

11-2 BPT モデルの共通のばらつきパラメータの推定とベイズ型予測

Inference of a common aperiodicity parameter and Bayesian forecast in the BPT model

統計数理研究所

The Institute of Statistical Mathematics

1. はじめに

地震調査研究推進本部（地震本部）による内陸主要活断層の長期評価において、Brownian Passage Time (BPT)分布更新過程のばらつきパラメータ α は4つの活断層の活動履歴から得た共通の最尤推定値 $\alpha = 0.24$ を全国一律に適用している¹⁾。しかし、一般に活動間隔数が十分でないデータによるばらつきパラメータの最尤推定は過小評価される傾向にあり²⁾、また推定に用いられた4つの活断層の活動履歴は現在までにその多くが修正されており、現在全国的に適用されている α の推定値 0.24 についてその信頼性が危ぶまれている。そこで、本報告では、地震本部が長期評価にて公表している内陸活断層の古地震カタログに基づき、共通のばらつきパラメータ α の推定値を与える。さらに、ばらつきパラメータ α の推定手法および地震予測手法の精度をシミュレーションにより比較検証する。

ある活断層における調査年代 $S \sim T$ 中に発生した地震活動年代の系列を $S < t_1 < \dots < t_n < T$ とおき、その活動間隔が平均 μ 、ばらつき α の BPT 分布に従うことを仮定したときの尤度を $L_{[S,T]}(\mu, \alpha | t_1, \dots, t_n)$ と表す。歴史地震を除いて、多くの地震活動年代は1時点に特定できず、地震痕跡周辺の堆積物の炭素年代などによる制約条件からしか推測することができない（第1図参照）。

そのような不確定な地震活動年代を、その確率密度関数 $\psi(t_1, \dots, t_n)$ により表現したとき、厳密な尤度は次のような不確定な地震活動年代に関する多重積分で表される³⁾。

$$L_{[S,T]}(\mu, \alpha | \psi) = \int_S^T \dots \int_S^T L_{[S,T]}(\mu, \alpha | t_1, \dots, t_n) \psi(t_1, \dots, t_n) dt_1 \dots dt_n \approx \frac{1}{N} \sum_{j=1}^N L_{[S,T]}(\mu, \alpha | t_1^{(j)}, \dots, t_n^{(j)}). \quad (1)$$

上式の尤度は数値積分あるいは確率密度関数 $\psi(t_1, \dots, t_n)$ からシミュレートした地震活動年代の乱数に基づくモンテカルロ積分により評価することができる⁴⁾。さらに、パラメータ μ, α について確率密度関数 $\pi(\mu, \alpha)$ をもつベイズ事前分布を与えたとき、ベイズ周辺尤度は $L_{[S,T]}(\pi | \psi) = \iint L_{[S,T]}(\mu, \alpha | \psi) \pi(\mu, \alpha) d\mu d\alpha$ と表される。このとき、将来期間 $T \sim T + \Delta$ における地震発生確率について、次式のベイズ予測⁴⁾ により評価することができる。

$$FB_{\psi}(\Delta | \pi) = \frac{\int_S^T \dots \int_S^T L_{[S,T]}(\mu, \alpha | t_1, \dots, t_n) \psi(t_1, \dots, t_n) \pi(\mu, \alpha) F_{T-t_n}(\Delta | \mu, \alpha) dt_1 \dots dt_n d\mu d\alpha}{L_{[S,T]}(\pi | \psi)} \\ \approx \frac{\sum_{j=1}^N L_{[S,T]}(\mu^{(j)}, \alpha^{(j)} | t_1^{(j)}, \dots, t_n^{(j)}) F_{T-t_n^{(j)}}(\Delta | \mu, \alpha)}{\sum_{j=1}^N L_{[S,T]}(\mu^{(j)}, \alpha^{(j)} | t_1^{(j)}, \dots, t_n^{(j)})}. \quad (2)$$

ただし、 $F_{T-t_n}(\Delta|\mu, \alpha)$ はパラメータ μ, α が与えられた下で、最新活動年代から $T-t_n$ 年経過時点での、将来 Δ 年間における地震発生確率である。上式のベイズ予測も数値積分あるいは確率密度関数 $\psi(t_1, \dots, t_n), \pi(\mu, \alpha)$ からの地震活動年代およびパラメータ μ, α の乱数に基づくモンテカルロ積分により評価される。

