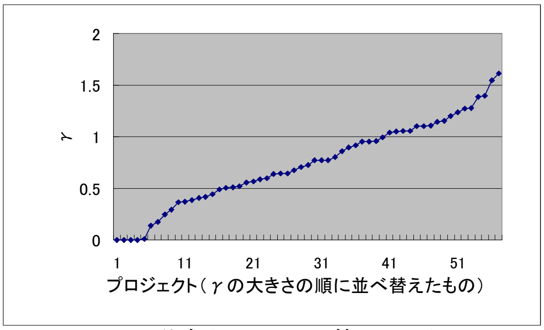

図３ γの分布(0＜γ＜２の範囲のみ)

３． 信頼限界の計算

３．１ 合成データによる信頼限界の計算

提案した方式による信頼限界の計算結果と推定残存欠陥数を平均値とするポアッソン分布を図４に示す。明らかに異なった分布をしていることがわかる。なお、信頼限界の推定計算のために用いた平均値関数は次の式で表されるものである。

$$y = 100 * (1 - e^{-1.5 t_i}) \tag{14}$$

ここで t_i=0, 1, 2,…,20 の値をとる。最新時刻 t_n=20 でｙは 95 となる。すなわち、欠陥検出比率が 0.95（全欠陥の 95%を検出）の時点で、提案した方式に従って残存欠陥数の確率分布を求めたものである。

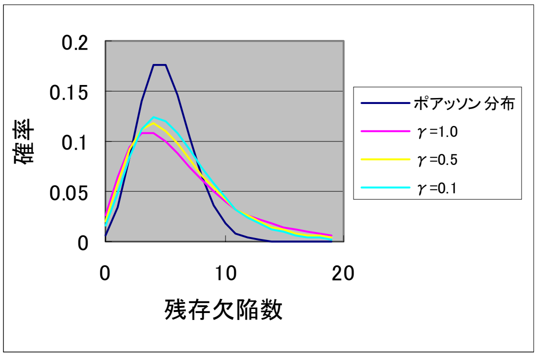

図４ 残存欠陥数の確率分布(N=100, yn=95)

３．２ 信頼限界に与えるパラメータの影響の分析

信頼限界が統合モデルのパラメータの変動によってどのように変化するかを合成データを対象として分析した。合成データの変動パラメータは、総欠陥数Ｎ、累積欠陥数の最新値 y_n（あるいは最新時刻での欠陥検出比率＝ y_n／N)、曲線の形状を表すγ、データ数ｎの４つである。

シミュレーションで用いた値は、総欠陥数N=100, 200、累積欠陥数の最新値 y_n=95（欠陥検出比率＝95%）及び 80（同 80%）、γ=1.0, 0.5, 0.1、n=20, 100 である。

（1） γの影響

累積欠陥比率を 0.95 とした場合の残存欠陥数の推定分布を図 4 に、同じく 0.8 とした場合の残存欠陥数の推定分布を図 5 に示す。検出比率によって明らかに分布が大きく異なることがわかる。同じ検出比率でもγによって微妙に異なり、γが小さいほど分布の形状は鋭くなり、ピーク値は大きな値となる。

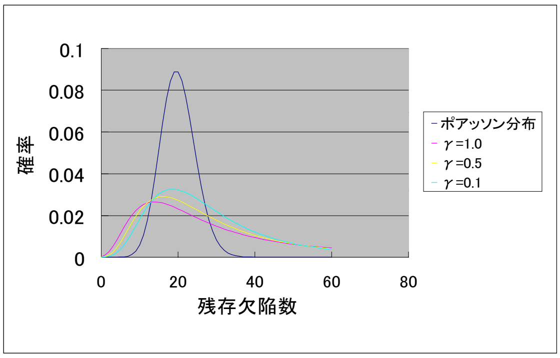

図 5 残存欠陥数の確率分布（N=100, yn=80）

（2）検出比率の影響

γを 1 とした場合の総欠陥数の確率分布を図 6 に示す。検出比率によって大きく分布は変わる。検出比率が高い場合は分布のピーク位置は実際の総欠陥数N（＝100）の値に近いが、検出比率が低くなるに従ってピーク位置は総欠陥数よりも小さくなる。しかし、それぞれの期待値は、検出比率が低くなるに従って逆にNよりも大きくなる（表 1）。これは検出比率が低くなるに従って推定総欠陥数の大きな値の確率分布が長い裾野をもつためである。

