対数ポアソン実行時間ソフトウェア信頼性モデルに関する一考察 — 余震発生モデルとの関連性を通して —

伊藤将貴, 岡村寛之 (01013754), 土肥正 (01307065)
広島大学大学院工学研究科情報工学専攻

1. はじめに

本稿では，ソフトウェア信頼性モデルの中で最も実用性が高いと言われている対数ポアソン実行時間モデル (logarithmic Poisson execution time model) [1, 2] に関する考察を行う．対数ポアソン実行時間モデルは非定常ポアソン過程 (NHPP) に基づいた確率モデルであり，その導出過程は強度関数に関する決定論的なふるまいをモデル化したものであった．

本稿では，まず最初に地震学において知られている，地震（本震）発生後の余震発生回数を記述する大森公式（例えば [3]）と対数ポアソン実行時間モデルが類似していることに着目し，対数ポアソン実行時間モデルに対する一般的な表現を与える．具体的には，現在の地震学において余震活動を記述する数理モデルとして最も良く知られている改良大森公式 [4] が，対数ポアソン実行時間モデルの拡張モデルとなっていることを示す．さらに，これらのソフトウェア信頼性モデルが一般化順序統計量モデルの枠組みで説明することが可能であることを示す．これにより，平均値関数が非有界な NHPP に従うソフトウェア信頼性モデルを，一般化順序統計量モデルの極限の意味で統一的に捉えることが出来る．

2. 改良大森公式 – 地震学からの知見

本震が発生した後に継続的に余震が発生するという現象は良く知られている．1894 年に Omori [5] は，本震発生後の経過時間と余震の発生回数の間に次のような関係式が成立することを発見した．

$$\lambda(t) = \frac{K}{t+c}. \tag{1}$$

ここで，$\lambda(t)$ は単位時間当たりの余震発生頻度であり，t は本震発生後の経過時間，$K\ (>0)$ と $c\ (>0)$ は任意定数である．上述の関係式は大森公式と呼ばれ，余震発生頻度が本震直後から減衰していく現象を記述している．さらに Utsu [4] は，1926 年以降の日本や世界の余震活動データを詳細に調べ，余震発生頻度の減衰が，大森公式よりはむしろ，

$$\lambda(t) = \frac{K}{(t+c)^p} \tag{2}$$

のような形状を示すことを指摘した．ここで，パラメータ $p\ (>0)$ は p 値と呼ばれ，時間減衰率を表す．式 (2) は改良大森公式と呼ばれており，余震発生頻度の時間的（かつ決定論的）挙動を記述する基本公式として世界中で広く認知されている．$p=1$ の時，改良大森公式は大森公式に一致することから，改良大森公式は大森公式の一般化表現であると言える．

余震活動を観測する時，式 (2) のパラメータ c は本震直後の複雑な変動状況を丸め込む役割を果たすが，通常その効果は数時間程度であるため，本震から数日以上経過すると t は c よりも十分大きくなり，近似的に $\lambda(t) \approx Kt^{-p}$ とみなすことが出来る．さらに両辺の対数をとると $\log\lambda(t) = \log K - p\log t$ となり，$-p$ は両対数プロット $(\log\lambda(t), \log t)$ の傾きを意味する．よって，単位時間当たりの余震発生回数のデータに線形回帰を適用することにより，改良大森公式（さらには当該地域の余震活動の指標）を特徴づける p 値を推定することが可能となる．このような決定論的なデータ解析法の欠点を克服するために，Ogata [6] は点過程に基づいた確率モデルの推定法について議論している．すなわち，余震発生回数が式 (2) の強度関数 $\lambda(t)$ を持つ NHPP に従うと仮定することにより，最尤法に基づいてパラメータ p, K, c を推定することを提案している．この確率モデルによる推定方法は，現在，余震活動予測の基本的なツールとして広く利用されている．

3. 対数ポアソン実行時間モデル

ソフトウェアのテスト工程においてフォールトが発見される（故障が発生する）現象をモデル化するために，単位テスト実行時間 (CPU 時間) 当たりに発見されるフォールト数を表す強度関数 $\lambda(t)$ に着目する．一般の点過程 $\{N(t), t \geq 0\}$ において，強度関数が時間のみに依存する関数である時，$N(t)$ は強度関数 $\lambda(t)$, 平均値関数

$$\Lambda(t) = \int_0^t \lambda(x)dx \tag{3}$$

を持つ NHPP に従うと言われる．Musa and Okumoto [1] は，ソフトウェアの単位テスト実行時間当たりに発見されるフォールト数が指数関数的に減少するものとし，

$$\lambda(t) = \lambda_0 \exp(-\theta\Lambda(t)) \tag{4}$$

を仮定した．ここで，パラメータ $\lambda_0\ (>0)$ は初期故障強度であり，$\theta\ (>0)$ はソフトウェアフォールト 1 個当たりの故障強度の減少率を表す．従って，微分方程式

$$\frac{d\Lambda(t)}{dt} = \lambda_0 \exp(-\theta\Lambda(t)) \tag{5}$$

を初期条件 $\Lambda(0)=0$ の下で解くことにより，

$$\lambda(t) = \frac{\lambda_0}{\lambda_0\theta t+1}, \tag{6}$$

$$\Lambda(t) = \frac{1}{\theta}\log(\lambda_0\theta t+1) \tag{7}$$

を得る．式 (7) で与えられる平均値関数 $\Lambda(t)$ を持つ NHPP

$$\Pr\{N(t)=n \mid N(0)=0\} = \frac{\Lambda(t)^n \exp(-\Lambda(t))}{n!} \tag{8}$$