2. 内陸活断層カタログの解析結果

地震本部が長期評価にて公表している 2018 年 6 月 1 日時点での内陸活断層カタログのうち、少なくとも最新活動と 1 つ前の活動の活動時期が判明している 73 の活動区間の活動履歴を用いて、パラメータ μ, α の推定および将来の地震発生確率の評価を行った。以下に提案する 3 つの手法を検討する。ただし、いずれの手法もばらつきパラメータ α は全活動区間で共通の値をもつものとする。

手法 1（中点データ、MLE plug-in 予測）：活動時期を推定区間の中点で近似した尤度を用いた最尤法によりパラメータを推定し、最尤推定値を代入した BPT 分布を用いて plug-in 予測を行う。

手法 2（区間データ、MLE plug-in 予測）：(1)式の厳密な尤度を用いた最尤法によりパラメータを推定し、最尤推定値を代入した BPT 分布を用いて plug-in 予測を行う。

手法 3（区間データ、Bayes 予測）：パラメータを特定の値に定めずにベイズ事前分布 $\pi(\mu, \alpha)$ (μ は 10~1,000,000 年の対数上の一様分布、 α は 0~1 の一様分布) を仮定し、(2)式を用いてベイズ予測を行う。

特に、ばらつき α に対する尤度の変化を第 2 図に示した。日本のカタログは活動区間ごとの活動履歴が少ないため、第 2 図の緑線、赤線から求まる α の最尤推定値 $\alpha = 0.34, \alpha = 0.42$ は α を過小評価している可能性が高い。なお、昨年報告より対象の活動区間を増やしたことにより、昨年の報告とは異なる α の最尤推定値を得ている⁴⁾。一方で、黒線で示した α のベイズ事後分布は、最尤推定に伴う過小評価が生じないため、 $\alpha = 0.50$ はより真に近い値と考えられる。

3. シミュレーションによる検証

以上の手法 1~3 の尤度に基づくパラメータ推定およびベイズ予測の精度について、数値実験により比較検討を行う。実験では、活動間隔の平均 $\mu=1000$ (年) と変動係数 $\alpha=0.5$ を与えた仮定の活断層から、BPT 分布の乱数によりシミュレートした活動履歴に基づいて、パラメータの推定および将来 30 年地震確率の評価を行う。ただし、活動時期は正確にはわからず、平均 100 年間隔のポアソン過程の乱数によりシミュレートした堆積物年代で挟まれる活動時期の推定区間を生成する。その結果、たとえば第 3 図のように、2 つの活動履歴が 1 つの区間内で特定されるケースも現れる。さらに、パラメータ推定および予測の対象とする活断層のシナリオとして

シナリオ 1：3 つの活動履歴が特定された単独の活断層

シナリオ 2：100 の活動履歴が特定された単独の活断層

シナリオ 3：3 つの活動履歴が特定された 100 の活断層（共通の α を仮定）

の 3 つを用意した。シナリオ 2 は現実的な活動数ではないが、十分なデータの下で正しい推定と予

測が行われるか確かめることを目的としている。シナリオ3では、現行の地震本部による内陸主要活断層の長期評価手法と同様にばらつきパラメータ α を全活断層で共通と仮定した場合の推定および予測精度を評価する。なお、シナリオ1では100,000組、シナリオ2, 3では1,000組の活動履歴をシミュレートする。

以上の3つのシナリオに対して、手法1~3を用いて、パラメータ推定および予測を行った結果を表1~3に示した。シナリオ1では、手法1, 2の最尤法により推定された α の値はいずれも平均して相当の過小評価となっていることがわかる。また、将来30年確率(%)の予測誤差を、真のパラメータの下での将来30年確率との平均二乗誤差により評価すると、手法3のベイズ予測は他の手法に比べ大きく改善している。さらに、予測誤差として将来30年間の対数尤度を評価すると、手法1, 2において評価不能(-Inf)となった。これは、手法1, 2において将来30年確率を0と予測したにもかかわらず30年以内に地震が起こるケースがあったためであり、 α の過小評価による影響は相当に大きいとわかる。

