

間欠故障データを含む場合の故障傾向解析 —EMアルゴリズムを基にしたテーブル化法による推定—

01503140 (財) 鉄道総合技術研究所 福岡 博 FUKUOKA Hiroshi

1. はじめに

一般に、現場で使用されている電子機器には、一時的に不安定になり故障現象が起きても原因不明のまま再使用されるというケースが少なくないと言われる。すなわち、発生した故障が図1のように原因が特定出来るもの、現象が再現せずに原因不明のまま一時的現象と見なされるもの、さらにその後に関欠的に同様の故障が起こるものとに分類できるというものである(例えば、仁川[1]など)。

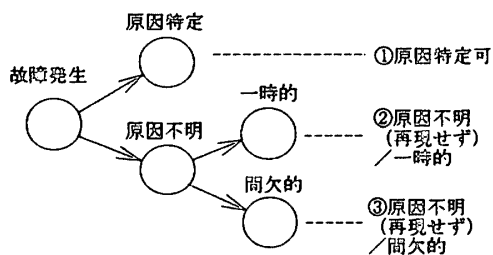


図1 現場で使用される電子機器の故障の分類

電子機器では、素子の劣化や変質等により環境条件(温度、湿度、振動、電磁環境など)の変化(ストレス)に対する受容性が(場合によっては一時的に)低下して動作が不安定となる場合があるが、前述の故障現象は通常の故障メカニズムにこのような故障モードがつけ加わったために起こるものと考えられる。すなわち、機器の状態変化が正常、故障の2種ではなく、正常、不安定、故障の3種になるというものである(図2)。



図2 不安定状態を含む故障形態

このように考えると、図1の故障分類において

① 原因特定可、

となる故障は、主として正常状態から不安定状態を通り過ぎて故障状態へ移行することにより起こる故障、

② 原因不明(再現せず) / 一時的

③ 原因不明(再現せず) / 間欠的

となる故障は、主として正常状態から不安定状態へ一度移行した後、不安定状態から故障状態へ移行することにより起こる故障と考えられる。他に、②、③の形態の特殊なものとして、見かけ上①の形態となるもの、あるいはその逆の場合も考えられるが、これは故障モード解析あるいは故障発生時の状況等により判断されることになる。

ここではこのような状況に対して、正常状態から不安定状態への移行を通常の故障発生と同様の現象とみなし、これを「不安定故障」と呼び、一方、不安定状態から故障状態への

移行を「間欠故障」と呼ぶことにする。

さて、このような故障発生形態に対しては、交代再生過程(不安定状態と正常状態の繰り返し)としてのモデル化が考えられる(たとえば、Malaiya[2]は並列冗長系でのユニット故障/システム故障を再生点にした)。しかし、通常の交代再生過程では不安定故障が認識可能であることが仮定されるが、ここで考えている不安定故障は容易には検出できない性質のものである。故障傾向解析の立場からは、これは故障データの一部分が観測できない不完全データの一つとしてみなせるものであり、通常の方法では解析できない。そこで、ここでは各試行における不安定故障発生時間の推定分布がベイズ統計における事後分布の形になることに着目して、寿命表を用いる方法(福岡[3]:EMアルゴリズム[4]をベースとしたもの。ここではこれをテーブル化法と呼ぶ)によって解析する。

2. 間欠故障発生モデル化

不安定故障の発生と間欠故障の発生との関係は図3に示すようになる。すなわち、不安定故障が起こった時点で不安定状態へ移行し、その後、不安定状態下における間欠故障の発生メカニズムによって故障が順次発生してゆくというものである。一般に、不安定故障は再生過程(再生故障)により、また間欠故障は修復過程により発生するもの(修復故障)とみなしてよい。

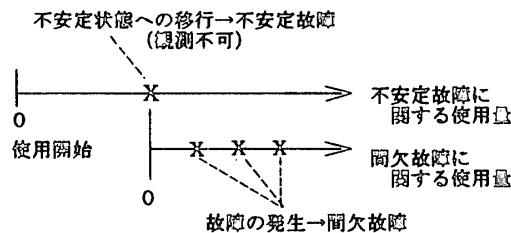


図3 不安定故障と間欠故障の関係

さて、不安定故障の発生までの使用量を T 、不安定状態に移行してから間欠故障が発生するまでの使用量をそれぞれ S_1, S_2, \dots とする (T と S_j は独立)。間欠故障発生時の使用開始からの使用量をそれぞれ U_1, U_2, \dots とすると、 $U_j = T + S_j$ となる。実際に観測されるのは、各サンプルごとの U_j の実現値であり、 T および S_j については情報は得られない。しかし、 U_j の分布は T および S_j の和の分布であり、たとえばこの分布形が扱いやすい形で表されていればこれを用いて T および S_j の分布形が推定可能である。しかし、一般的には T および S_j はともに特定の寿命分布に従うと考えられ、例えば適用例で述べる様に、両者に寿命分布としては一般的なワイブル分布を用いた場合には、これは困難である。そこで、ここではEMアルゴリズムをベースとしたテーブル化法によって数値的にこのデータを解析する。

3. テーブル化法

EMアルゴリズムは、不完全データの極めて一般的な解析法であるが、指数分布族のように十分統計量でデータの情報が取り扱えるとき以外は途中のステップでの関数の計算、最大化が非常に面倒になる。そのため、同様の繰り返し手法を用いる multiple imputation 法[5]などが新たに提案されている。しかし、multiple imputation 法は計算の各ステップでモンテカルロ法を実施するために、計算の手間が非常にかかると言われている。これに対してテーブル化法は、観測データを寿命表データで近似することによってEMアルゴリズムを一般的な問題に適用可能としたものであって、multiple imputation法と比較して計算量が少なく済むという特徴がある。