は対数ポアソン実行時間モデルと呼ばれ，ソフトウェア信頼性モデルの中で最も実用的なモデルであることが知られている．

上述の対数ポアソン実行時間モデルと **2.** で紹介した大森公式の類似性は明らかである．すなわち，$1/c=\lambda_0\theta$ 並びに $K/c=\lambda_0$ と置くことにより，大森公式と対数ポアソン実行時間モデルの平均値関数は一致する．この変数変換を改良大森公式に適用すると，

$$\lambda(t) = \frac{\lambda_0^p\theta^{p-1}}{(\lambda_0\theta t+1)^p}, \tag{9}$$

$$\Lambda(t) = \frac{(\lambda_0\theta)^{p-1}}{\theta(p-1)}\left\{1-\left(\frac{1}{\lambda_0\theta t+1}\right)^{p-1}\right\} \tag{10}$$

となり，$p\to 1$ の極限において大森公式と一致するため，より一般的な表現であると言える．

4. 対数ポアソン実行時間モデルの確率論的解釈

前節の微分方程式に基づいた対数ポアソン実行時間モデルの導出は，NHPP の平均値関数（強度関数）に対する時間的なふるまいの観点から導出されたものである．これに対して本節では，対数ポアソン実行時間モデルに対して確率論的な解釈を与える．ソフトウェア信頼性評価モデルにおいて，NHPP に基づいたソフトウェア信頼性モデルを統一的に取扱う枠組みとして，一般化順序統計量モデルが提案されている [7]．一般化順序統計量モデルはソフトウェアフォールトの発生メカニズムをある母集団からの確率的抽出過程（デバッギング）として捉えることを可能としている．次のような仮定を設定する．

(仮定 A) ソフトウェアのテスト中にソフトウェア故障が発生した場合，原因となるフォールトは瞬間的に発見・除去される．

(仮定 B) プログラム中に含まれる初期フォールト数 N_0 は平均 ω (>0) のポアソン分布に従う．

(仮定 C) ソフトウェア故障は各々独立かつ時間に関してランダムに発生し，ソフトウェアフォールト 1 個当りのフォールト発見時刻は確率分布 $F(t)$，密度関数 $f(t), t\geq 0$ に従って分布する．

いま，$\{X_i,\ i=1,\ldots,N_0\}$ を各ソフトウェア故障発生時刻を表す確率変数列として定義する．初期時刻 $t=0$ においてソフトウェア内に n 個のフォールトが潜在している条件の下で，時刻 t までに m 個のフォールトが発見される確率は

$$\Pr\{N(t)=m|N_0=n\} \tag{11}$$

$$= \binom{n}{m}\{F(t)\}^m\{1-F(t)\}^{n-m} \tag{12}$$

となる．仮定 B より初期フォールト数はパラメータ ω のポアソン分布に従うため，時刻 t までに発見されたフォールト数の確率分布は

$$\Pr\{N(t)=m\} = \frac{\{\omega F(t)\}^m}{m!}\exp\{-\omega F(t)\} \tag{13}$$

によって表現される．これは平均値関数 $\omega F(t)$ をもつ NHPP と等価であり，ソフトウェアフォールト 1 個当たりの発見時間分布 $F(t)$ により種々のソフトウェア信頼度成長現象を表現することが可能である．

単一フォールト発見時間分布に第 2 種パレート分布

$$F(t) = 1-\frac{1}{(1+\lambda_0\theta t)^{p-1}} \tag{14}$$

を仮定すると，上述の議論により，フォールト発見事象は強度関数

$$\lambda(t) = \frac{\omega\lambda_0\theta(p-1)}{(1+\lambda_0\theta t)^p} \tag{15}$$

をもつ NHPP となる．いま，初期フォールト分布がパラメータ

$$\omega = \frac{(\lambda_0\theta)^{p-1}}{\theta(p-1)} \tag{16}$$

のポアソン分布に従うとすると，式 (15) の強度関数は式 (9) と一致する．よって $p\to 1$ の極限操作により対数ポアソン実行時間モデルに収束することがわかる．これより，対数ポアソン実行時間モデルは，非常に多数の初期フォールトを有するソフトウェアに対して，各フォールトの発見時刻がパレート分布に従って発見される現象を表現していることが理解できる．

参考文献

[1] J. D. Musa and K. Okumoto, "A logarithmic Poisson execution time model for software reliability measurement," *Proc. 7th Int'l Conf. Software Eng.*, 230–238, IEEE CS Press (1984).

[2] J. D. Musa, A. Iannino and K. Okumoto, *Software Reliability, Measurement, Prediction, Application*, McGraw-Hill, NY (1987).

[3] 尾形良彦, "地震学とその周辺の地球科学分野における統計モデルと統計的手法," 日本統計学会誌, **22** (3), 413–463 (1993).

[4] T. Utsu, "A statistical study on the occurrence of aftershocks," *Geophys. Mag.*, **30**, 521–605 (1961).

[5] F. Omori, "On the aftershocks of earthquakes," *J. Coll. Sci. Tokyo Imp. Univ.*, **7**, 111–200 (1894).

[6] Y. Ogata, "Likelihood analysis of point processes and its application to seismological data," *Bull. Int'l Statist. Inst.*, **50**, 943-961 (1983).

[7] N. Langberg and N. D. Singpurwalla, "Unification of some software reliability models," *SIAM J. Sci. Comput.*, **6**, 781-790 (1985).