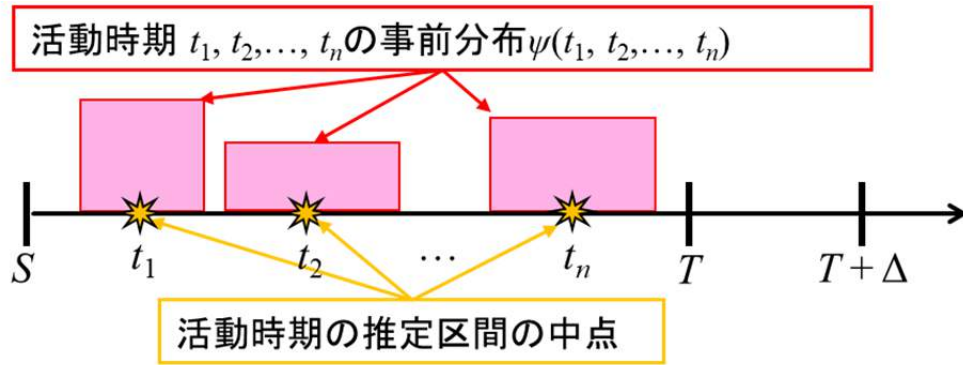
シナリオ2の活動履歴が十分にある状況では、手法2の厳密な尤度による最尤推定値は真の値の近くに推定できているが、手法1の活動時期を中点近似した尤度では α の最尤推定値が真の値より過大評価されていることがわかる。このシナリオにおいても、手法3のベイズ予測が予測精度で若干優れている。

シナリオ3は、活動履歴の少ない活断層を100集めることでパラメータの推定精度の改善を図るものであるが、手法1, 2の最尤法により推定された α の値は依然として過小評価されている。一方、手法3について、実際に予測に用いているわけではないが α のベイズ推定値を評価すると真の値に十分近い推定値を得ることができる。予測精度も手法3が最も優れている。以上の数値実験結果から、ばらつきパラメータ α の推定および予測において、手法3のベイズ予測が最も良い予測精度をもつことが確認された。

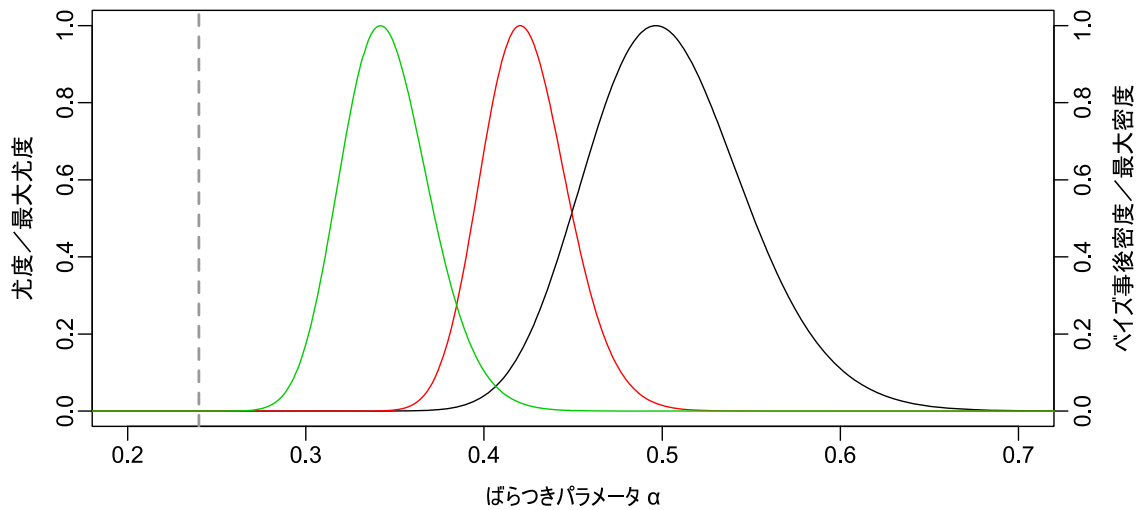
(野村俊一, 尾形良彦)

参考文献

- 1) 地震調査研究推進本部 (2001) 長期的な地震発生確率の評価手法について
- 2) Nomura, Ogata, Komaki and Toda (2011) *J. Geophys. Res.*, 116, B04315.
- 3) Ogata (1999) *J. Geophys. Res.*, 104, 17,995-18,014.
- 4) 統計数理研究所 (2017) 連絡会報, 第98巻 (11-5), 443-448.