表 1 総欠陥数の期待値の変動

欠陥検出比率	総欠陥数の期待値
0.95	101.9
0.90	105.3
0.80	120.2
0.70	123.6

（3）総欠陥数（残存欠陥数）の違いによる影響

ポアッソン分布では、分散の値が平均値と等しいため、分散の平方根で表される標準偏差は平均値に比例しない。したがって、従来のポアッソン分布を用いた推定方式では、残存欠陥数が 2 倍になっても標準偏差が 1.4 倍にしかならず、信頼限界は残存欠陥数に比例しない。例えば、残存欠陥数の信頼限界は平均の期待値をμ、標準偏差をσとし、信頼限界をμ±2σとすると、μが 10 件の場合の信頼限界は 4 件から 16 件であるが、μが 40 件の場合は 40±12 件すなわち 28 件から 52 件となる。推定残存欠陥数に対する比率でみると、前者が±60%であるのに対し、後者は±30%となる。

提案した方式における残存欠陥数の大きさの影響を分析した結果を図 4 と図 7 に示す。図 7 は図 4 で得られた結果に対して、同一の検出比率及び曲線形状をもっていて総欠陥数Nだけが 2 倍（N＝200）の場合の推定残存欠陥数の確率分布を示したものである。図 4 と図 7 をもとに、N=100 の場合と 200 の場合の推定信頼限界の比較を表 2 に示す。明らかに、総欠陥数が増えると、推定信頼限界の総欠陥数に対する比率は下がる。

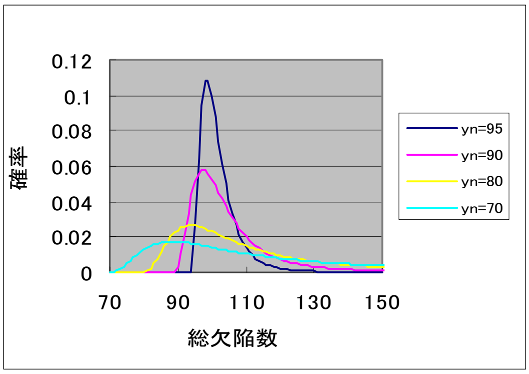

図 6 総欠陥数の確率分布（N=100, γ=1.0）

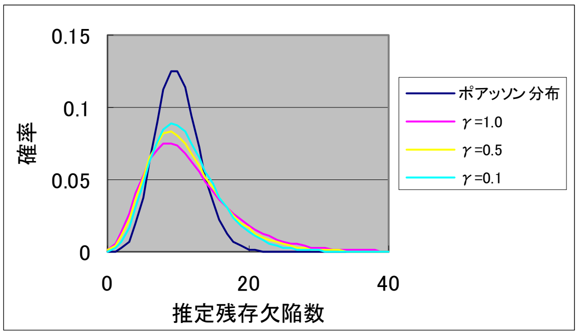

図 7 残存欠陥数の確率分布（N=200, yn=190）

表 2 残存欠陥数の信頼限界(95%)の総欠陥数による比較

分布形状	N = 100	N = 200
ポアッソン分布	2-9 (2%-9%)	5-16 (2%-8%)
γ=1.0	1-22 (1%-22%)	3-27 (2%-14%)
γ=0.5	1-17 (1%-17%)	3-24 (2%-12%)
γ=0.1	1-15 (1%-15%)	3-22 (2%-11%)

(注) ( )内は総欠陥数に対する比率

（4）データ数の影響

同一欠陥検出比率、同一曲線形状、同一推定残存欠陥数で、データ数nだけが異なる場合の推定残存欠陥数の確率分布を図 8 に示す。データ数は 20 の場合と 100 の場合を比較したものである。データ数nが増加すると確率分布のピークが大きな値にシフトするが、その差は非常に小さい。