さて、テーブル化法の基本的なアイデアは、次の通りである。すなわち、テーブル化法ではEMアルゴリズムで言う「完全データ」を区間分けされたデータ（すなわち、寿命表）とする。そして、この（一般的な意味では）不完全なデータを、区間数ある程度大きく（正確には各区間の幅を小さく）とっておくことにより、近似的に非区間分けデータのように扱う。いま、 $0 = s_0 < s_1 < \dots < s_N$ によって区間分けが行われ、観測データ、非観測データがともに分類されてテーブル化（寿命表の構成）ができ、非区間データのような取扱い（近似）がされている状況を想定する。このとき、EMアルゴリズムのEステップで評価する関数 Q は、

$$Q = E(\log f(x|\phi)|y, \phi') \quad (3.1)$$

$$= E(\log \prod_{r=1}^N q(u_r|\phi)^{a_r+V_r} | y, \phi') \quad (3.2)$$

$$= \sum_{r=1}^N [(a_r + E(V_r|y, \phi')) \log q(u_r|\phi)] \quad (3.3)$$

ただし、

x, y : EMアルゴリズムでの完全データ、不完全データ

$f(\cdot)$: EMアルゴリズムでの完全データの標本密度

ϕ, ϕ' : EMアルゴリズムでの $p+1$ 、 $p+1$ でのパラメータ値

u_r : テーブル化したときの分類 r の代表値

a_r : 完全データの分類 r での故障数

V_r : 不完全データの分類 r での故障数の確率変数

$q(u_r|\phi)$: 分類 r でのサンプルの標本確率

と表すことができる。これによって関数 Q を計算すれば、Mステップでは完全データに対する最尤法と同様の手続きを用いればよいことになる。

4. 間欠故障データを含む場合のテーブル化法の適用

間欠故障データを含む故障傾向解析問題へのテーブル化法の適用における基本的なアイデアは次の通りである。 T の実現値は、見方を変えると U_j の分布に関する位置パラメータであると考えることができ、そのとき、 T はこの位置パラメータに関する事前分布を表す確率変数とみなすことができる。それゆえ、 S_1, S_2, \dots が観測されたもとのこの位置パラメータに関する事後分布が推定できる。これにより、各試行ごとに不安定故障発生に関する寿命分布が推定できるので、これをもとにしてEステップを実行する。すなわち、試行 i における不安定故障の発生時間に関する事後分布を $c_{i,post}(x)$ とすると、以下のように表せる：

$$c_{i,post}(x) = c_{pre}(x|\phi) \prod_{j=1}^{n_i} d_{i,j}(u_{i,j}|\phi) \quad (4.1)$$

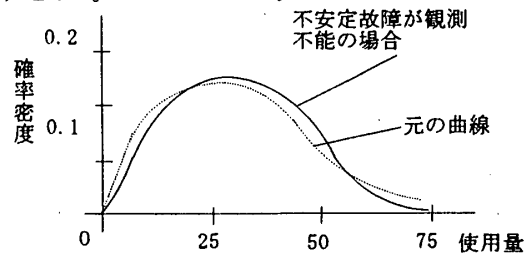
ここで、

$c_{pre}(x|\phi)$: 不安定故障の発生時間に関する事前分布

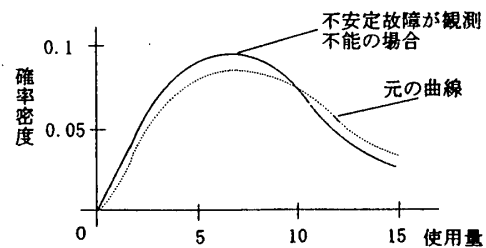
$d_{i,j}(u_{i,j}|\phi)$: 試行 i における j 番目の間欠故障データに関する標本確率

5. 適用例

図2はシミュレーションにより得られた間欠故障を含む場合のデータを上記の手法により解析した結果と元のデータ（不安定故障の発生時点が既知の場合）を解析した結果を比較したものである。シミュレーションでは間欠故障で観測機器を再生するか否かはランダムに決まる（再生する確率は0.3）とした。



(1) 不安定故障の確率密度関数



(2) 間欠故障の確率密度関数

図2. 間欠故障を含む場合と元のデータとの結果の比較

解析結果は、間欠故障を含む場合： $m'=2.43$ 、 $c'=38.5$ 、 $m=2.05$ 、 $C=9.99$ 、元のデータの場合： $m'=2.24$ 、 $c'=38.2$ 、 $m=2.00$ 、 $C=9.40$ （試行数：20）であり、比較的良く一致している（ m' 、 c' および m 、 C は、それぞれ不安定故障、間欠故障に対するワイブルの形状、尺度パラメータ値）。

6. おわりに

寿命表を使うことによる近似誤差については、区間数を十分大きくとることで、実用上問題ない程度には回避できる。収束については、適当な打ち切り基準を設ければ問題はない。今後は、推定量のばらつきの評価等を行う予定である。

[参考文献] [1] 仁川 清他：デバイス・部品の故障解析、日科技連、1992 [2] Malaiya, Y.K., IEEE Trans. Rel. Vol. 31, No. 2, 211-215, 1982 [3] 福岡 博：1989年度OR学会秋期研究発表会アブストラクト集、1989 [4] Dempster, D.B., et al., J. of Roy. Stat. Soc., ser. B, Vol. 39, 1-37, 1977 [5] Tanner, M. A. et al. : JASA, Vol. 82, No. 398, 528-540, 1987