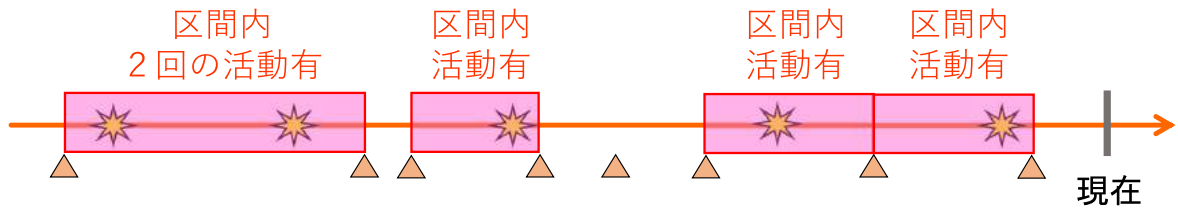


第 1 図 不確かな地震活動年代データと中点データ
 Fig. 1 Uncertain occurrence dates and their middle points



第 2 図 共通のばらつきパラメータ α を固定したときの平均活動間隔 μ に関する最大尤度 (赤：中点データ, 緑：区間データ). 黒線は, パラメータ μ, α に無情報事前分布を仮定したときの α の事後分布の確率密度関数である. 緑, 赤, 黒線はそれぞれ $\alpha = 0.34, 0.42, 0.50$ で最大となる. 縦点線は地震本部の推定値 $\alpha = 0.24$ を表す.

Fig. 2 Maximum likelihoods with respect to mean recurrence times μ with fixed common aperiodicity parameter α obtained from middle-point datasets (red) and interval datasets (green). Black curve shows Bayesian posterior density of α with non-informative priors for μ and α . Green, red and black curves take their maximum at 0.34, 0.42 and 0.50, respectively. The vertical dashed line shows the estimate $\alpha = 0.24$ given by the Headquarters for Earthquake Research Promotion.



★ : BPT分布更新過程による活動履歴のシミュレーション

▲ : ポアソン過程による堆積物のシミュレーション

第3図 仮想の活断層における活動年代の推定区間のシミュレーション
 Fig. 3 Simulation for estimated intervals of occurrence dates in virtual active faults

表1 シナリオ1 (3つの活動履歴が特定された単独の活断層) の数値実験結果
 Table 1 Results of simulation in Scenario 1 (sole active fault whose 3 historical events are specified)

比較手法	μ の推定値 の平均	α の推定値 の平均	予測誤差 (二乗誤差)	予測誤差 (対数尤度)
手法1 (中点 MLE)	1018.0	0.325	31.1	-Inf
手法2 (区間 MLE)	1019.9	0.278	65.3	-Inf
手法3 (区間 Bayes.)			8.3	-0.133

表2 シナリオ2 (100の活動履歴が特定された単独の活断層) の数値実験結果
 Table 2 Results of simulation in Scenario 2 (sole active fault whose 100 historical events are specified)

比較手法	μ の推定値 の平均	α の推定値 の平均	予測誤差 (二乗誤差)	予測誤差 (対数尤度)
手法1 (中点 MLE)	1001.4	0.524	0.423	-0.122
手法2 (区間 MLE)	1001.5	0.494	0.399	-0.121
手法3 (区間 Bayes)			0.395	-0.121

表3 シナリオ3 (3つの活動履歴が特定された共通のばらつきパラメータ α をもつ 100 の活断層) の数値実験結果

Table 3 Results of simulation in Scenario 3 (100 active faults with common aperiodicity parameter α whose 3 historical events are specified)

比較手法	μ の推定値 の平均	α の推定値 の平均	予測誤差 (二乗誤差)	予測誤差 (対数尤度)
手法1 (中点 MLE)	1017.1	0.399	7.9	-0.133
手法2 (区間 MLE)	1019.6	0.359	9.2	-0.136
手法3 (区間 Bayes.)		0.500※	5.8	-0.130

注) ベイズ事後分布 (尤度) における α の平均 (予測には直接使用されていない)