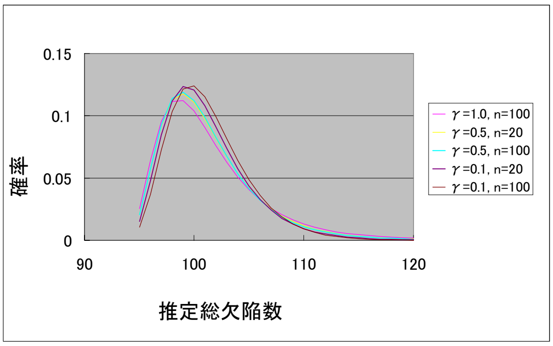

図 8 データ数の違いによる推定総欠陥数の確率分布(N=100, yn=95)

3．3 実データによる信頼限界の推定

実データとして、ある事務処理系のソフトウェアの開発途中の累積欠陥データを分析した。このデータの横軸は相対座標である。図 9 に累積欠陥データとその最尤推定曲線を示す。この例では総欠陥数の最尤推定値Nは 36.6、曲線の形状を表すパラメータγは 0.73 である。γの値からこの最尤推定曲線は、指数型モデルと遅延S字型モデルのちょうど中間の曲線となっていることがわかる。欠陥の検出比率は 31／36.6＝0.85（85%）である。

この実データに対して、本方式で計算した推定残存欠陥数の確率分布を図 10 に示す。95%の信頼限界は 1～119 であり、80%の信頼限界は 3～57 である予想以上に大きい。

図 10 では、実データだけでなく、実データの最尤推定モデルを検出欠陥の変動のない理想データとみなして、それを対象に推定残存欠陥数の確率分布を重ね合わせている。実データと理想データの間に確率分布の差は全くみられない。この結果から、累積欠陥データの累積欠陥のデータがｙ＝H(t) にぴったりの場合も、得られた平均値関数ｙ＝H(t) に対して上下に大きく振れている場合も全く同じ信頼限界をもつことになる。

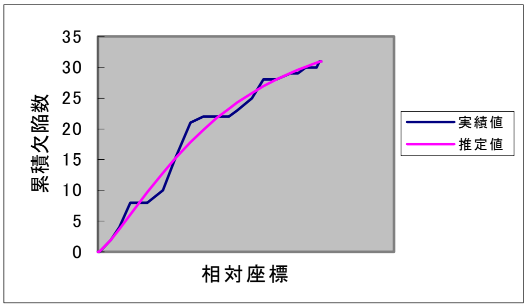

図 9 実績データとその推定値(N=36.6, γ=0.73)

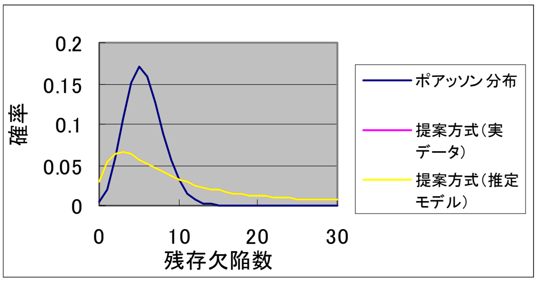

図 10 実データに対する残存欠陥数の確率分布

3．4 考察

シミュレーションの結果は、残存欠陥数の期待値が、最尤推定による残存欠陥数と異なるという意外なものであった。この原因としては次のことが考えられる。与えられた累積欠陥データから得られる推定モデルのパラメータは、最尤推定モデルパラメータを中心にそれより大きな値と小さな値に平等に広がっていくが、推定残存欠陥数が大きくなる方向に対しては無限大まで可能性が広がり、その値が小さくなる場合より強調されるためであると考えられる。最尤推定値としての残存欠陥数の確率分布を図 11 に示す。また、さまざまな欠陥検出率に対して、残存欠陥数の期待値と確率分布のピーク位置を表 3 に示す。一般に、ピーク位置は最尤推定残存欠陥数のよりも小さい値となるが、残存欠陥数の期待値は逆に最尤推定残存欠陥数より

も大きな値になる。

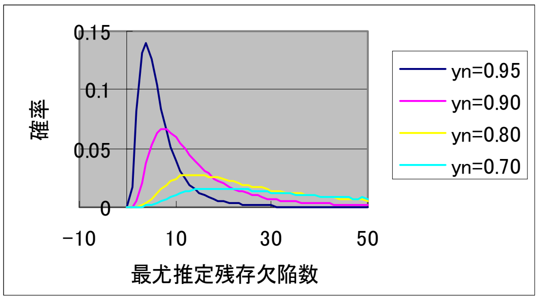

図 11　最尤推定残存欠陥数の確率分布
（N=100, γ=1.0）

表 3　最尤推定残存欠陥数の期待値と確率分布のピーク位置

yn	期待値	ピーク位置	最尤推定残存欠陥数
0.95	7.1	4	5
0.90	16.2	7	10
0.80	38.1	14	20
0.70	47.7	16	30

４．おわりに

信頼度のデータの利用法のひとつとして次のようなものがある。ソフトウェア信頼度成長モデルは、テストが理想的に行われ、欠陥の検出過程が完全に確率過程に従うときにその有効性が増す。しかし、一般にはテストは理想的に管理された状態で行われるとは限らない。そのようなとき現れるのは、例えばリリース後の検出欠陥数が推定残存欠陥数の信頼限界から大きくはずれると考えられる。このような場合は、テスト法に問題がなかったということを調べるきっかけとなりうる。

今回提案した､SRGM を用いて残存欠陥数を推定する際にその信頼限界を計算する新しい方法は、まず、与えられた累積欠陥データから、推定モデルのインスタンスがそのようなデータを生み出す確率を計算する。次に、各モデルのインスタンスごとに最尤推定に基づく残存欠陥数を推定する。最後に各モデルのインスタンスごとの最尤推定残存欠陥数を平均とするポアッソン分布を計算してその和から最終的な残存欠陥数を推定するものである。

シミュレーションを行った結果、与えられた累積欠陥データから最尤推定モデルのパラメータを求め、そのモデルが与える残存欠陥数を平均とするポアッソン分布から残存欠陥数の確率分布を求めるという従来の方法に比べると、信頼限界ははるかに大きなものであることがわかった。このことは、残存欠陥数はこれまで考えられているよりは、最尤推定値から大きく外れている可能性があるということを示すものである。信頼限界に最も大きな影響を与えるパラメータが欠陥検出比率であることは容易に理解できる。しかし、残存欠陥数の確率分布のピーク値は最尤推定値よりも小さくなるが、期待値は最尤推定値よりも大きくなることや、累積欠陥のグラフがｙ＝H(t)にぴったりの場合も、得られた平均値関数ｙ＝H(t)に対して上下に大きく振れている場合も同じ確率分布、したがって同じ信頼限界をもつことは、意外な結果であり、今後さらにその理由を分析する必要がある。

謝辞：本研究で用いた実データは、株式会社構造計画研究所から提供して頂いた。株式会社構造計画研究所の関係各位に感謝します。

[参考文献]

[1] 尾崎俊治, “ 非定常ポアッソン過程モデル ”, 情報処理, Vol. 31, No. 12, pp. 1631-1640 (1990).
[2] 山田　茂：ソフトウェア信頼性モデル――基礎と応用，日科技連，p. 195（1994）.
[3] 古山恒夫、中川 豊: ソフトウェア信頼度成長曲線に関する統合モデルと有効性の検証、情報処理学会ソフトウェア工学研究会、Vol. 97-10, pp. 73-80 (1994).
[4] Furuyama,T. and Nakagawa,Y.:A Manifold Growth Model that Unifies Software Reliability Growth Models, *Int. J. of Reliability, Quality and Safety Engineering*, Vol. 1, No. 2, pp.161-184 (1994).